1-2024

Statistical discrimination and duration dependence in a semistructural model

Ismail BAYDUR
Singapore Management University, ismailb@smu.edu.sg

Jianhuan XU
Singapore Management University, jhxu@smu.edu.sg

Citation

BAYDUR, Ismail and XU, Jianhuan. Statistical discrimination and duration dependence in a semistructural model. (2024). *International Economic Review*. 65, (3), 1-30.
Available at: https://ink.library.smu.edu.sg/soe_research/2758

INTERNATIONAL ECONOMIC REVIEW
Vol. 00, No. 0, xxxx 2024 DOI: 10.1111/iere.12696

STATISTICAL DISCRIMINATION AND DURATION DEPENDENCE IN A SEMISTRUCTURAL MODEL*

BY ISMAIL BAYDUR AND JIANHUAN XU

Singapore Management University, Singapore

This article develops a job-search model with unobserved worker heterogeneity and learning about worker types from unemployment duration. The model features negative duration dependence that stems from unobserved heterogeneity, skill depreciation, and statistical discrimination. We estimate job-finding rates implied by our model using microlevel data from the Current Population Survey. We find that removing interview costs counterfactually, thereby eliminating statistical discrimination, substantially increases the job-finding rates of the long-term unemployed. The performance of low-skill workers at the interview stage with discriminating firms plays a key role in explaining our counterfactual result.

1. INTRODUCTION

The fraction of the unemployed who find a job within a month falls dramatically by unemployment duration. This feature of the U.S. labor market, which is referred to as negative duration dependence, has been studied by a large body of literature. More recently, Kroft et al. (2013) show in an experimental study with fictitious resumes that the callback rates to an interview decline with unemployment duration. This phenomenon, called statistical discrimination, could be an important source in explaining the large difference between short- and long-term job-finding rates.

In this article, our goal is to study statistical discrimination in a job-search model and quantify its contribution to the negative duration dependence for the U.S. economy. To this end, we develop a job-search model where workers randomly meet with vacancies at an exogenously given rate. However, matching does not guarantee hiring. Firms matched with a worker decides whether to go through an interview process, and hiring occurs only if the outcome of the interview is successful. The outcome of the interview is a random event. At the interview stage, the matched pair randomly draws match-specific productivity and the firm hires the worker only if this random draw is above a minimum productivity level required for a match to be viable.

There are two worker types, which we call high-skill and low-skill. A high-skill worker draws match-specific productivity from a distribution that first-order stochastically dominates that of a low-skill worker so that he is better at turning interviews into job offers. Worker types are not constant over time. High-skill workers lose their abilities and become low-skill at a constant rate as they remain in the unemployment pool; that is, they experience skill

*Manuscript received July 2022; revised December 2023.

We are grateful to the editor, Iourii Manovskii, for comments and guidance, as well as to three anonymous referees. We also thank Nicolas Jacquet, Thomas Sargent, Sephorah Mangin, Costas Meghir, Gueorgui Kambourov, Dirk Krueger, Jeremy Greenwood, Toshihiko Mukoyama, Wouter den Haan, and Russell Cooper for their comments and suggestions. We also thank seminar and conference participants at City University of Hong Kong, GRIPS, Kyushu University, National University of Singapore, Singapore Management University, Econometric Society conferences in Seoul and Auckland, and Midwest Macro Meetings in Michigan State University. Qiugu He provided valuable research assistance. This study is funded through a research grant (Fund No: C244/MSS18E001) from Singapore Management University. All errors are our own. Please address correspondence to: Ismail Baydur, Singapore Management University, 90 Stamford Road, Singapore, 178903. E-mail: *ismailb@smu.edu.sg*.

depreciation. Similarly, low-skill workers become high-skill at a constant rate while employed; that is, they experience on-the-job learning.

A key feature of our model is hiring firms do not observe a worker's true type, but they can infer the probability that the worker is high-skill from his most recent unemployment duration. We call this probability the *resume* of the worker. Firms update these resume values based on the differences between the job-finding rates of the high- and low-skill workers and interview applicants only if the prospects are favorable relative to the cost of the interview. There are two types of firms: discriminating and nondiscriminating. A discriminating firm draws a random interview cost and selectively interviews the applicants, whereas a nondiscriminating firm interviews every applicant at zero cost.[1] Because more low-skill workers are present at longer durations, discriminating firms are less likely to interview the long-term unemployed, which leads to statistical discrimination in our model.

To measure the importance of statistical discrimination, we consider the counterfactual job-finding rates after removing the interview costs so that every firm indiscriminately interviews every applicant. We calculate the percentage increase in the long-term job-finding rates relative to the baseline equilibrium with statistical discrimination and use this metric to assess the importance of statistical discrimination. We show this metric is the product of two ratios: (i) the share of discriminating firms relative to that of nondiscriminating firms, and (ii) the interview success probability of a low-skill worker with a discriminating firm relative to that with a nondiscriminating firm. Intuitively, discriminating firms very rarely interview the long-term unemployed workers in the baseline equilibrium, because these workers are mostly low-skilled. The first ratio determines the percentage increase in the probability of getting an interview for a long-term unemployed worker after removing the interview costs. However, the overall increase in the long-term job-finding rates also depends on how the low-skill workers perform at these interviews with the discriminating firms, which is not possible in the baseline equilibrium due to statistical discrimination. The second ratio captures this second effect. For statistical discrimination to have a big impact on the long-term job-finding rates, our metric requires both a large share of discriminating firms and a high success probability for the low-skill workers with these firms.

Our model delivers a parametric hazard model for job-finding rates. We obtain a reduced-form representation of these job-finding rates and estimate its parameters via maximum likelihood using individual-level data from the Current Population Survey (CPS). An advantage of our reduced-form approach is that one of the reduced-form parameters corresponds to the metric we defined in the previous paragraph to measure the importance of statistical discrimination. This feature allows us to estimate the impact of removing interview costs without having to identify the underlying structural model parameters. We provide a formal discussion about identification of this parameter. Our key insight is that the higher the job-finding rate of a medium-term unemployed worker relative to that of a long-term unemployed worker, the higher the value of this parameter.

Intuitively, both the medium- and long-term unemployed workers are predominantly low-skill workers, and in either case, most of the hired workers are low-skill workers.[2] Therefore, the difference between their job-finding rates is largely driven by the type of the hiring firm. Nondiscriminating firms interview everyone, but a medium-term unemployed worker has a greater chance of being interviewed by a discriminating firm, because this group has relatively more high-skill workers. If a large difference exists between the medium- and long-term job-finding rates, a sizable amount of discriminating firms must be present. This requirement alone is not sufficient, because the job-finding rates also depend on the outcome of these interviews. Recalling that medium-term unemployed workers are mostly low-skill workers, a

[1] In the data, the job-finding rates approach a value that is significantly greater than zero. We capture this feature of the data in our model with the nondiscriminating firms.

[2] As a point of reference, we consider six months as medium term. Based on our estimates, the share of high-skill workers at this unemployment duration is only about 10%. The longest reported unemployment duration is 24 months (104 weeks) in our sample.

large difference also requires that low-skill workers have a good chance of forming a productive match with discriminating firms. Noting the reduced-form parameter governing the importance of statistical discrimination also depends on these two criteria, we conclude that the difference between the medium- and the long-term job-finding rates is informative about the importance of statistical discrimination in accounting for the declining job-finding rates by unemployment duration.

The reduced-form parameter estimate for the impact of statistical discrimination is 1.716, which indicates that the long-term job-finding rates increase by 171.6% in the absence of interview costs. This estimate, which implies almost a threefold increase, is statistically significant at conventional levels and robust to various sample restrictions. Consistent with our intuition above, the difference between medium- and long-term job-finding rates is sizable. At the estimated parameter values and after removing the effects due to observed heterogeneities, the job-finding rate at six months of unemployment duration is about 40% higher than that of 24 months of unemployment duration.

To understand the model mechanism, we recover the structural model parameters from the reduced-form parameter estimates. In the baseline equilibrium, the average job-finding rates monotonically decrease from 0.506 to 0.121 over the course of 24 months, implying about a 40-percentage-point difference between a recently unemployed and a long-term unemployed worker. When discriminating firms can also freely interview every applicant, we find the job-finding rate of a long-term unemployed worker increases by about 20 percentage points to 0.326, whereas the job-finding rate of a recently unemployed worker increases marginally to 0.515. We conclude that statistical discrimination alone accounts for half of the original difference in job-finding rates. The interview success probability of the low-skill workers with the discriminating firms plays an important role in driving this result. Our estimates imply this probability is roughly equal to half of the interview success probability of a high-skill worker with a discriminating firm, but it is still significantly greater than zero.

In a recent paper, Jarosch and Pilossoph (2019) similarly study the effects of statistical discrimination on job-finding rates of long-term unemployed workers. Contrary to our findings, they argue that eliminating statistical discrimination does not have any significant effect on the job-finding rates of long-term unemployed workers. Motivated by the findings in Kroft et al. (2013), they also work with a structural model in which firms make explicit interview decisions. They assume that firms have minimum hiring requirements and are heterogeneous with respect to these requirements. Interview probability in their model falls with unemployment duration because workers at the bottom of the skill set are less likely to satisfy these requirements. Therefore, these workers remain in the unemployment pool longer. Consequently, only a small fraction of the firms actually find value in interviewing the long-term unemployed.

The implied interview success probabilities take a stark form in Jarosch and Pilossoph (2019), particularly for low-skill workers who are more likely to be long-term unemployed. They fail their interviews with most firms because of their strict minimum skill requirements, and regardless of their unemployment duration, they are hired predominantly by the firms that have no skill requirement and indiscriminately interview everyone. When they remove the interview cost, low-skill workers see a large increase in their interview rates but fail to satisfy the skill requirement for most of these additional interviews. This specification is different from ours, where low-skill workers can convert these additional interviews to actual job offers with a probability that is not necessarily equal to 0.

Our model can be viewed as introducing randomness at the interview stage in addition to the skill requirements imposed by the production function. Our approach is to parameterize this process in a way that allows for the specification in Jarosch and Pilossoph (2019), and estimate it directly from the data. We are able to generate their main result by making the interview success probability of low-skill workers with discriminating firms small, but the data do not support this parameter restriction. Consequently, we reach a very different conclusion regarding the role of statistical discrimination. We illustrate our point more formally by writing the job-finding rates in our model in a more general form which also nests the job-finding rates in Jarosch and Pilossoph (2019).

Our modeling approach for skill heterogeneity in forming a productive match and learning from employment histories is based on the theoretical work of Gonzalez and Shi (2010). Building on their framework, Doppelt (2016) and Fernandez-Blanco and Preugschat (2018) stress the importance of unobserved heterogeneity and learning in accounting for negative duration dependence in an equilibrium model. Doppelt (2016) infers the role of statistical discrimination by comparing the job-finding rates in a benchmark equilibrium with incomplete information and learning about worker types to that in an equilibrium with full information. Fernandez-Blanco and Preugschat (2018) compare job-finding rates in a nonsequential search framework with the optimal ranking of applicants based on their unemployment duration with a counterfactual setting in which firms are forced to test every applicant for suitability for the job. Despite methodological differences, both papers find that statistical discrimination plays an important role in generating heterogeneity in job-finding rates.[3]

Alvarez et al. (forthcoming) decompose duration dependence in an optimal stopping time model. Duration dependence at the individual level is captured by two distinct parameters. Unlike our model, which contains only two worker types, many distinct worker types can exist in their model, and these types can be nonparametrically estimated provided that at least two spells of unemployment of the same individual are observed. Using microlevel data from Austria, they show unobserved heterogeneity is a critical source in explaining duration dependence. However, their model does not feature statistical discrimination.

Our article is also related to a recent literature that studies the role of heterogeneity in estimating an aggregate matching function. Using microlevel CPS data, Barnichon and Figura (2015) and Hall and Schulhofer-Wohl (2018) find unemployment duration is an important determinant of job-finding rates. However, both papers treat unemployment duration as an explanatory variable. We address a more fundamental question in this article related to sources of negative duration dependence.

The rest of the article is organized as follows: The next section describes our model. We then describe our data set and estimation procedure. In Sections 4 and 5, we study the equilibrium properties of our model and perform our counterfactual analysis, respectively. The last section concludes.

2. MODEL

2.1. *Environment.* The economy is populated by infinitely lived risk-neutral workers with a unit measure and a large number of risk-neutral firms. Time is continuous and the discount rate is $\rho > 0$. In what follows, we focus on a stationary equilibrium.

At any instant, a worker can either be employed or unemployed. An employed worker engages in production and becomes unemployed at a constant rate $\delta > 0$. While employed, he cannot search for a job. An unemployed worker searches for a job in a frictional labor market where job opportunities arrive at an exogenously given rate η. To be hired by a firm, an unemployed worker needs to not only meet a vacancy but also go through an interview process with the firm. Hiring occurs only if the outcome of the interview is successful in that the matched worker–firm pair generates a positive surplus. Output in a match depends on match-specific productivity, y. This match-specific productivity is randomly drawn at the interview stage and remains constant thereafter. Firms have a minimum productivity requirement in that the match-specific productivity draw must be sufficiently large for the worker to be hired.

Workers are heterogeneous with respect to their skills. Each worker can be one of the two possible types, H and L, which we call high-skill and low-skill workers, respectively. A

[3] In a related paper, Kospentaris (2021) finds that unobserved heterogeneity is more important than skill loss in accounting for duration dependence although he does not provide a separate measure for the role of statistical discrimination.

high-skill worker is not only more productive but also draws match-specific productivity from a distribution that first-order stochastically dominates that of a low-skill worker. The latter of these assumptions implies a high-skill worker is more successful at the interview stage than a low-skill worker. A worker's skill type may change over time. While unemployed, a high-skill worker loses his skills at a constant rate $\gamma \geq 0$ (e.g., skill depreciation), and becomes low-skill. A low-skill worker, on the other hand, becomes high-skill while employed (e.g., on-the-job learning) at a constant rate $\sigma \geq 0$.

Hiring firms cannot observe the *true* skill type of an unemployed worker, but only his current unemployment duration, τ. Let $r(\tau)$ denote the proportion of high-skill workers at unemployment duration τ. We call $r(\tau)$ the resume of the worker. The resume values change with unemployment duration because workers are heterogeneous with respect to their skills. Firms factor this information into their interview decisions, where they compare the expected value from interviewing a candidate with its cost.

Firms are heterogeneous in terms of their interview technology. There are two types of firms. α fraction of the firms face interview costs and may decide not to interview the applicant, whereas the remaining $(1-\alpha)$ fraction of the firms can costlessly interview every applicant. By construction, only the former type of firms statistically discriminate against unemployed workers at the interview stage. Therefore, we refer to these firms as discriminating and nondiscriminating firms and denote them with D and N, respectively. We also allow the match-specific productivity distributions and the minimum productivity requirements to possibly depend on the firm type.

2.2. *Match Surplus.* We assume the following "up-to-task" production function for $s = \{H, L\}$ and $j = \{D, N\}$:

$$f_j^s(y) = \begin{cases} x^s + y, & \text{if } y \geq \bar{y}_j \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where $\bar{y}_j$ is an exogenously given firm-specific minimum productivity requirement for a match to be viable and $x^H > x^L$ are skill-specific productivity levels. We further assume $\bar{y}_j + x^s$ is sufficiently larger than the flow value of unemployment, b, for all the possible combinations of s and j so that a match is formed as long as $y \geq \bar{y}_j$.

We allow the distribution of y to depend on both firm type and worker type. Let $\psi_j^s(y)$ denote the cumulative density function of the match-specific productivity distribution for worker s and firm j. First-order stochastic dominance implies that $\psi_j^H(y) \leq \psi_j^L(y)$ with strict inequality for some y. In particular, we assume $\psi_j^H(\bar{y}_j) < \psi_j^L(\bar{y}_j)$ for at least one of $j = \{D, N\}$, so that a high-skill worker is relatively more successful at the interview stage with at least one of the firm types. In what follows, we refer to this conditional probability as the interview success probability and denote it as $p_j^s = 1 - \psi_j^s(\bar{y}_j)$.[4]

Because labor markets are frictional, a firm–worker match creates a surplus. We assume the firm gets the entire surplus value (e.g., workers have no bargaining power). Under this assumption, the wage rate is equal to b. We further normalize the value of an open position to 0.[5] The value of a filled position at a firm of type j with a high-skill worker, $J_j^H(y)$, and with a low-skill worker, $J_j^L(y)$, satisfies the following Bellman equations, respectively:

$$\rho J_j^H(y) = y + x^H - b - \delta J_j^H(y), \tag{2}$$

$$\rho J_j^L(y) = y + x^L - b - \delta J_j^L(y) + \sigma\left(J_j^H(y) - J_j^L(y)\right) \tag{3}$$

[4] Our specification for forming a productive match is similar to Gonzalez and Shi (2010) except that the output is higher in expectation for a high-skill worker conditional on firm type.

[5] This assumption is consistent with the free-entry condition in the standard Diamond–Mortensen–Pissarides matching model.

for $y \geq \bar{y}_j$. Intuitively, a filled position is an asset for the firm, and its discounted value, $\rho J_j^s(y)$, is equal to the return on this asset given by the terms on the right-hand side. A filled position generates a flow profit of $(y + x^s - b)$, and the job becomes vacant at rate δ, in which case the firm loses $J_j^s(y)$. In the case of a low-skill worker, an additional term reflects gains from becoming a high-skill worker while employed for the same firm. After rearranging Equations (2) and (3), we obtain:

$$J_j^H(y) = \begin{cases} \dfrac{y + x^H - b}{\rho + \delta}, & \text{if } y \geq \bar{y}_j \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

and

$$J_j^L(y) = \begin{cases} \dfrac{(\rho + \delta)(y + x^L - b) + \sigma(y + x^H - b)}{(\rho + \delta)(\rho + \delta + \sigma)}, & \text{if } y \geq \bar{y}_j \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The match surplus not only directly depends on the worker's skill type via his productivity at a given match, but also indirectly via the distributional assumptions on y. We explain how the match surplus affects interview decisions of D-type firms in the next section.

2.3. *Interview Decisions and Job-Finding Rates.* Conditional on matching with a worker, N-type firms interview every applicant at no cost, whereas D-type firms draw a random interview cost. The interview takes place only if the D-type firm pays the interview cost. For both types of firms, hiring occurs if the outcome of the interview is successful. Otherwise, both the firm and the worker continue to search for alternatives. If the interview cost is sufficiently high, a D-type firm may opt not to interview the applicant at all.

Specifically, each D-type firm matched with a worker randomly draws a cost for interviewing the applicant, κ, from a fixed distribution with a cumulative distribution function $C(\kappa)$. At this stage, the firm observes the resume value of the applicant. Without knowing his true skill type, the firm decides whether to pay the cost and interview the applicant or continue searching for alternatives. If the firm eventually decides to interview the applicant, the worker–firm pair learns the productivity of their match. For a given firm type j, a high-skill worker draws $y \geq \bar{y}_j$ with probability p_j^H, in which case an employment relationship is formed and the firm–worker pair starts producing. Otherwise, with probability $1 - p_j^H$, hiring does not occur. A low-skill worker draws $y \geq \bar{y}_j$ with probability p_j^L. Our assumption regarding first-order stochastic dominance ensures that $p_j^L < p_j^H$ for at least one firm type. Furthermore, the skill-specific average match surplus conditional on hiring, Q_j^s, is given by:

$$Q_j^s = \int J_j^s(y) d\psi_j^s(y | y \geq \bar{y}_j). \tag{6}$$

When making a decision about interviewing an applicant with a resume value r, a D-type firm compares the cost of the interview with its *expected* value, $W_D(r)$, which is equal to:

$$W_D(r) = r p_D^H Q_D^H + (1 - r) p_D^L Q_D^L. \tag{7}$$

We consider productivity distributions that imply $Q_j^L < Q_j^H < \infty$ so that both values are finite and high-skill workers on average are employed in higher surplus matches. Accordingly, the expected value of the interview is linear and monotonically increasing in r and the interviewing decision has a cutoff property in that the firm interviews the applicant only if $\kappa < W_D(r)$. For a worker with resume value r, the probability of being interviewed, conditional on a match, is equal to $C(W_D(r))$, which is also increasing r. In other words, conditional on a match, workers with high resume values are more likely to be interviewed by D-type firms.[6]

We further assume the interview cost comes from a uniform distribution between $W_D(0)$ and $W_D(\hat{r})$ for some parameter $\hat{r} \geq r(0)$, so the interview probability with a D-type firm is equal to $r(\tau)/\hat{r}$. Overall, a worker with resume value $r(\tau)$ expects to be interviewed with probability $(1-\alpha+\alpha r(\tau)/\hat{r})$. We can write the job-finding rate for an unemployed worker at duration τ conditional on his true skill type, $s = \{H, L\}$, as follows:

$$\lambda^s(\tau) = \eta[(1-\alpha)p_N^s + \alpha r(\tau)p_D^s/\hat{r}]. \tag{8}$$

Because the interview probability is increasing with the resume value, workers with high resume values find jobs faster for a given skill type. Moreover, for a given resume value of r, high-skill workers can find jobs faster because they are more likely to form a productive match at the interview stage.[7]

2.4. *Resume Updating.* The resume values correspond to the proportion of high-skill workers at unemployment duration τ and help firms make an inference about the true skill type of an unemployed worker from his unemployment duration. Bayesian updating implies the following updating rule for the resume values at duration τ:

$$\frac{dr(\tau)}{d\tau} = -\gamma r(\tau) - (1-r(\tau))r(\tau)\big(\lambda^H(\tau) - \lambda^L(\tau)\big). \tag{9}$$

We obtain this expression as a limiting case of the Bayes rule in a discrete-time approximation and defer the details to Online Appendix A.

An important feature of the solution to Equation (9) is that $r(\tau)$ is strictly decreasing in τ and approaches 0 in the limit. To see this point, note that, by definition, the resume values cannot be negative or be greater than 1. Suppose further that either $r(0) < 1$ or $\gamma > 0$ (or both). First-order stochastic dominance implies that $(\lambda^H(\tau) - \lambda^L(\tau)) > 0$ for any $0 < r(\tau) \leq r(0)$. Then, from Equation (9), $dr(\tau)/d\tau$ is negative for any $0 < r(\tau) \leq r(0)$ and it is equal to 0 if $r(\tau) = 0$. Hence, $r(\tau)$ monotonically decreases from $r(0)$ to 0.[8] Intuitively, the high-skill workers leave the unemployment pool faster, and the unemployment pool is eventually populated mostly by low-skill workers. We refer to this fact further below when we analyze counterfactual job-finding rates and discuss identification.

We also show in Online Appendix A that Equation (9) has a solution of the form $r = G(\tau)$, in which we have a closed-form solution for the *inverse* of the function $G(\tau)$. This equation becomes very handy in the estimation part. For example, we are able to address a potential time-aggregation bias in our estimation without resorting to alternative methods such as solving a discrete-time approximation of our model.

2.5. *Equilibrium.* Because the worker types are not observed, it is convenient to define $\lambda^r(\tau)$ as the *average* job-finding rate at duration τ as follows:

$$\lambda^r(\tau) = r(\tau)\lambda^H(\tau) + (1-r(\tau))\lambda^L(\tau). \tag{10}$$

[6] As far as the expected value from interviewing a candidate is concerned, one could allow $Q_D^H = Q_D^L$: for example, $x^H = x^L$ and $\psi_D^s(y)$'s are Pareto with the same scale value that is greater than 1 and lower bounds are such that $\underline{y}_D^L < \underline{y}_D^H < \bar{y}_D$. The expected value from interviewing the candidate would still be linear and increasing in r.

[7] If we further assume that $\psi_j^s(y)$ and $\bar{y}_j$ are the same for $j = \{D, N\}$, Equation (8) reduces to a proportional hazard model with random frailty, which has been extensively studied in the literature.

[8] Noting that $\lambda^s(\tau)$ in Equation (8) are linear in $r(\tau)$, Equation (9) is a cubic ordinary differential equation, that is, an Abel equation of the first kind with constant coefficients. Hence, there are two other $r(\tau)$ values that satisfy $dr(\tau)/d\tau = 0$. However, one can show that one of them is greater than 1 and the other one is negative. Therefore, neither can be a valid resume value.

Let $\Gamma(\tau)$ denote the measure of the unemployed workers at duration $\tau \geq 0$. We can write the change in $\Gamma(\tau)$ as follows:

$$\frac{d\Gamma(\tau)}{d\tau} = -\lambda^r(\tau)\Gamma(\tau). \tag{11}$$

Equation (11) states that the measure of the unemployed decreases with unemployment duration at a rate given by the average job-finding rate at that duration. This average job-finding rate depends on the share of the high- and low-skill workers given by the resume value at that duration. Let us also define the unemployment rate, u, as the integral over all unemployment durations, that is, $u = \int_0^\infty \Gamma(\tau)d\tau$. We also note the measure of unemployed at $\tau = 0$ is given by the flow into the unemployment pool: $\delta(1-u)$. With this initial condition, we can solve the distribution of workers by unemployment duration as

$$\Gamma(\tau) = \delta(1-u)\exp\left(-\int_0^\tau \lambda^r(\omega)d\omega\right). \tag{12}$$

In the special case in which $\lambda^r(\tau)$ is constant, say, $\bar{\lambda}^r$, Equation (12) and the definition of u imply $\bar{\lambda}^r u = \delta(1-u)$. Rearranging for u yields an expression for the equilibrium unemployment rate that is standard in job-search models.

Finally, the calculation of $\Gamma(\tau)$ depends on the initial resume value, $r(0)$, which is equal to the proportion of high-skill workers among employed workers. Its value is determined endogenously in equilibrium and satisfies the following flow equation:

$$\int_0^\infty \lambda^H(\tau)r(\tau)\Gamma(\tau)d\tau + \sigma(1-r(0))(1-u) = r(0)\delta(1-u). \tag{13}$$

The first term on the left-hand side is the flow of high-skill workers from the unemployment pool, and the second term corresponds to the low-skill employed workers who become high-skill through on-the-job learning. The right-hand side is the flow of the high-skill workers who exogenously separate into the unemployment pool.

Given these equations, we can define a stationary equilibrium as follows:

Definition: A stationary equilibrium is a set of job-finding rates, $\lambda^s(\tau)$, resume values, $r(\tau)$ and $r(0)$, and a duration distribution of unemployed workers $\Gamma(\tau)$ such that:

1. $\lambda^s(\tau)$ are given in Equation (8);
2. $r(\tau)$ is updated according to Equation (9), where $r(0)$ satisfies Equation (13); and
3. $\Gamma(\tau)$ satisfies Equation (12), where $u = \int_0^\infty \Gamma(\tau)d\tau$.

2.6. *Long-Term Job-Finding Rates.* We are eventually interested in calculating counterfactual job-finding rates after removing the interview costs for all firms to gauge the effect of statistical discrimination on long-term job-finding rates. We close this section by deriving an expression for the increase in long-term job-finding rates after removing the interview costs and discussing the role of some of the structural parameters.

Recall that the proportion of high-skill workers declines with unemployment duration and approaches 0 in the limit. Therefore, a long-term unemployed worker with a resume value close to 0 is very rarely interviewed by D-type firms. Consequently, the job-finding rates monotonically decrease by unemployment duration to $\eta(1-\alpha)p_N^L$ as the resume values approach 0. By contrast, when the interview cost is removed, the job-finding rate approaches $\eta((1-\alpha)p_N^L + \alpha p_D^L)$, because D-type firms also interview every applicant regardless of their resume value.

We can express the percentage increase in the job-finding rates of the long-term unemployed workers after removing the interview cost as follows:

$$\zeta = \frac{\alpha}{1-\alpha}\frac{p_D^L}{p_N^L}. \tag{14}$$

This quantity, denoted by ζ, is a direct measure of the impact of statistical discrimination on long-term unemployed workers. The first term, $\alpha/(1-\alpha)$, captures the increase in the number of interviews for long-term unemployed workers. Holding everything else constant, a larger share of discriminating firms would be associated with a larger increase in the number of interviews after removing the interview cost.

However, the overall increase in the long-term job-finding rates also depends on how the low-skill workers perform in these interviews. This effect is captured by the second term, p_D^L/p_N^L. In the next section, we directly estimate ζ and test whether it is significantly different from 0.

We also note our model can reproduce the main finding in Jarosch and Pilossoph (2019). Setting α to a value that is significantly greater than zero and p_D^L to a value that is close to zero yields a large increase in the number of interviews but virtually no change in long-term job-finding rates after removing the interview costs. We compare and contrast the two models in greater detail in Section 5.

3. ESTIMATION

In this section, we discuss how we estimate our model using the observed job-finding rates in the data. We first describe the data set. Then, we derive a reduced-form representation of the job-finding rates in our model and discuss the identification of its parameters. A key result is that one of the reduced-form parameters corresponds to ζ in Equation (14), and we are able to directly estimate the increase in long-term job-finding rates after removing the interview costs.

One of the important aspects of the equilibrium in our model is that calculating the equilibrium requires the job-finding rates, $\lambda^s(\tau)$, which, from Equation (8), depend on interview success probabilities, p_j^s. These probabilities ultimately reflect our assumptions on the production function and the productivity distributions. However, one can still calculate the equilibrium in our model for a given set of values for p_j^s *without* the full information on these underlying assumptions. Therefore, we use these interview success probabilities in deriving the reduced-form representation of the job-finding rates in our model. One advantage of this approach is that we avoid making specific functional-form assumptions about the productivity distributions.[9]

3.1. *Data Description and Sample Restrictions.* We use microlevel data from the CPS obtained from the Integrated Public Use Microdata Series (IPUMS-CPS) database (Flood et al., 2020). CPS uses rotation groups in which each individual is interviewed for four consecutive months, not interviewed for eight months, and then reinterviewed for another four months before leaving the sample. The rotating panel design of CPS allows us to observe the employment status of each individual next month. In addition, unemployed individuals are asked to report their duration of unemployment, which allows us to calculate resume values throughout their current unemployment spell.

We perform a thorough consistency check for the reported unemployment duration in consecutive months and removed inconsistent observations, for example, more than a five-week

[9] We only maintain that high-skill workers draw from first-order stochastically dominant productivity distributions so that they are more successful at the interview stage for at least one type of firm, that is, either $p_N^H > p_N^L$ or $p_D^H > p_D^L$ or both.

increase in the unemployment duration in the next month. Moreover, some unemployed individuals in our sample move out of the labor force for a month before coming back to the unemployment pool in the next month. For most of these unemployed individuals, this period out of the labor force is added in the original data to the unemployment duration reported in the previous month. We treat the middle month as a classification error and change the individual's employment status in this month to unemployed. We drop the remaining individuals for whom the unemployment spell is reset to a smaller value.[10] We also remove individuals who move from the out-of-labor force to the unemployment pool for a month and then move out of the labor force again. By doing so, we focus on individuals who are attached to the labor force.[11]

We also collect information about age, gender, education, region, occupation, industry, class of worker, race, marital status, and reason for unemployment. The CPS monthly files contain some inconsistent information. Therefore, we drop the individuals whose gender or race changes across different surveys. We also drop individuals when their age decreases or when their age increases by two or more years.

A worker's age changes while unemployed, but the reference month for the reported age is the age during the interview and not the age of the worker when he entered the unemployment pool. For example, if the worker remains in the unemployment pool for two years, his age is substantially different from his age on entry to the unemployment pool. To address this concern, we randomly assigned a month component to the reported age in years and we indicate each individual's age at the time of the interview in months. We restrict our sample to those individuals who were at least 24 years old when their unemployment spell started. Consequently, we first normalized the age variable so that an individual who is 24 years old and is randomly assigned to January is one month old. Then, we excluded unemployment spells if an individual's adjusted age is negative when he started his current unemployment spell. We also dropped individuals whose adjusted age is greater than 384 months, or 55 years, because interview and hiring decisions might be substantially different for those near retirement (e.g., due to short expected tenure at the time of the hiring).

We create four occupation categories: (a) professional and managerial, (b) personal services, (c) sales and office, and (d) production. The industry variable indicates whether the individual is in the goods or services producing sector. We define three education categories: high school and less than high school, some college, and at least a four-year college degree. The marital-status variable indicates whether the respondent is married, divorced or widowed, or never married and single. We excluded family workers and self-employed individuals from our sample.

IPUMS-CPS provides six categories for the reason for unemployment: (a) job loser or on layoff, (b) other job losers, (c) end of temporary job, (d) job leaver, (e) new entrant, and (f) reentrant. Fujita and Moscarini (2017) report that many temporary-layoff workers are called back by their previous employer, and their duration dependence is different from that of other workers. Therefore, we exclude the first category from our main analysis.[12] Moreover, new entrants do not report an occupation or industry. Consequently, we do not have any observations in this category.

Apart from the microlevel CPS data, we obtained state-level vacancy data from the Conference Board's Help Wanted Online (HWOL) database. These data are only available from May 2005 to October 2018. Therefore, our main analysis focuses on this time period. We

[10] For these individuals, whether they take a short-term job that is not recorded or their employment status has a classification error and their subsequent unemployment duration has a measurement error is unclear.

[11] See Elsby et al. (2015) for their "deNUNification" methodology to adjust for classification error to measure gross labor market flows.

[12] In Online Appendix C, we perform two robustness analyses for this variable. In the first one, we add the temporary layoff category back to our sample. In the second one, we further drop job leavers from the sample, which may include individuals who have quit into unemployment and think they will be able to quickly find a job. Neither of these sample restrictions affects our main conclusion.

combined these series with the state-level unemployment rate from the Bureau of Labor Statistics to create time series for market tightness.

We note market tightness enters the estimation as a time-varying regressor. We also include monthly seasonal dummies in our regression. Finally, CPS uses a highly stratified sampling scheme. To have a representative sample of the population, we use the longitudinal sampling weights for the two consecutive months available in the IPUMS-CPS database. Even though we link individuals across different samples, their contribution to the likelihood function is through this sampling weight.

3.2. *Estimation and Identification.* We estimate the model parameters by maximizing a likelihood function based on the job-finding rates in Equations (8) and (9), and observed individual characteristics. In this subsection, we describe our estimation framework and provide a detailed discussion about identification.

Parameters assigned via prior information: Although $r(0)$ is determined endogenously in equilibrium, we take its value as given when we estimate our model. Later, we calibrate the value of σ, which does not directly affect the job-finding rates, to match the value of $r(0)$ from Equation (13). As in Jarosch and Pilossoph (2019), we also assume that $r(0) = \hat{r}$ so that a recently unemployed worker is always interviewed. We directly assign the value of $r(0)$ to be 0.75. Gregory et al. (2021) identify three types of workers based on employment transition rates in the United States. Of these types, the "gamma" type, which refers to the group of workers with the lowest job-finding rate, makes up 25% of the recently unemployed workers. Our benchmark value for $r(0)$ is consistent with their calculations. To ensure that the specific value of $r(0)$ does not affect our results, we estimate the job-finding rates in our model for a wide range of values for $r(0)$ in Online Appendix C and confirm that our main conclusion is robust to the choice of $r(0)$.

Job-finding rates in reduced form: Let us collect the remaining structural parameters into a few reduced-form parameters and rewrite the job-finding rates for $s = \{H, L\}$ as follows:

$$\lambda^H(\tau) = \tilde{\eta}[1 + \phi + (\zeta + \xi)r(\tau)/\hat{r}], \tag{15}$$

$$\lambda^L(\tau) = \tilde{\eta}[1 + \zeta r(\tau)/\hat{r}], \tag{16}$$

where we define four reduced-form parameters: $\tilde{\eta}$, ϕ, ξ, and ζ. They are related to the structural parameters as follows:

$$\tilde{\eta} = \eta(1 - \alpha)p_N^L, \tag{17}$$

$$\phi = \frac{p_N^H}{p_N^L} - 1, \tag{18}$$

$$\xi = \frac{\alpha}{1 - \alpha}\frac{p_D^H - p_D^L}{p_N^L}, \tag{19}$$

and ζ is given in Equation (14). Similarly, we can rewrite the resume-updating rule as follows:

$$\frac{dr(\tau)}{d\tau} = -\gamma r(\tau) - \tilde{\eta}(1 - r(\tau))r(\tau)(\phi + \xi r(\tau)/\hat{r}). \tag{20}$$

Because worker types are unobserved in the data, we construct the likelihood function using the average job-finding rates defined as follows:

$$\lambda^r(\tau) = r(\tau)\lambda^H(\tau) + (1 - r(\tau))\lambda^L(\tau)$$
$$= \tilde{\eta}\{1 + (\phi + \zeta/\hat{r})r(\tau) + \xi[r(\tau)]^2/\hat{r}\}. \tag{21}$$

Finally, job-finding rates vastly vary by observed individual characteristics. To account for these differences, we assume that labor markets are segmented as in Barnichon and Figura (2015), and search efficiency in a market segment depends on observable characteristics that make the worker more or less likely to make a contact with a firm. Formally, let x_i denote a vector of explanatory variables for individual i. We adopt a standard assumption in the literature and assume that job opportunities arrival rate is log-linear in x_i. Because η and $\tilde{\eta}$ are proportional to each other in Equation (17), we have the following specification:

$$\tilde{\eta}_i = \exp(x_i\beta), \tag{22}$$

where β is a vector of coefficients including a constant.[13]

Equations (20), (21), and (22) form the basis for our estimation which depend on the coefficient vector β, the reduced-form parameters ϕ, ξ, and ζ, as well as the structural parameter γ. The resume values are complex functions of observable characteristics and these parameters. Their complexity poses a challenge especially for score-vector calculations. We implement a chain-rule based method in the vein of automatic differentiation techniques, which are relatively new in the literature. The details of our estimation procedure are available in Online Appendix B.

Our main parameter of interest is ζ, which measures the increase in long-term job-finding rates after removing the interview cost. An attractive feature of our estimation is that we are able to obtain a direct estimate for this measure without pinning down all the structural parameters and directly test for the significance of statistical discrimination. Therefore, we defer the discussion on how we recover the remaining structural parameters from the reduced-form parameter estimates until Section 4 and focus on the parameter estimates based on the reduced-form representation of the job-finding rates in this section.

Identification: One major drawback of relying on numerical methods to estimate our model is that we are silent on the identification of the model parameters. Given the lack of an explicit closed-form solution for $r(\tau)$, we are unable to formally prove that the observed data at our disposal uniquely pins down the model parameters. Furthermore, even if the model is identified, it is not clear what feature of the data identifies a certain model parameter. To shed light on these identification issues, our strategy is to obtain an approximation to the resume values which captures the job-finding rates reasonably well for a subset of individuals in our sample.

The complexity of $r(\tau)$ function is due to the presence of higher-order terms in Equation (20). However, recall that the resume values are strictly decreasing in unemployment duration and approach 0 as τ goes to infinity. Hence, for a sufficiently large value of τ, we can approximate Equations (20) and (21) by ignoring the second- and third-order terms in $r(\tau)$. Let τ^* denote such an unemployment duration. To simplify our discussion further, assume that we can group the observations based on a set of dummy variables. Then, for any $\tau > \tau^*$, we have the following approximation for the job-finding rates:

$$\lambda_i^r(\tau) \approx \tilde{\eta}_i[1 + r_i(\tau^*)(\phi + \zeta/\hat{r})\exp\{-(\gamma + \tilde{\eta}_i\phi)(\tau - \tau^*)\}] \text{ for } \tau > \tau^*, \tag{23}$$

[13] There are also time-varying regressors, such as market tightness, in our estimation. To account for their presence, we slightly change our notation and rewrite Equation (22) in Online Appendix B, where we detail our estimation procedure.

where $r_i(\tau^*)$ is the resume value at τ^* for observations with x_i.[14] Equation (23) describes a fully parametric model of the job-finding rates for $\tau > \tau^*$ and can be used to construct a likelihood function. In particular, using the observations with a sufficiently long duration, one can obtain estimates for β, γ, ϕ, and $r_i(\tau^*)(\phi + \zeta/\hat{r})$ following the standard procedures in maximum likelihood estimation.

Equation (23) is very helpful, because we can now directly observe how the reduced-form parameters interact with the observed data, which are not possible with Equations (20) and (21). Given the explicit form in Equation (23), we can explain what data feature identifies each parameter as follows: First, the job-finding rates of the long-term unemployed workers conditional on individual characteristics identify the coefficient vector β, because $\lambda_i^r(\tau)$ in Equation (23) approaches to $\tilde{\eta}_i$ as $\tau \to \infty$. Second, γ appears as the coefficient in front of unemployment duration, whereas ϕ appears as the coefficient in front of the interaction of unemployment duration with $\tilde{\eta}_i$. Noting that $\tilde{\eta}_i$ depends on observable characteristics from Equation (22), we conclude that γ is identified by data on unemployment duration, and the interaction of unemployment duration with the observable characteristics identifies ϕ.[15] This result is intuitive, because skill depreciation is common to everyone, whereas ϕ reflects dynamic selection and depends on individual characteristics.[16]

Finally, from Equation (23), the job-finding rates of the unemployed workers with duration τ^* yield an estimate for $r_i(\tau^*)(\phi + \zeta/\hat{r})$ for each group i. We can write this relationship as follows:

$$\frac{\lambda_i^r(\tau^*)}{\tilde{\eta}_i} - 1 \approx r_i(\tau^*)(\phi + \zeta \hat{r}). \tag{24}$$

At this point, there are two more reduced-form parameters left, ζ and ξ. Note also that $r_i(\tau^*)$ is an endogenous outcome and we cannot separate it from $(\phi + \zeta/\hat{r})$ in Equation (24). However, $r_i(\tau^*)$ negatively depends on ξ but does not depend on ζ, because the latter does not enter into the resume-updating Equation (20). Therefore, given all the other parameter estimates, Equation (24) represents a positively sloped curve when plotted on a $(\xi - \zeta)$ plane.

To separate ζ and ξ, we can utilize the job-finding rates of workers with $\tau \le \tau^*$. In particular, the job-finding rate of a recently unemployed worker $(\tau = 0)$ relative to the job-finding rate of long-term unemployed workers is given by:

$$\frac{\lambda_i^r(0)}{\tilde{\eta}_i} - 1 = \phi\hat{r} + \zeta + \xi\hat{r}, \tag{25}$$

where we use $r_i(0) = \hat{r}$. Contrary to Equation (24), Equation (25) represents a negatively sloped curve when plotted on a $(\xi - \zeta)$ plane. Consequently, Equations (24) and (25) uniquely identify ζ and ξ.

Statistical discrimination: With the help of Equations (24) and (25), we can analyze what data feature identifies ζ, our parameter of interest. Holding everything else constant, a higher

[14] Note that the resume values for $\tau > \tau^*$ for group i are approximated by:

$$r_i(\tau) \approx r_i(\tau^*) \exp\{-(\gamma + \tilde{\eta}_i\phi)(\tau - \tau^*)\}.$$

[15] We need at least one variable that affects the contact rate but not skill depreciation and interview success probabilities. If $\tilde{\eta}_i$ were common across the groups, say $\tilde{\eta}$, then data on unemployment duration would yield a single estimate for the term $(\gamma + \tilde{\eta}\phi)$. Although the long-term job-finding rates still identify $\tilde{\eta}$, we would not be able to separate γ and ϕ from a single estimate for $(\gamma + \tilde{\eta}\phi)$.

[16] From Equation (20), dynamic selection is stronger in high $\tilde{\eta}_i$ markets and this feature is reminiscent of the proportionality of job-finding rates to the contact rate in Equation (8). The mixed proportional hazard model, which has been extensively studied in the literature, has a similar feature. See subsection 5.3 in Van den Berg (2001) for an insightful and detailed discussion on the role of interaction between duration and explanatory variables in the observed hazard in identification of the mixed proportional hazard model.

value of $\lambda_i^r(\tau^*)$ yields a higher the value of ζ (and a lower value for ξ) from Equations (24) and (25). In other words, ceteris paribus, the higher the job-finding rate of medium-term unemployed workers relative to the long-term unemployed workers, the larger the effect of statistical discrimination.[17]

To understand this result, note that statistical discrimination is driven by the interview decisions of D-type firms. If statistical discrimination is an important source of negative duration dependence, the share of D-type firm, α, must be large in the first place. Moreover, both the medium- and long-term unemployed workers are mostly low-skill workers, and their respective job-finding rates are driven by these workers. However, relatively more high-skill workers are among the medium-term unemployed workers and a D-type firm is more likely to interview a worker in this group. If the difference between the job-finding rates of medium- and long-term unemployed workers is large, low-skill workers must have a good chance to convert an interview with a D-type firm to an actual job offer. That is, p_D^L must be large. From Equation (14), the reduced-form parameter ζ is large when both α and p_D^L are large.

3.3. *Estimation Results.* The structural parameters in our model may differ for different demographic groups. To alleviate this concern, we keep our sample relatively homogeneous by restricting it to white males only, which is the largest group based on gender and race. We treat this group as our main sample and perform two robustness analysis by further restricting our sample.

First, we exclude unemployment spells falling into the Great Recession period, which corresponds to the months between December 2007 and June 2009 (both inclusive). Our motivation for excluding this time period stems from the fact that we restrict $r(0)$ to be constant in our regressions and a dramatic change in the flows into the unemployment pool could have occurred during this recessionary period. A constant $r(0)$ is compatible with a steady-state assumption, and we think our restricted sample is better suited for this assumption.

Second, in Equation (22), we assume that observable characteristics matter only for the matching process but not for the interview stage and the skill depreciation process. However, education, occupation, and industry of the worker may potentially matter for the subsequent stages after the matching process. As another robustness check, we further restrict our sample to white males who have no college education and were previously employed in a goods producing sector with a production related occupation. This group corresponds roughly to one-third of our main sample and it is the largest group within white males.

We present our estimation results in Tables 1 and 2. In each table, the first column uses our main sample with white males only; the second column excludes the Great Recession period; and the third column uses our sample with further restrictions on education, occupation, and industry.

Table 1 shows our estimates for the reduced-form parameters of our model governing the negative duration dependence together with the skill depreciation rate. In each row, we report three numbers. The first number is our point estimate. The second and the third numbers are the likelihood ratio (LR) test statistic and the p-value associated with this statistic. In each case, the parameter of interest is set equal to 0 under the null hypothesis. The p-values are not retrieved from the χ^2 distribution. Because we are testing at the boundary of the parameter space, the asymptotic properties of the maximum likelihood estimator are not valid.[18] To address this issue, we simulated the LR test statistic under the null hypothesis using our sample. More specifically, for each observation in our sample, we randomly simulated employment status based on the job-finding rates under the null hypothesis. Then, we estimated the LR test statistic for this random sample. We repeated this process for 1,000 random samples

[17] To give the readers a sense of how big τ^* is, the resume values based on our estimation are near or below 0.1 beyond six months of unemployment.

[18] See Andrews (2001) for a detailed discussion.

Table 1
MAXIMUM LIKELIHOOD ESTIMATION RESULTS FOR THE REDUCED-FORM PARAMETERS GOVERNING DURATION DEPENDENCE AND SKILL DEPRECIATION

Parameter	I	II	III
ζ:	1.716	1.855	1.444
	22.561	23.793	9.549
	<0.001	<0.001	0.001
ϕ:	0.083	0.000	0.511
	0.034	0.000	0.357
	0.320	1.000	0.247
ξ:	1.907	2.028	2.367
	7.028	6.421	3.008
	0.006	0.003	0.025
γ:	0.204	0.192	0.126
	28.671	29.873	13.692
	<0.001	<0.001	<0.001
# of observations:	59,340	45,131	20,157
# of unemployment spells:	35,338	27,677	12,264

Note: Column I uses our main sample with white males. Column II excludes the Great Recession period. Column III uses white males with no college education who were previously employed in a goods producing industry with an occupation in the production category.

Table 2
MAXIMUM LIKELIHOOD ESTIMATION RESULTS FOR EXPLANATORY VARIABLES

Variable	I	II	III
Market tightness:	0.354	0.299	0.423
	(0.047)	(0.057)	(0.104)
Age in months ($\times 10^{-3}$):	−1.041	−0.942	−1.034
	(0.183)	(0.214)	(0.372)
Education			
Some college:	−0.050	−0.055	–
	(0.037)	(0.042)	–
At least college:	0.008	−0.003	–
	(0.047)	(0.053)	–
Occupation			
Personal services:	0.088	0.060	–
	(0.058)	(0.064)	–
Sales and office:	−0.066	−0.097	–
	(0.051)	(0.058)	–
Production:	0.078	0.048	–
	(0.049)	(0.055)	–
Industry (services):	0.007	−0.009	–
	(0.040)	(0.040)	–
Reason for unemployment			
End of temporary job:	0.325	0.298	0.350
	(0.056)	(0.065)	(0.106)
Job leaver:	0.343	0.349	0.290
	(0.065)	(0.078)	(0.133)
Reentrant:	0.162	0.163	0.184
	(0.044)	(0.051)	(0.095)
Marital status			
Widowed/divorced:	−0.163	−0.138	−0.211
	(0.045)	(0.052)	(0.091)
Never married/single:	−0.221	−0.213	−0.271
	(0.043)	(0.051)	(0.093)
# of observations:	59,340	45,131	20,157
# of unemployment spells:	35,338	27,677	12,264

Note: Standard deviations are indicated in parentheses. Column I uses our main sample with white males. Column II excludes the Great Recession period. Column III uses white males with no college education who were previously employed in a goods producing industry with an occupation in the production category.

to obtain an approximation for the distribution of the LR test statistic associated with that particular null hypothesis.[19]

Our main parameter of interest is ζ, which is a direct estimate of percentage increase in long-term job-finding rates after counterfactually removing the interview costs. Our point estimate is equal to 1.716, implying almost a threefold increase in long-term job-finding rates, and it is statistically different than 0. Our estimates under columns II and III are also large and statistically significant despite a significant drop in sample sizes due to additional sample restrictions. Recall from Equation (14) that ζ captures the combined effect of the share of discriminating firms, α, and the interview success probability of low-skill workers, p_j^L's. We able to assess the importance of statistical discrimination for the long-term unemployed without further restrictions on these parameters. In the next section, we impose identification restrictions on some of the structural model parameters to pin down α and p_j^L's separately, and calculate counterfactual job-finding rates at all unemployment durations.

Our point estimate for skill depreciation, γ, indicates that every month, about 20% of unemployed workers lose their skills. This number is consistent with Doppelt (2016), who finds that it takes an average time of six months for someone to experience a drop in skills. Our estimate for ξ is large and statistically significant, which implies that the response of high-skill job-finding rates to the resume values is stronger. Finally, the point estimate for ϕ using our full sample is small and not significantly different from 0, and it is equal to 0 with our sample excluding the Great Recession. From Equation (18), this estimate suggests that interview success probabilities with N-type firms are similar for high- and low-skill workers.

Table 2 gives a breakdown of coefficient estimates for a selected set of explanatory variables that enter the expression in Equation (22). The standard errors are indicated in parentheses. Overall, our estimates for the remaining parameters are consistent with the findings in Barnichon and Figura (2015), which estimate a generalized matching function. In particular, we used the natural logarithm of market tightness in our estimation, which is consistent with the Cobb–Douglas matching function specification used in their paper.[20] Therefore, our estimate for the coefficient in front of the market tightness is directly comparable to theirs. We also find the job-finding rate increases with market tightness, and our point estimate is similar in magnitude to the estimate in Barnichon and Figura (2015).

To conduct an overall assessment of the fit of our model, we compare the monthly job-finding probabilities predicted by the parameter estimates from our model with those non-parametrically estimated from the data. As in the literature, we use the Kaplan–Meier formula, which estimates the job-finding probability at a given unemployment duration as the share of the unemployed who find a job within a month. To obtain the predicted monthly job-finding probabilities from our model, we first calculate the predicted job-finding probability for each observation in our sample using the parameter estimates along with the explanatory variables. Then, we group individuals based on their reported unemployment duration and calculate the weighted average of these monthly job-finding probabilities. Constructed this way, the predicted job-finding probabilities retain the impact of explanatory variables and are comparable to the Kaplan–Meier estimates. In Figure 1, we plot both of these measures by unemployment duration using our full sample. The circles are proportional to the sample sizes in each bin. Overall, our parametric model tracks the job-finding probabilities calculated using the Kaplan–Meier formula very closely even at longer durations for which we have fewer observations.

[19] In our case, because we restrict the parameter space, the log-likelihood of the unrestricted model often coincides with that of the restricted model.

[20] Although the job-finding rates derived from our model are consistent with a Cobb–Douglas matching function, we did not model the equilibrium determination of the market tightness. In the standard Diamond–Mortensen–Pissarides matching model, the market tightness is pinned down via a free-entry condition in equilibrium. In our setup, the uniqueness of a steady-state equilibrium is not guaranteed, due to skill deprecation. See Pissarides (1992) for a detailed analysis.

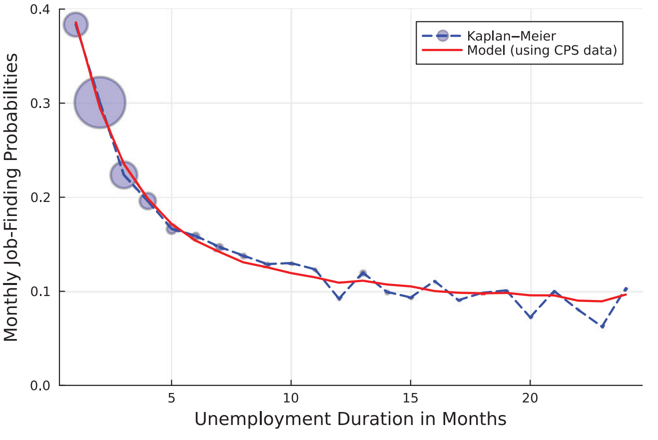


FIGURE 1

MONTHLY JOB-FINDING PROBABILITIES BY UNEMPLOYMENT DURATION: MODEL FIT USING CPS DATA VERSUS KAPLAN–MEIER ESTIMATES

In Figure 1, the job-finding probabilities at 6 and 24 months are 0.154 and 0.097, respectively. The fact that there is about a 60% difference accords with our analysis that relates ζ to the medium- and long-term job-finding rates. However, one should be mindful of dynamic selection due to observed heterogeneities in Figure 1. In the next section, we verify that the difference remains large even after removing the effects due to observed heterogeneities.

4. CHARACTERIZATION OF THE EQUILIBRIUM

Our analysis has so far focused on estimating the job-finding rates, using the reduced-form representation described in Equations (20) and (21). Although our estimation strategy allows us to directly estimate the change in counterfactual long-term job-finding rates after removing the interview costs, we cannot shed light on the model mechanism without the structural parameter estimates. In this section, we first describe how we recover the structural parameters of our model. Then, we use the values of the structural parameters to describe the quantitative features of our model. We study the counterfactual job-finding rates by unemployment duration in the next section.

4.1. *Baseline Calibration.* We have a direct estimate for the skill depreciation rate, γ, and we set its value to 0.204 from column I of Table 1. We also set the interview cost parameter $\hat{r}$ equal to 0.75 as in the estimation part.

We have six structural parameters remaining that directly appear in the job-finding rates: four p_j^s's, α, and η. Four Equations ((14), (17), (18), and (19)) describe their relationship with the reduced-form parameters. We use our reduced-form parameter estimates for ζ, ϕ, and ξ from column I of Table 1 in these equations. Moreover, the data contain observed individual heterogeneities, captured by Equation (22), but we assumed them away in the theoretical model. We calculate the sample average of the fitted values to obtain a single value for $\tilde{\eta}$ to be used in Equation (17). Taken together, we have four equations in six unknown structural parameters. To obtain a point estimate for these structural parameters, we need to impose (at least) two more restrictions.

The first restriction we impose is that the interview success probability of a high-skill worker is independent of the type of the firm: $p_D^H = p_N^H = p^H$. We have two motivations for this restriction. First, Equation (14) suggests that the long-term job-finding rates largely depend on the interview success probabilities of the low-skill workers with different types of

TABLE 3
STRUCTURAL MODEL PARAMETERS

Parameter	Value
$\hat{r}$: Interview cost parameter	0.750
γ: Skill depreciation rate	0.204
η: Job-opportunity arrival rate	1.012
α: Share of D-type firms	0.770
Interview success probabilities:	
$p^H_D = p^H_N = p^H$: (high-skill)	0.559
p^L_D: (low-skill/D-type)	0.265
p^L_N: (low-skill/N-type)	0.515
δ: Separation rate	0.016
σ: On-the-job-learning	0.016

firms. Therefore, we focus on obtaining separate estimates for p^L_D and p^L_N. Second, our restriction is also in line with Jarosch and Pilossoph (2019), which imposes that the highest skill type can form a viable match with any firm.

The second restriction we impose is a normalization assumption on the interview success probabilities. η in Equation (8) multiplies through p^s_j's and we can normalize the value of one of these parameters without changing the job-finding rates. We choose to normalize the average interview success probability of the recently unemployed to 0.5:

$$r(0)p^H + (1 - r(0))\left((1-\alpha)p^L_N + \alpha p^L_D\right) = 0.5. \tag{26}$$

Using these restrictions together with Equations (14), (17), (18), and (19), we are able to obtain separate estimates for p^L_D and p^L_N, which are crucial for the long-term job-finding rates.

Our choice for the normalization value is innocuous for the counterfactual analysis in the next section. That is, we obtain identical counterfactual results if we use a value other than 0.5 even though we obtain different values for η and p^s_j's. To see this point, note that we obtain a value for α using the reduced-form parameter estimates by relying only on the first restriction and its value is independent of the other structural parameter values. Moreover, the interview success probabilities in the reduced-form Equations (14), (18), and (19) are all proportional to p^L_N and independent of η. Therefore, we can obtain p^L_N from Equation (26), and then η from Equation (17). Note that p^L_D is proportional to the normalizing constant and changing the normalization value changes p^s_j's and η in opposite directions by the same proportion without changing the skill-specific job-finding rates in Equation (8). The counterfactual job-finding rates also depend on the specific value of α, because discriminating firms also interview everyone. However, the value of α does not depend on the normalization value. Thus, the counterfactual job-finding rates are the same regardless of the normalization value.[21]

Two more parameters in the model do not enter the job-finding rate equations directly but are required to calculate equilibrium: the separation rate, δ, and the on-the-job-learning parameter, σ. Of these two parameters, we set $\delta = 0.016$ so that the steady-state unemployment is 6%. For σ, we note that $r(0)$ is determined in equilibrium, and we set its value equal to $\hat{r}$ in the estimation part. Accordingly, we set σ so that $r(0) = \hat{r}$ in equilibrium from the flow equation in (13). The calibrated value of σ is equal to 0.016. A summary of the structural model parameter values is presented in Table 3.

Using the calibrated parameter values in Table 3, we calculate the average monthly job-finding rate as 0.254, but it shows great variation by unemployment duration. The solid red line in Figure 2 displays the equilibrium job-finding rates by unemployment duration in our

[21] In Online Appendix B, we show that the first restriction guarantees that all the interview success probabilities are less than 1 for normalization to an arbitrary value less than or equal to $\hat{r}$. We also show that the counterfactual job-finding rates do not depend on the normalizing constant.

FIGURE 2

MONTHLY JOB-FINDING RATES BY UNEMPLOYMENT DURATION: MODEL UNDER BASELINE CALIBRATION VERSUS MODEL FIT USING CPS DATA

model. On entry into the unemployment pool, the job-finding rate is slightly above 0.5, but it sharply drops to 0.2 within only four months. After this point, the job-finding rate continues to drop, albeit at a slower pace, reaching 0.12 by 15 months of unemployment and it remains constant thereafter.

In the remainder of this section, we describe some of the features of the baseline equilibrium and discuss their implications for the job-finding rates by unemployment duration.

4.2. *Observed Heterogeneity.* The calibrated value of η implies that workers receive on average about one job opportunity in a month. We use a constant value for η in calculating the equilibrium above, but our estimates in Table 2 suggest considerable variation in job-opportunity arrival rates. Such differences in observed individual characteristics would imply an even steeper decline in job-finding rates by unemployment duration due to a composition effect. For example, young individuals find a job much faster than the population average and they tend to be short-term unemployed; therefore, they disproportionately contribute more(less) to the calculation of short(long)-term job-finding rates.

To assess the role of observed heterogeneity in job-finding rates, we calculate predicted job-finding rates for our sample using our parameter estimates in Tables 1 and 2 and the full set of explanatory variables. We follow a procedure that is identical to calculating the monthly job-finding probabilities in Figure 1 except that we now calculate job-finding rates, as opposed to probabilities, at the beginning of a given month. The dashed blue line in Figure 2 shows the predicted monthly job-finding rates from our CPS data, which display variation in job-opportunity arrival rate by construction.

A comparison between the job-finding rates from our model with constant η and the predicted job-finding rates with the full CPS data gives us an idea about the contribution of observed individual heterogeneity to unemployment duration dependence. An inspection of Figure 2 reveals that observed individual heterogeneity plays a nontrivial role in explaining the negative duration dependence, especially at short durations. As we move from the shortest duration to the longest duration in the CPS data, the job-finding rate declines from about 0.6 to around 0.1. In our model with constant η, the job-finding rate for the recently unemployed is only about 0.5 and declines down to 0.12. These numbers suggest about one-fifth of the duration dependence is attributable to observed differences in job-finding rates. Our following analysis focuses on the remaining part.

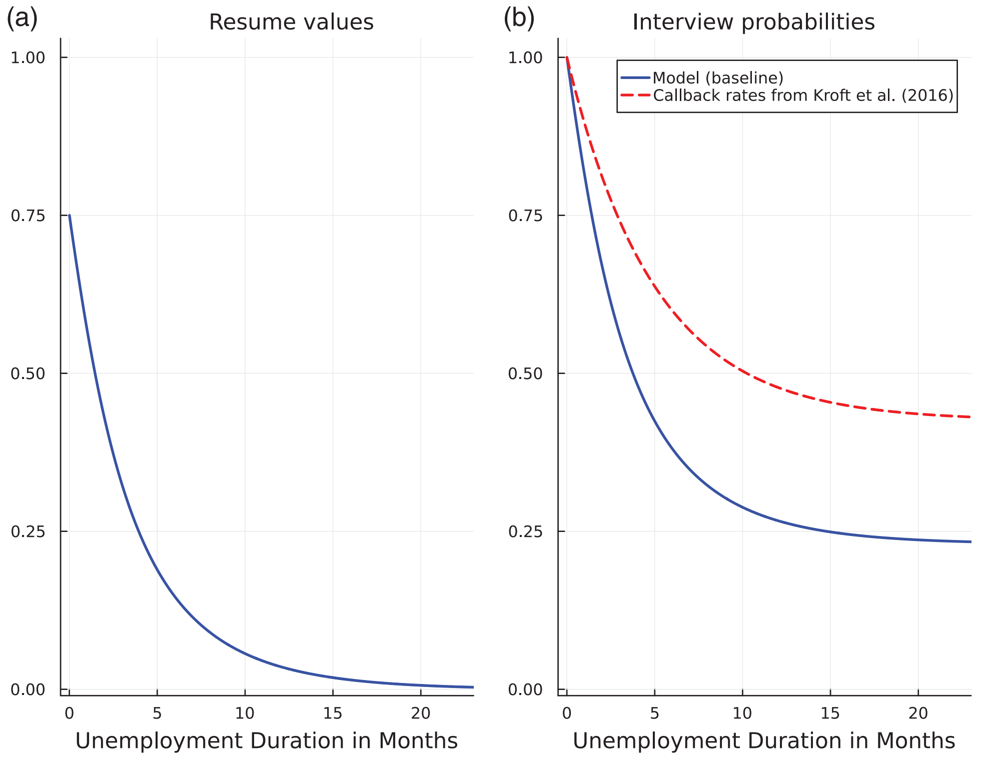


FIGURE 3

(A) RESUME VALUES UNDER BASELINE CALIBRATION AND (B) INTERVIEW PROBABILITIES UNDER BASELINE CALIBRATION VERSUS RELATIVE CALLBACK RATES IN KROFT ET AL. (2016)

Note that the job-finding rate at six months of unemployment duration in Figure 2 is about 40% higher than the long-term job-finding rates even after removing the effects due to observed heterogeneities (0.169 vs. 0.121).[22] This sizable difference suggests an important role of statistical discrimination, which we verify in the next section.

4.3. *Resume Values and Interview Probabilities.* The decline in job-finding rates in Figure 2 is driven by the changes in the share of high-skill workers and the interview decisions of the firms in response to these changes. In Figure 3(a), we plot the resume values by unemployment duration, $r(\tau)$, which reflects the share of high-skill workers. The interview probabilities, conditional on a match, depend on the resume values and the share of D-type firms, and it is equal to $(1-\alpha+\alpha r(\tau)/\hat{r})$. Figure 3(b) displays interview probabilities by unemployment duration.

The resume values start at 0.75, which we target in the baseline equilibrium, and smoothly decline to 0 over time. The interview probability for a recently unemployed is equal to 1, which we also target in the baseline equilibrium, and it tracks the pattern in resume values. Because the resume values approach 0 with unemployment duration, the long-term unemployed workers are largely interviewed with N-type firms, the share of which is equal to $(1-\alpha)$. The value of α in Table 3 implies that the interview probability approaches $1-0.770=0.230$.

[22] Kroft et al. (2016) fit a "double" exponential model to the job-finding rates observed in the monthly CPS data from 2002 to 2013. In their figure 7, panel A, the medium-term job-finding rates are larger than the long-term job-finding rates and this difference is similar to ours.

In Figure 3(b), we also plot the relative callback rates for an interview (red dashed line) by unemployment duration from Kroft et al. (2016). They fit a negative exponential model using the experimental data with fictitious resumes from Kroft et al. (2013). Specifically, their fitted model is $A(\tau) = 0.425 + (1 - 0.425)\exp(-0.199\tau)$, where τ is unemployment duration and $A(\tau)$ is the measured relative callback rates. Given this specification, the long-term callback rate is 0.425. Through the lens of our model, their estimate suggests that the share of N-type firms should be larger so that the interview probability for the long-term unemployed approaches 0.425 as opposed to 0.230.

We interpret our estimate as an alternative to Kroft et al. (2016) because our sample differs from theirs in important ways. The fictitious resumes in Kroft et al. (2013) are identical by design, and they may not match well with our CPS sample. Their sample is also relatively young and inexperienced.[23] Roughly two-thirds of their sample is female, whereas our sample is restricted to white males. Moreover, their experiment is conducted from August 2011 and July 2012 when the job-finding rates were still below their historical average. Given the slow recovery after the Great Recession and the sensitivity of job-finding rates to aggregate conditions, it is likely that their sample displays a flatter profile than normal times.[24]

The share of discriminating firms, α, directly affects the interview probabilities by unemployment duration, but the prevailing job-finding rates also depend on the ability of the workers to turn these interviews into actual job offers. The latter depends on the interview success probabilities, p_j^s, which we study extensively in the next section.

5. COUNTERFACTUAL ANALYSIS

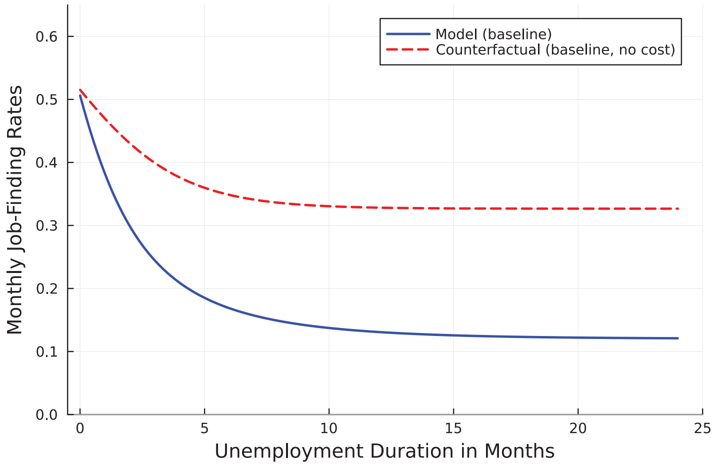
Our main focus in this section is quantifying the importance of statistical discrimination on long-term job-finding rates. We find statistical discrimination has a large effect on long-term job-finding rates. Because our main finding regarding the role of statistical discrimination is in stark contrast to Jarosch and Pilossoph (2019), who argue that removing the interview cost has a negligible effect on the job-finding rates of long-term unemployed workers, we provide a detailed discussion comparing two papers in a separate subsection.

5.1. *Removing Interview Costs.* Firms discriminate against long-term unemployed workers in our model, due to the presence of the costly interview stage in the hiring process. The first counterfactual exercise we perform is to remove the interview costs so that every firm interviews every worker in the new equilibrium. Because this exercise follows Jarosch and Pilossoph (2019), our results in this section are directly comparable to theirs.

Figure 4 shows the equilibrium monthly job-finding rates by unemployment duration under baseline calibration together with the counterfactual job-finding rates in the new equilibrium after removing the interview costs. At the estimated parameter values (the blue solid line in the figure), the average job-finding rate of a recently unemployed worker is about 0.506, and it decreases by about 40 percentage points to 0.120 after 24 months. After removing the interview cost (the dashed red line in the figure), the job-finding rate of a worker with 24 months of unemployment duration jumps by about 20 percentage points from 0.120 to 0.326. Therefore, the statistical discrimination due to the presence of interview costs accounts for about half of the decline in the job-finding rates moving from short-term to long-term unemployment duration (20 of 40 percentage points difference). The remaining half is attributable to the skill depreciation and the composition effect due to differential job-finding rates.

[23] Kroft et al. (2013) report that the sample average of age is 27.903.

[24] As an alternative, we have considered replacing the identification assumption $p_D^H = p_N^H$ in Subsection 4.1 with $1 - \alpha = 0.425$. However, p_j^s's are not guaranteed to be between 0 and 1 in this case for an arbitrary normalization that is less than or equal to $\hat{r}$. Furthermore, we reestimate our model in Online Appendix C under a restriction on the reduced-form parameters that implies $1 - \alpha = 0.425$. Our point estimate for the reduced-form parameter ζ under this restriction is still large.



NOTES: Counterfactual job-finding rates are calculated in the new equilibrium after removing the interview costs.

FIGURE 4

MONTHLY JOB-FINDING RATES BY UNEMPLOYMENT DURATION UNDER BASELINE CALIBRATION: MODEL VERSUS COUNTERFACTUAL

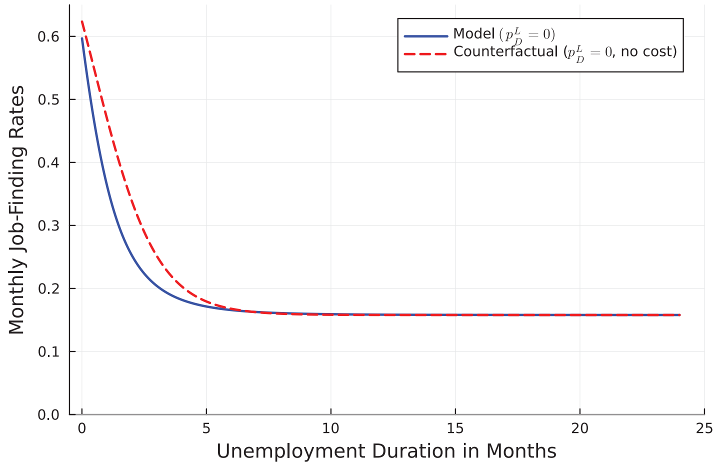
Equivalently, the increase from 0.120 to 0.326 implies that the job-finding rates of long-term workers increase by 171.6% after removing the interview costs. This number coincides with our estimate for the reduced-form parameter ζ (times 100) in column I of Table 1. Recall that ζ corresponds to a direct measure of the impact of removing the interview costs on the long-term job-finding rates. The similarity between the two numbers verifies our interpretation of this reduced-form parameter.

The expression for ζ in Equation (14) suggests that our counterfactual result can be explained by the interview success probability of low-skill workers with D-type firms. Intuitively, long-term unemployed workers are mostly low-skill, and their resume values are close to 0. Hence, these workers are interviewed and hired in equilibrium mostly by N-type firms. When one removes the interview cost, D-type firms also interview every applicant regardless of their unemployment duration, and we would see an increase in the interview probability, particularly for long-term unemployed workers. At a very long unemployment duration, the interview probability goes up from $(1-\alpha)$ to 1. However, these long-term unemployed workers would still mostly be low-skill workers and the prevailing change in the long-term job-finding rates would depend on how low-skill workers perform in these free interviews.

This analysis suggests that if the value of p_D^L is small, removing the interview cost would not have any effect on the long-term job-finding rates. To verify this claim, we repeated the counterfactual exercise above under $p_D^L = 0$.[25] In Figure 5, the solid blue line corresponds to the equilibrium job-finding rates under $p_D^L = 0$. The dashed red line, on the other hand, shows the counterfactual job-finding rates under $p_D^L = 0$ after removing the interview costs. The figure shows the long-term job-finding rates are unchanged under this parameter restriction.

5.2. *Comparison to Jarosch and Pilossoph (2019).* In this section, we compare and contrast our counterfactual result with that in Jarosch and Pilossoph (2019) in great detail. Our main point is that the interview success probability of low-skill workers with D-type firms is key to understanding the differences between our article and theirs. We stress that our model

[25] To maintain comparability to the baseline equilibrium, we also set $\eta = 1.329$ and $\sigma = 0.014$ so that the job-finding rates and the initial resume values are equal to those in the baseline equilibrium. The rest of the parameters are set equal to their values in Table 3.



NOTES: Counterfactual job-finding rates are calculated in the new equilibrium after removing the interview costs.

FIGURE 5

MONTHLY JOB-FINDING RATES BY UNEMPLOYMENT DURATION UNDER $p_D^L = 0$: MODEL VERSUS COUNTERFACTUAL

delivers their main result under a certain parameter restriction, $p_D^L = 0$, but when confronted with data, we reject this parameter restriction.

5.2.1. *A generalization of the job-finding rates.* We begin with a brief description of the model in Jarosch and Pilossoph (2019). Workers and firms are heterogeneous with respect to their skill and productivity levels, respectively, and meet randomly in the labor market. Firms observe the worker's unemployment duration but not their true skill level when deciding on whether to interview the applicant. Although firms face the same interview cost, they still statistically discriminate against long-term unemployed workers when calling for an interview, because their skill levels tend to be lower, thus generating a lower surplus. A more distinctive feature of their model is the assumption on the production function: a worker–firm pair produces a positive output only if the worker's skill is at least as good as the firm's productivity. This specification places a rather strong assumption on the interview success probabilities: given the worker's skill level, the interview success probability is either 0 or 1, depending on the interviewing firm's productivity. For the long-term unemployed workers, who tend to be relatively low-skill, this probability is 0 for most firms, and they are exclusively hired by the firms at the bottom of the firm distribution, which have low or no minimum hiring requirements. When the interview cost is counterfactually removed, these workers see a large increase in the number of interviews received but systematically fail to get hired. Consequently, the job-finding rates of long-term unemployed workers remain largely unchanged after removing the interview cost.

To facilitate a formal comparison between the two models, we rewrite the job-finding rates more generally as follows:

$$\lambda^s(\tau) = \eta((1-\alpha)p_N^s + \alpha\iota(\tau)p_D^s(\tau)), \tag{27}$$

where $\iota(\tau)$ is the interview probability with a worker at unemployment duration τ, and $p_D^s(\tau)$ is the skill-specific probability of getting hired after an interview with a D-type firm at unemployment duration τ. It is straightforward to verify our model follows this specification. Specifically, our model implies $d\iota(\tau)/d\tau < 0$, $\iota(\tau) \to 0$ as $\tau \to \infty$, and p_j^s are constants.

The interview probabilities, $\iota(\tau)$, and the outcome of these interviews, $p_D^s(\tau)$, are more complex in Jarosch and Pilossoph (2019), because they depend on the endogenously determined distribution of workers by unemployment duration as well as the productivity distribution of interviewing firms. We write both quantities as functions of τ to emphasize this dependence. In the remainder of this section, we establish that the job-finding rates in Jarosch and Pilossoph (2019) follow the same specification.[26]

Worker and firm heterogeneity: In Jarosch and Pilossoph (2019), a continuum of worker and firm types exists, both of which lie in the unit interval. Let $s \in [0, 1]$ and $y \in [0, 1]$ denote worker skill (denoted by x in the original paper) and firm productivity, respectively. By assumption, a match produces $s + y$ only if $s \geq y$, so that hiring requires that the worker's skill exceeds the interviewing firm's productivity. Moreover, firm productivity is uniformly distributed with a positive mass at the bottom of the distribution. This mass of firms, denoted by $F(0)$ in their paper, is akin to N-type firms in our model, because they have no minimum skill requirement: that is, $s \geq 0$ for all s, and hence, they interview every applicant. The remaining firms make nontrivial interview decisions depending on their productivity level. Therefore, they are similar to D-type firms in our model. Accordingly, we can associate α directly with (1-$F(0)$) in Jarosch and Pilossoph (2019).

Interview decisions: Considering the interview decisions of D-type firms, $\iota(\tau)$, they are not driven by random interview costs but instead by the interviewing firm's productivity level and the distribution of workers at duration τ. Although Jarosch and Pilossoph (2019) have no simple expression for this interview probability, it also declines with unemployment duration and approaches 0 at long durations as in our model, that is, $d\iota(\tau)/d\tau < 0$ and $\iota(\tau) \rightarrow 0$ as τ becomes arbitrarily large.[27] Intuitively, D-type firms are willing to interview a worker with unemployment duration τ as long as enough high-skill workers exist at that unemployment duration.[28]

Hiring decisions: In Jarosch and Pilossoph (2019), the firms at the bottom of the distribution hire every worker that they interview. Therefore, $p_N^s = 1$ for any worker type. Our estimates for p_N^s's and p_N^H in Table 3 are consistent with their assumption, in that both skill types in our model have similar interview success probabilities with N-type firms, although they are not strictly equal to unity.

$p_D^s(\tau)$ in Equation (27) has a different interpretation in Jarosch and Pilossoph (2019), in that it corresponds to the share of firms for which the worker is eligible within the set of interviewing D-type firms. Although the set of these interviewing firms is determined in equilibrium, $p_D^s(\tau)$ takes rather a simple form for two skill types. First, the highest skill type, $s = 1$, turns every interview into an offer. Our equality restriction on the interview success probability of a high-skill worker with either type of firm is also consistent with this assumption, although we do not set them strictly equal to unity. Second, the lowest skill type, $s = 0$, is not eligible for a job with *any* D-type firm by assumption. Moreover, the estimated worker skill distribution is bimodal in Jarosch and Pilossoph (2019), and the job-finding rates are largely driven by the workers near these two extreme points in the skill distribution.[29]

Counterfactual job-finding rates: The counterfactual job-finding rates are easier to characterize in Jarosch and Pilossoph (2019). Once they remove the interview cost, every firm interviews every worker as in our model. Through the lenses of Equation (27), we can think of this exercise as setting $\iota(\tau) = 1$ for all τ. Moreover, a worker is hired by a D-type firm only if his

[26] Jarosch and Pilossoph (2019) set up their model in discrete time, but this distinction is immaterial for our comparison.

[27] We also note that both papers assume $\iota(0) = 1$, although this assumption does not have a direct impact on long-term job-finding rates.

[28] In their baseline equilibrium, interview probabilities are equal to 1 for an extended period (up until five months) before dropping sharply in a few months and remain relatively constant thereafter.

[29] Specifically, they assume a Beta distribution for worker types. About 90% of the workers at the population level have a skill level that is either less than 0.1 or greater than 0.9.

skill is greater than the firm productivity. Given that the firm productivity is distributed uniformly, the probability of getting hired by a D-type firm conditional on an interview is equal to the worker's skill type, s. Accordingly, we can write the counterfactual job-finding rate for a given skill type $s \in [0, 1]$ as $\eta((1-\alpha)+\alpha s)$.[30] Moreover, when τ becomes arbitrarily large, the job-finding rates in the baseline equilibrium with statistical discrimination are equal to $\eta(1-\alpha)$ for every skill type, because only N-type firms interview and hire workers. Using these two expressions, one can write the percentage increase in the long-term job-finding rate for a given skill type after removing the interview costs as $s\alpha/(1-\alpha)$, which is a special case of Equation (14).

For a given value of α, the gains are smaller for low-skill workers, and it is equal to 0 for the lowest skill type. Hence, the overall increase in the long-term job-finding rates after removing the interview costs is limited by the average skill of the long-term unemployed workers. Quantitatively, the low-skill workers near zero disproportionately make up a large fraction of the long-term unemployed workers in their baseline equilibrium.[31] Consequently, the counterfactual long-term job-finding rates in Jarosch and Pilossoph (2019) are not very different from those in their baseline equilibrium.

Intuitively, the job-finding rates of low-skill workers are not really affected by statistical discrimination in their baseline equilibrium. A worker with small $s \geq 0$ randomly receives interviews and the callback rate falls with duration due to statistical discrimination. However, *by assumption*, a low-skill worker is not eligible for most of the firms with $y > 0$. Hence, his job-finding rates decline with unemployment duration but only slightly. Thus, eliminating statistical discrimination by giving free interviews cannot have a big impact on their job-finding rates, because they are not affected much by it in the first place.

This analysis suggests that the restrictions on the eligibility of the low-skill workers in Jarosch and Pilossoph (2019) imposed by their assumptions on the production function matter for the counterfactual job-finding rates.[32] One interpretation of strict skill requirements is that skill stands for easily verifiable qualifications of the applicant, such as education and work experience, and hiring decisions are based on these qualifications. For example, someone with a high-school degree would immediately be disqualified for a physician position at a local hospital. However, in the field experiment of Kroft et al. (2013), the resumes are identical and such information would be readily available on the resume. Therefore, thinking of skills as qualifications that are not so easily verifiable from the worker's resume but are correlated with his unemployment duration, such as soft skills, would be more suitable. Evaluation of the applicant's skills, in this case, would be noisy in a way that affects the outcome of the interview.[33]

One can view our model as a generalization of the production function in Jarosch and Pilossoph (2019), albeit with two distinct worker types, which allows for the possibility that a low-skill worker has a chance to be hired by a D-type firm. Our production function in Equation (1) does not preclude the case in which a low-skill worker is ineligible for a job with any D-type firm, for example, a degenerate distribution for $\psi_D^L(y)$. Instead of taking a

[30] This relationship holds regardless of the prevailing variation in output in viable matches. In this counterfactual setting, interviews are also freely available for D-type firms and all that matters for hiring decisions is the viability of matches, which is simply equal to the skill of the worker.

[31] In equilibrium, about 85% of the unemployed workers who are unemployed for at least six months have a skill level that is less than 0.1.

[32] We note that Jarosch and Pilossoph (2019) also estimate their model using the job-finding rates from the CPS data, instead of some other output or wage data, despite some methodological differences. More specifically, they employ a simulated method-of-moments approach to estimate the structural model parameters governing the job-finding rates and interview callback rates in their model. They target empirical moments of the (relative) job-finding rates from the CPS data along with the (relative) interview callback rates in Kroft et al. (2016). The implied variation in output levels is quantitatively negligible, and worker heterogeneity is largely about the viability of matches. See their footnote 29.

[33] Even more generally, some other factors could be independent of the worker's skill and the firm's productivity. For example, fatigue during the interview may negatively affect the outcome of an interview even if the worker is qualified for the job.

stand on their chances with a D-type firm, we parameterize the probability of forming a viable match with the parameter p_D^L, and estimate it with the job-finding rates from the CPS data.[34]

Finally, given the bimodal distribution of worker types in Jarosch and Pilossoph (2019), our model with two discrete worker types provides a good approximation to the overall job-finding rates in their paper.[35] In Figure 5, we are able to reproduce their main finding regarding the counterfactual job-finding rates after setting $p_D^L = 0$ in our model.[36] This finding confirms that the eligibility of the workers at the lower end of the skill distribution plays an important role in explaining the different counterfactual results. We also note job-finding rates at 6 and 24 months are indistinguishable from each other in Figure 5. Recalling that their difference is informative about the impact of statistical discrimination and is sizable in the CPS data, Figure 5 illustrates why the data do not support assumption $p_D^L = 0$.

5.2.2. *Interview costs.* Jarosch and Pilossoph (2019) argue the effect of statistical discrimination on job-finding rates increases with the interview cost. Therefore, verifying the interview costs in our model are comparable to theirs is crucial. Jarosch and Pilossoph (2019) calibrate the cost of interviewing an applicant as 10% of the average monthly output based on the direct measures of man-hours spent by company personnel to screen, evaluate, and interview the applicants. The average match surplus value is about 65% of the average monthly output in their equilibrium. Accordingly, the interview cost in their model is about 15.4% of the average match surplus.

In our model, the interview cost is random and the average cost incurred per interview is calculated as follows:

$$\frac{\displaystyle\int_0^\infty \alpha \frac{r(\tau)}{r_0}\left[\frac{W_D(r(\tau)) + W_D(0)}{2}\right]\Gamma(\tau)d\tau}{\displaystyle\int_0^\infty \left(1 - \alpha + \alpha\frac{r(\tau)}{r_0}\right)\Gamma(\tau)d\tau}.$$

Our estimation left the match surplus values unspecified. To calculate the average interview cost, we need to set values for the (expected) match surplus with a D-type firm. For comparability to Jarosch and Pilossoph (2019), we set its value with a high-skill and low-skill worker to a value that corresponds to the maximum and minimum surplus values reported in their equilibrium, respectively. The match-surplus value in their equilibrium ranges from 56.9% for a match between the lowest-skill worker and the lowest-productivity firm to 73.0% for a match between the highest-skill worker and the highest-productivity firm. Using these numbers, we calculate the average incurred interview cost relative to the average match surplus value as 19.0%, which is close to the interview cost in Jarosch and Pilossoph (2019).[37] Therefore, we argue the difference between the two papers is not due to the difference in the measured interview costs.

Revisiting our counterfactual analysis from a policy standpoint, we can now calculate the cost of eliminating statistical discrimination. Assume firms still face interview costs but are fully subsidized so that they indiscriminately interview every applicant. In this case, the

[34] Analogously, one could extend the production function in Jarosch and Pilossoph (2019) by allowing for randomness at the interview stage.

[35] In particular, the workers with the lowest skill type makes up about 25% of the recently unemployed workers (after discretization). See figure 2 in their paper. Targeting $r(0) = 0.75$ in our model is consistent with this number.

[36] Jarosch and Pilossoph (2019) estimate the fraction of the firms that have no skill requirement and interview every applicant as 0.398 so that $\alpha = 0.602$, which is smaller than our estimate α in Table 3. From Equation (14), a larger α implies a larger increase in the interview probabilities after removing the interview costs. Despite an increase in the interview probabilities that is larger than those in Jarosch and Pilossoph (2019), the long-term counterfactual job-finding rates in Figure 5 when $p_D^L = 0$ are still virtually unchanged relative to the benchmark equilibrium.

[37] Jarosch and Pilossoph (2019) calibrate the interview cost parameter using the numbers reported in Silva and Toledo (2009) and Barron and Bishop (1985), which are based on hiring costs actually incurred by the firms. For comparability with these data sets, we report the interview costs conditional on an interview actually taking place.

average interview cost is equal to 28.8% of the average match surplus, which is up from 19.0% under the benchmark case in the absence of any subsidy.[38] The gain from such a policy is a reduction in the unemployment rate. After the subsidy, the unemployment rate goes down from 6% to 3.61%.

5.3. *Job-Finding Rates and Average Resume Value.* In this subsection, we forbid firms from looking at the workers' unemployment duration when making a decision to interview the worker. Compared with subsidizing interview costs of the firms, making it illegal for firms to ask about the applicant's unemployment duration is more feasible and policy-relevant. We also note that the results from this section are *not* directly comparable to Jarosch and Pilossoph (2019).

Suppose firms cannot observe their applicants' resume values, but they are fully informed about the average of the resume value in the unemployment pool, $\bar{r}$. Under this restriction, the job-finding rate conditional on the true type, s, becomes

$$\lambda^s(\bar{r}) = \eta((1-\alpha)p_N^s + \alpha\bar{r}/r_0 p_D^s). \tag{28}$$

Note the average resume value is determined in equilibrium. Therefore, we write the job-finding rate explicitly as a function of $\bar{r}$. Although the interview probability is the same for every worker, skill depreciation and unobserved heterogeneity still lead to declining job-finding rates by unemployment duration because the unemployment pool is mixed with high- and low-skill workers and they have different job-finding rates. The proportion of high-skill workers changes by unemployment duration according to the following resume-updating rule, which is now given by the following equation:

$$\frac{dr(\tau;\bar{r})}{d\tau} = -\gamma r(\tau;\bar{r}) - (1 - r(\tau;\bar{r}))r(\tau;\bar{r}))(\lambda^H(\bar{r}) - \lambda^L(\bar{r})), \tag{29}$$

given $r(0;\bar{r})$.[39] By definition of the average resume value, we also have the following relationship:

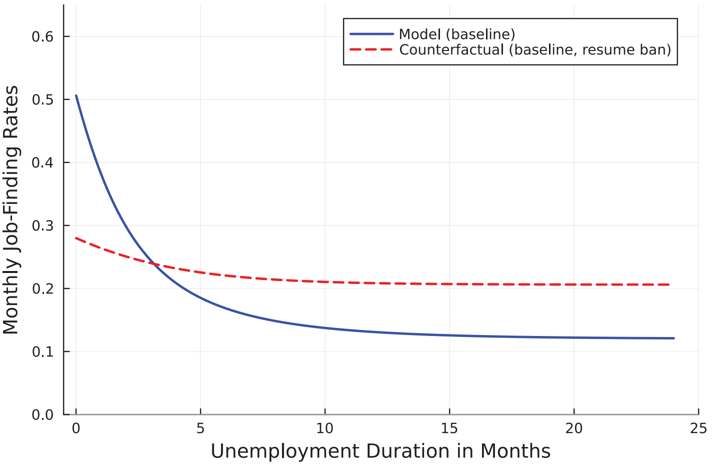
$$\bar{r} = \frac{\int_0^\infty r(\tau;\bar{r})\Gamma(\tau;\bar{r})d\tau}{\int_0^\infty \Gamma(\tau;\bar{r})d\tau}, \tag{30}$$

where $\Gamma(\tau;\bar{r})$ is the measure of the unemployed workers at duration $\tau \geq 0$ as before. Both $r(0;\bar{r})$ and $\bar{r}$ are determined in equilibrium.

In Figure 6, we plot the job-finding rates by unemployment duration when firms cannot observe the individual resume values together with our baseline results from the previous section. The job-finding rate for the long-term unemployed clearly increases at the expense of the short-term unemployed, and the negative duration dependence becomes much weaker. Therefore, forbidding firms from asking about their applicant's unemployment duration is an effective policy for eliminating statistical discrimination. However, the average job-finding rate decreases from 0.254 in the baseline case to 0.241 under this policy restriction because the resume value of recently unemployed workers declines from 0.75 to 0.70 (not shown). Consequently, the unemployment rate increases from 6% to 6.32%. We conclude that this policy is not an effective tool for reducing the overall unemployment rate.

[38] A policymaker would be more interested in the number of new jobs created after the subsidy. The interview cost per hire is equal to 41.3% of the average match surplus under the benchmark, and it increases to 47.6% with the subsidy.

[39] The differential equation in (29) has a closed-form solution. See Online Appendix A for details.



NOTES: Counterfactual job-finding rates are calculated in the new equilibrium where firms observe only the average resume value in the unemployment pool.

FIGURE 6

MONTHLY JOB-FINDING RATES BY UNEMPLOYMENT DURATION UNDER BASELINE CALIBRATION: MODEL VERSUS COUNTERFACTUAL

6. CONCLUSION

We study unemployment duration dependence in a job-search model with unobserved worker heterogeneity and learning about their types from their most recent unemployment duration. Firms facing interview costs discriminate against long-term unemployed workers because they tend to be low-skill. Our model delivers a parametric hazard model for job-finding rates, which we estimate using microlevel data from CPS via maximum likelihood. The long-term job-finding rates increase substantially after removing the interview costs relative to those in the baseline equilibrium with statistical discrimination. We argue that the difference between the medium- and long-term job-finding rates observed in the data are informative about the role of statistical discrimination in accounting for the steep decline in job-finding rates by unemployment duration.

Our conclusion with respect to the role of statistical discrimination in explaining the negative duration dependence is in stark contrast to the findings in Jarosch and Pilossoph (2019). We provide a formal comparison between the two models in an extension to our baseline model. Although these two models have several features in common, they have different assumptions about the interview success probability of low-skill workers with discriminating firms and deliver contrasting counterfactual outcomes. We are able to generate their counterfactual result in our model under a certain parameter restriction, but we did not find empirical support for this specification.

Examining the interview performance of long-term unemployed workers is crucial for understanding the role of statistical discrimination. From an empirical standpoint, interview success probabilities are hiring probabilities conditional on an interview. A more direct approach to estimating these probabilities would require a richer data set, for example, information on interviews and their outcomes. Such estimation is not possible with the CPS data, because we observe only the overall hiring rate without any information about interviews. We leave this task for future research when a richer data set becomes available.

DATA AVAILABILITY STATEMENT The basic monthly CPS data were derived from IPUMS-CPS available in the public domain: https://cps.ipums.org/cps/. The state-level unemployment data were derived from the U.S. Bureau of Labor Statistics available in the public domain: https://data.bls.gov/cgi-bin/srgate. The state-level vacancy data are available from the

Conference Board. Restrictions apply to the availability of the vacancy data, which were used under license from Haver Analytics for this study.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table A1: True parameter values assumed for the data generating process.
Table A2: Distribution of parameter estimates from simulated data.
Table A3: 95 percent confidence intervals for the reduced-form parameters governing duration dependence and skill depreciation.
Table A4: Maximum likelihood estimation results for the reduced-form parameters governing duration dependence and skill depreciation.
Table A5: Maximum likelihood estimation results for the reduced-form parameters governing duration dependence and skill depreciation.
Figure A1: Interview probabilities under alternative parameterization based on Table A5 vs. relative callback rates in Kroft et al. (2016).
Figure A2: Log-likelihood function and ζ for various values of r(0).
Data S1

REFERENCES

Alvarez, F. E., K. Borovičková and R. Shimer, "Decomposing Duration Dependence in a Stopping Time Model," *Review of Economic Studies* forthcoming.

Andrews, D. W., "Testing When a Parameter is On the Boundary of the Maintained Hypothesis," *Econometrica* 69 (2001), 683–734.

Barnichon, R., and A. Figura, "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics* 7 (2015), 222–49.

Barron, J. M., and J. Bishop, "Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity," *Economic Inquiry* 23 (1985), 363–82.

Doppelt, R., "The Hazards of Unemployment: A Macroeconomic Model of Job Search and Résumé Dynamics," Working Paper, 2016.

Elsby, M. W., B. Hobijn, and A. Şahin, "On the Importance of the Participation Margin for Labor Market Fluctuations," *Journal of Monetary Economics* 72 (2015), 64–82.

Fernandez-Blanco, J., and E. Preugschat, "On the Effects of Ranking by Unemployment Duration," *European Economic Review* 104 (2018), 92–110.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. R. Warren, "Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]," IPUMS, Minneapolis, MN, 2020. https://doi.org/10.18128/D030.V7.0.

Fujita, S., and G. Moscarini, "Recall and Unemployment," *American Economic Review* 107 (2017), 3875–916.

Gonzalez, F. M., and S. Shi, "An Equilibrium Theory of Learning, Search, and Wages," *Econometrica* 78 (2010), 509–37.

Gregory, V., G. Menzio, and D. Wiczer, "The Alpha Beta Gamma of the Labor Market," Working Paper 28663, National Bureau of Economic Research, 2021.

Hall, R. E., and S. Schulhofer-Wohl, "Measuring Job-Finding Rates and Matching Efficiency with Heterogeneous Job-Seekers," *American Economic Journal: Macroeconomics* 10 (2018), 1–32.

Jarosch, G., and L. Pilossoph, "Statistical Discrimination and Duration Dependence in the Job Finding Rate," *Review of Economic Studies* 86 (2019), 1631–65.

Kospentaris, I., "Unobserved Heterogeneity and Skill Loss in a Structural Model of Duration Dependence," *Review of Economic Dynamics* 39 (2021), 280–303.

Kroft, K., F. Lange, and M. J. Notowidigdo, "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment," *Quarterly Journal of Economics* 128 (2013), 1123–67.

———, ———, and ———, and L. F. Katz, "Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation," *Journal of Labor Economics* 34 (2016), S7–S54.

PISSARIDES, C. A., "Loss of Skill during Unemployment and the Persistence of Employment Shocks," *Quarterly Journal of Economics* 107 (1992), 1371–91.

SILVA, J. I., and M. TOLEDO, "Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment," *Macroeconomic Dynamics* 13 (2009), 76–96.

VAN DEN BERG, G. J., "Duration Models: Specification, Identification and Multiple Durations," *Handbook of Econometrics* 5 (2001), 3381–460.