Estimating a search equilibrium model of statistical discrimination: racial wage and employment disparities in the US*

Bruno Decreuse †   Linas Tarasonis ‡

December 2014

Abstract

In the US, black workers spend more time in unemployment, lose their jobs more rapidly, and earn lower wages. This paper offers a model of statistical discrimination in an equilibrium search and matching environment that accommodates these facts. We build on Moscarini (2005) seminal work and extend it to allow for two types of workers and hiring discrimination. We estimate the model on US employment and wage data. Identification strategy relies on the heterogenous impacts of statistical discrimination along the wage distribution. Residual employment and wage disparities are decomposed in ex-ante differences in prior distributions about match quality and precision of output signals vehicled by match duration.

Keywords: employment and wage differentials, discrimination, job search.

JEL codes: J31; J64; J71.

*We thank participants to the GEDIM workshop on gender and ethnic discrimination and the AMSE-Banque de France conference that both took place in Aix-en-Provence (2014). Bruno Decreuse thanks the Institut Universitaire de France for financial support.

†Aix-Marseille School of Economics (Aix-Marseille University), EHESS & CNRS, bruno.decreuse@univ-amu.fr, `https://sites.google.com/site/brunodecreuseecon/`

‡Aix-Marseille School of Economics (Aix-Marseille University), EHESS & CNRS, linas.tarasonis@univ-amu.fr, `http://linas.econ.lt/`

1 Introduction

Black workers earn lower wages, stay unemployed longer and lose their jobs more rapidly. Though a substantial part of such disparities can be attributed to heterogeneity, residual differences seem to be large and persist over time. In their survey on discrimination theories, Lang and Lehmann (2012) summarize the evidence as follows: the residual wage gap is about 10 percentage points and the duration of unemployment is about 25% longer for Blacks. These features have motivated a body of research devoted to the estimation of structural models of search unemployment with discrimination. Such models put additional structure to the relationships between discrimination motives, worker heterogeneity and economic outcomes for the different demographic groups. This allows discrimination to be identified from worker unobserved heterogeneity.

All the efforts have been made on taste-based models of discrimination: Black (1995), Eckstein and Wolpin (1999), Bowlus and Eckstein (2002), Borowczyk-Martins et al. (2014), for racial discrimination, but also Flabbi (2010a,b) for gender discrimination. Following Becker (1971), these papers assume that some employers have a distaste for minority workers. Identification relies on the fact that not all employers discriminate. Thus the wage distribution for discriminated workers is a mix of two partial distributions, for workers in discriminating and non-discriminating jobs. This combined with parametric assumptions on the wage distributions ensure identification. The general result is that discrimination can account for a very substantial share of racial wage and employment gaps.

Taste-based search discrimination models are not entirely convincing. Lang and Lehmann (2012) note these models estimate that the proportion of prejudiced employers is very large, typically around 50%, and that these employers are very strongly prejudiced. To quote them: "We take the evidence from the surveys and the Implicit Association Tests as suggesting that credible models of discrimination based on prejudice may rely on the presence of strong prejudice among a relatively small portion of the population and/or weak prejudice among a significant fraction of the population, but not on widespread strong prejudice. It does not seem likely that a large proportion of employers, for example, are willing to forego significant profits in order to avoid hiring blacks." Moreover, discrimination attitudes have dramatically fallen over time in the US, whereas employment gaps have remained stable, and wage gaps have slightly decreased.

In this paper, we follow a different route. We develop and estimate the structural parameters of an equilibrium search model of statistical discrimination. This type of discrimination arises in the labor market when employers believe that minority workers are less productive (Arrow, 1973) or when employers have greater difficulty assessing their productivity despite they have similar observable characteristics (Phelps, 1972). Conditional on employers' beliefs, statistical discrimination is a rational response to nonidentical distributions of unobserved characteristics or to nonidentical output observability among different groups of workers. Thus it is consistent with profit maximization and likely to persist in equilibrium.

Our model of statistical discrimination (Section 2) is an adaptation of Moscarini (2005). He introduces job turnover à la Jovanovic (1984) in the standard equilibrium search unemployment framework. Workers and firms are ex-ante identical. However, there is an unobserved match-specific shock that the worker and the firm gradually learn. Learning is complicated by Gaussian output noise. The model predicts there is endogenous separation. Job loss occurs when the belief on match-specific component becomes lower than a threshold specific to each group. Wage bargaining over the match surplus implies there is a mapping from the ergodic distribution of beliefs to the stationary wage distribution. Under additional restrictions, the wage distribution is single-peaked with a Pareto tail.

We introduce two types of workers in the model, Blacks and Whites. The distribution of ability differs between these groups. Thus employers have different beliefs regarding them. In Moscarini (2005), the initial draw has a unique possible value. Conditional on meeting, the worker is always hired. To account for hiring discrimination, we allow for a continuous support of the initial draw. Then, beliefs evolve like in the initial model. The model predicts job finding, job loss and the wage distribution for both groups of workers.

The model has three sources of differences between Blacks and Whites: observed worker heterogeneity, unobserved heterogeneity, and precision of the output signal vehicled by match duration. Each component has a specific impact on model outcomes. This allows us to disentangle the effects of employers' beliefs from the direct effects of heterogeneity. In particular, observed heterogeneity only affects wages, whereas ex-ante differences in the distribution of employers' beliefs imply that the racial differential in wage quantile tends to decrease with quantile. Moreover, differences in output signal precision imply that black wages tend to be more concentrated at the bottom of the wage scale than white wages. By combining these three sources of ethnic differentials, we can match very heterogenous patterns in the wage distributions.

We estimate the parameters of the model on a cross-section of CPS data (Section

3). We focus on prime-aged low-skilled male workers and compute their job-finding and job-separation rates as well as the residual wage distribution by mean of a Mincer wage regression. We then use indirect inference and target 14 moments that characterize labor market outcomes for both groups: mean job-finding probability, mean job loss probability, five deciles of the unconditional wage distribution, and five deciles of the initial wage distribution. The parameter estimates are such that output observability and the initial belief distribution are both better for Whites than for Blacks. The estimated model reproduces reasonably well differences in job-finding rates, as well as racial differences in initial and unconditional wage distributions. By contrast, it fails to reproduce the job-separation rate differential.

In our data, the overall return to tenure is roughly the same between Blacks and Whites. This is also what our estimates predict. This results from two opposite forces. On the one hand, ex-ante differences in belief distributions imply that the predicted match quality is initially lower for Blacks than for Whites. As jobs last, low quality matches gradually disappear. Belief updating accounts for this fact, and the extent of statistical discrimination decreases. This force tends to raise the black return to tenure above the white one. On the other hand, differences in output observability imply that the black wage is less sensitive to tenure than the white one. This force tends to reduce the black return to tenure. These two forces offset each other and the returns to tenure are roughly the same. Finally, the return to experience is larger for Whites because they enjoy longer tenures.

Some of the key ideas of this paper have been inspired by Fryer Jr et al. (2013). They develop a stylized three-period micro model of statistical discrimination where the employer and the employee learn the value of the match. They emphasize that the return to tenure should be larger for Blacks than for Whites. We make two contributions with respect to theirs. First, we complete their approach by sketching a full model and estimating its parameters. This allows us to quantify the different mechanisms leading to discrimination. Second, in their sample, Mincer regressions imply that the return to tenure is 1.1% larger for Blacks than for Whites. In our data, the return to tenure does not differ across racial groups. This is explained by the lower quality of the output signal for Blacks.

We are not the first to estimate a structural model of statistical discrimination for the labor market. Moro (2003) estimates a model for racial discrimination. The model has complementarities between skilled and unskilled workers and may feature multiple equilibria. Moro discusses identification and equilibrium selection. Gayle and Golan (2011) focus on gender gaps. Statistical discrimination arises because women are less attached to the labor force, which creates turnover costs. Both papers tell important lessons on wage inequality. However, they abstract from search frictions and, therefore, cannot account for racial differences in unemployment duration and job separation.

Beyond their ability to describe unemployment disparities between Blacks and Whites, search and matching models provide an interesting framework to study discrimination. A couple of papers make the intriguing point whereby discrimination can arise in equilibrium despite employers have no taste for discrimination and Blacks and Whites have similar characteristics. In Rosén (1997), match-specific productivity is random and workers have private information about it. If employers discriminate against Blacks, then these workers tend to apply for low-quality matches, thereby creating the type of belief that leads employers to discriminate them. In Mailath et al. (2000), employers can direct their search towards Blacks or Whites, whereas workers make a pre-market investment in skills. If employers do not send offers to Blacks, these workers invest less in human capital, justifying employers' behavior. In Holden and Rosén (2014), match quality is random and workers in a bad match search on the job. As dismissal is costly, employers trapped in a bad match hope that the worker finds another job very rapidly. Now, if employers discriminate against Blacks, these workers find alternative jobs less rapidly, and thus become less attractive to employers. We do not explore the rich possibilities offered by such models. Instead, we stick to the seminal contribution of Phelps (1972) where skills exogenously differ between Blacks and Whites, and where the output signal precision varies across ethnic groups.

The remainder of this paper is organized as follows. Section 2 presents the theoretical model and examines its predictions in terms of employment and wage discrimination. The proofs of the different results lie in the Appendix. Section 3 is devoted to the estimation of the theoretical model. We introduce the data, the estimation method, and comment our estimates. Section 4 concludes.

2 Theory

The model we develop in this section builds on Moscarini (2005) who nests a job matching model à la Jovanovic (1984) into a Mortensen and Pissarides (1994)-type equilibrium search environment. We add three things: there are two (ethnic) groups of workers, they differ in productive type within groups, and the initial signal on match quality is nontrivial.

The economy is populated by a continuum of risk-neutral workers of measure one and a continuum of firms that post vacancies in a labor market with search frictions. Firms are identical. Workers differ in type α and in (ethnic) group $i = B, W$, where B stands for Black and W for White. The measure of each group is m_i, with $m_B + m_W = 1$. The distribution of type in the whole population is group-specific. The probability function and density of such distributions are, respectively, $\Psi_i(\alpha)$ and $\psi_i(\alpha)$.

Unemployed workers receive job offers at rate λ irrespective of their type. Then the firm and the worker decide if they form a match. Employed workers may lose their job: the match is terminated if the match surplus falls to zero or for exogenous reasons at rate δ. The endogenous measure of group-i unemployed is u_i. There is no on-the-job search.

The output of a firm-worker pair is $y_{\alpha\mu} = \alpha\mu$, where μ is match-specific quality. It is imperfectly observed at hiring, and gradually learnt over tenure. When a firm and a worker meet, they draw a common signal p_0 about the average productivity of the match. In Moscarini (2005), this signal takes a single deterministic value. To account for hiring discrimination, we assume the signal is drawn from the nontrivial group-specific distribution G_i^0 with density g_i^0. If they decide to form a match, then they draw the actual match-specific parameter, μ, which can take two values, $\mu_H > \mu_L$, whereas the initial signal determines the probability of being in a good match, $p_0 = \Pr(\mu = \mu_H) = 1 - \Pr(\mu = \mu_L) \in (0, 1)$.

Match productivity is subject to an additional source of idiosyncratic noise. The cumulative output of a match of tenure t follows a Brownian motion with drift $\alpha\mu$ and type-specific variance $\alpha^2\sigma_{Xi}^2$:

$$X_{\alpha it} = \alpha\left(\mu t + \sigma_{Xi} Z_t\right) \sim N(\alpha\mu t, \alpha^2\sigma_{Xi}^2 t), \tag{1}$$

where Z_t is a Wiener process that keeps μ hidden. The variance is log-linear in type. As output is also log-linear in α, the variance-to-output ratio does not depend on α.

After observing flow match output, $dX_{\alpha it}$, firms and workers update their belief with regard to match quality using Bayes' rule. Let $p_{\alpha it}$ be the probability that the match is good. The standard Wonham (1964) result implies that $p_{\alpha it}$ follows a diffusion process:

$$dp_{\alpha it} = \sigma_{pi}(p_{\alpha it}) d\overline{Z}_{\alpha it}, \tag{2}$$

where

$$\sigma_{pi}(p) = p(1-p)\frac{\mu_H - \mu_L}{\sigma_{Xi}} \tag{3}$$

is the diffusion parameter and

$$d\overline{Z}_{\alpha it} = \frac{dX_{\alpha it} - p_{\alpha it}\alpha\mu_H dt - (1 - p_{\alpha it})\alpha\mu_L dt}{\alpha\sigma_{Xi}} \tag{4}$$

is the innovation process, the normalized difference between realized and unconditionally expected flow output. The variable $\overline{Z}_{\alpha it}$ follows a standard Wiener process.

Equations (2)-(4) imply that beliefs depend on worker group and job tenure, but not on worker type. Hereafter we simply neglect index α and denote p_{it}.

2.1 Value functions

Because p determines the future prospects of a match, it serves as a state variable for the value functions. Let $W_{\alpha i}(p)$ denote the discounted total payoffs that an employed worker expects to receive. Also, let $U_{\alpha i}$ denote the worker's value of unemployment and $J_{\alpha i}(p)$ be the value of a firm employing this worker.

The worker's values solve the Hamilton-Jacobi-Bellman (HJB)[1] equations:

$$rU_{\alpha i} = b\alpha + \lambda \int \max\left\{W_{\alpha i}(p) - U_{\alpha i}, 0\right\} g_i^0(p)dp, \tag{5}$$

$$rW_{\alpha i}(p) = w_{\alpha i}(p) + \frac{1}{2}\sigma_{pi}^2(p)W''_{\alpha i}(p) + \delta[U_{\alpha i} - W_{\alpha i}(p)], \tag{6}$$

where $b\alpha$ is unemployment income and $w_{\alpha i}(p)$ is the wage.

The value of opening a vacancy is set to zero. In the Appendix, we deduce this condition from an enriched model where there is a standard constant return to scale matching function, entry is costly, and firms are free to enter. The value of a filled job with belief p solves the following HJB equation:

$$rJ_{\alpha i}(p) = \alpha\overline{\mu}(p) - w_{\alpha i}(p) + \frac{1}{2}\sigma_{pi}^2(p)J''_{\alpha i}(p) - \delta J_{\alpha i}(p), \tag{7}$$

where $\overline{\mu}(p) \equiv p\mu_H + (1 - p)\mu_L$ is the expected flow output.

2.2 Equilibrium wages and reservation values

The equilibrium wage is pinned down by a generalized Nash bargaining solution, giving the worker a fraction $\beta \in [0, 1]$ of total match surplus:

$$W_{\alpha i}(p) - U_{\alpha i} = \beta S_{\alpha i}(p), \tag{8}$$

where $S_{\alpha i}(p) \equiv W_{\alpha i}(p) - U_{\alpha i} + J_{\alpha i}(p)$ is the total match surplus.

[1]See Appendix B for more details on the derivation of the HJB equations.

Using (6) and (7), we can rewrite the total surplus as the following differential equation:

$$S_{\alpha i}(p) = \frac{\alpha\overline{\mu}(p) + \frac{1}{2}\sigma_{pi}^2(p)S''_{\alpha i}(p) - rU_{\alpha i}}{r+\delta}, \tag{9}$$

subject to value matching, $S_{\alpha i}(\underline{p}_{\alpha i}) = 0$, and smooth pasting, $S'_{\alpha i}(\underline{p}_{\alpha i}) = 0$. Following Moscarini (2005) , we solve this differential equation and obtain:

$$S_{\alpha i}(p) = c_{\alpha i} p^{\frac{1}{2}-\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}}(1-p)^{\frac{1}{2}+\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}} + \frac{\alpha\overline{\mu}(p) - rU_{\alpha i}}{r+\delta} \tag{10}$$

where $s_i \equiv (\mu_H - \mu_L)/\sigma_{Xi}$ is the signal/noise ratio. The coefficient $c_{\alpha i}$ and the optimal stopping point $\underline{p}_{\alpha i}$ solve the system of value matching and smooth pasting equations. The total surplus function is an increasing function of belief $p \in [\underline{p}_{\alpha i}, 1]$.

The Nash bargaining solution implies $\beta J''_{\alpha i}(p) = (1-\beta)W''_{\alpha i}(p)$. Using these facts yield a simple expression for the equilibrium wage:

$$w_{\alpha i}(p) = \beta\alpha\overline{\mu}(p) + (1-\beta)rU_{\alpha i}, \tag{11}$$

which is increasing in belief p.

Having solved for the equilibrium total surplus, we can plug the general solution (10) into the value of unemployment (5) using the Nash bargaining solution (8) to obtain the following return to search:

$$rU_{\alpha i} = \frac{(r+\delta)b\alpha + \beta\lambda\int_{\underline{p}_{\alpha i}}\left[(r+\delta)c_{\alpha i}p^{\frac{1}{2}-\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}}(1-p)^{\frac{1}{2}+\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}} + \alpha\overline{\mu}(p)\right]g_i^0(p)dp}{r+\delta+\beta\lambda(1-G_i^0(\underline{p}_{\alpha i}))}. \tag{12}$$

Lemma (Worker type heterogeneity) *The following statements hold for all $p \in [0,1]$ and all $i = B, W$:*

(i) the functions $U_{\alpha i}$, $W_{\alpha i}$, $J_{\alpha i}$ and $w_{\alpha i}$ are proportional to α, i.e. $U_{\alpha i} = \alpha U_i$, $W_{\alpha i}(p) = \alpha W_i(p)$, $J_{\alpha i}(p) = \alpha J_i(p)$, $w_{\alpha i}(p) = \alpha w_i(p)$;

(ii) the belief threshold $\underline{p}_{\alpha i}$ does not depend on worker type, i.e. $\underline{p}_{\alpha i} = \underline{p}_i$.

The Lemma shows that worker type heterogeneity has very simple effects on model outcomes. In particular, α has a log-linear impact on the wage, it does not affect the threshold belief below which workers stay unemployed or lose their job, and the distribution of worker type conditional on worker group is the same in employment and in unemployment. Three assumptions are key for this result. First, output is log-linear in worker type and match quality. Second, the standard deviation of observed cumulative output is proportional to worker type. Finally, unemployment income is also proportional to worker type.

2.3 Belief distribution

Let $g_i(p)$ be the unnormalized density function of the ergodic belief distribution among employed workers of group i. For beliefs below the threshold $\underline{p}_i$, this pdf is zero, i.e. $g_i(p) = 0$ for all $p \in [0, \underline{p}_i)$. For beliefs above the threshold, the Kolmogorov forward equation describes the dynamics of this measure. Imposing stationarity we obtain:

$$0 = \frac{dg_i(p)}{dt} = \frac{d^2}{dp^2}\frac{1}{2}\sigma_{pi}^2(p)g_i(p) + \lambda u_i g_i^0(p) - \delta g_i(p), \tag{13}$$

where $u_i = m_i - \int g_i(p)dp$. The first term balances all flows that are due to learning. The second term describes flows of workers to $g_i(p)$ from unemployment. The last term captures the attrition due to exogenous separations.

The forward equation is subject to two boundary conditions. Moscarini (2005) names the first condition *no time spending at* $\underline{p}_i$:

$$\frac{1}{2}\sigma_{pi}^2(\underline{p}_i)g_i(\underline{p}_i) = 0. \tag{14}$$

As $\sigma_{pi}^2(\underline{p}_i) \neq 0$, this condition implies that $g_i(\underline{p}_i) = 0$. Hence, the flow number of workers at the bottom of the belief distribution must be 0. By continuity of the function g_i, this number must also be very small just above the threshold $\underline{p}_i$.

As for the second boundary condition, we integrate equation (13) over the interval $[\underline{p}_i, 1]$ using the condition $g_i(\underline{p}_i) = 0$. This leads to the following equation that balances total flows out of and into unemployment:

$$\lambda u_i \int_{\underline{p}_i} g_i^0(p)dp = \delta \int g_i(p)dp + \frac{1}{2}\sigma_{pi}^2(\underline{p}_i)g_i'(\underline{p}_i). \tag{15}$$

The first and the second terms of the right-hand side measure the total flow into unemployment due to exogenous and endogenous separation, respectively. Additional algebra allows us to express the measures of unemployed workers of both types as a function of exogenous fundamentals and $\underline{p}_i$:

$$u_i = m_i - \int g_i(p)dp = \frac{\delta m_i + \frac{1}{2}\sigma_{pi}^2(\underline{p}_i)g_i'(\underline{p}_i)}{\delta + \lambda \int_{\underline{p}_i} g_i^0(p)dp}. \tag{16}$$

Equation (16) defines $g_i'(\underline{p}_i)$. This condition is used in the numerical implementation of the model.

Let $\tilde{g}_i$ be the density of the ergodic distribution among employed workers. We have $\tilde{g}_i(p) = g_i(p)/\int_{p\geq\underline{p}_i} g_i(p)dp$. Similarly, $\tilde{G}_i(p) = G_i(p)/\int_{p\geq\underline{p}_i} g_i(p)dp$.

2.4 Negative stereotyping

The underpinning assumption here is that employers hold negative (rational) beliefs about the quality of black workers. This assumption can be incorporated by making the group-W distribution of initial signal $G_W^0(p)$ stochastically dominate the group-B distribution.

Assumption 1 (Restricted Stochastic Dominance): *The distribution $G_W^0(p)$ stochastically dominates the distribution $G_B^0(p)$ at first order*

$$G_B^0(p) \geq G_W^0(p), \, \forall p \in [0,1]$$

and

$$\exists p \in [\underline{p}_W, 1] \, s.t. \, G_B^0(p) > G_W^0(p).$$

Moreover, $\sigma_{XB} = \sigma_{XW} = \sigma_X$.

It implies that the variance of the belief updating process is the same across groups, which we simplify to $\sigma_{pW}^2(p) = \sigma_{pB}^2(p) = \sigma_p^2(p)$. The direct consequence of Assumption 1 is the following proposition.

Proposition 1 (Outside Option Differential I). *Under Assumption 1, the per-type value of unemployement of group-W workers is higher than that of group-B workers, that is, $U_W > U_B$.*

The proof of Proposition 1 is stated in the Appendix. This result is entirely due to between-group heterogeneity. As matches with Black workers are less productive on average, wage and employment expectations are poorer for Blacks. This translates into lower returns to search.

We first examine employment discrimination. The model predicts the ethnic-specific job-finding rate jfr_i, and the ethnic-specific job loss rate jlr_i, $i = B, W$. Their combination affects unemployment odds for both groups.

Proposition 2 (Employment Discrimination I). *Under Assumption 1, the following statements hold:*

(i) $\underline{p}_W > \underline{p}_B$,

(ii) the job-finding rate differential is

$$jfr_W - jfr_B = \lambda \left[G_B^0(\underline{p}_B) - G_W^0(\underline{p}_W) \right], \tag{17}$$

(iii) the job-loss rate differential is

$$jlr_B - jlr_W = \frac{1}{2} \left[\frac{\sigma_p^2(\underline{p}_B) g_B'(\underline{p}_B)}{\int g_B(p) dp} - \frac{\sigma_p^2(\underline{p}_W) g_W'(\underline{p}_W)}{\int g_W(p) dp} \right]. \tag{18}$$

The model yields ambiguous employment outcomes. Part (i) thus shows that White workers are submitted to hiring discrimination. The cause is given by Proposition 1. As the return to search is lower for Blacks, the match surplus conditional on belief p is always larger for Blacks than for Whites. Thus a belief threshold sustaining a positive surplus will be lower for Blacks, i.e. $\underline{p}_W > \underline{p}_B$. It follows that Blacks will be negatively selected with respect to Whites, and the expected match quality is higher for Whites.

Taste-based models of discrimination do not necessarily predict that the selection into employment of one type is necessarily stronger than the other. In such models, a fixed share of employers have a distaste for Black workers. They only hire Blacks when the match productivity is very large. This implies that Blacks lose a number of job opportunities, which negatively affects their outside options. Blacks, therefore, are cheaper, and unprejudiced employers take advantage of that. They tend to hire Blacks in low productivity matches as a result. Thus it implies positive selection with prejudiced employers and negative selection with unprejudiced ones: the overall effect is ambiguous.

Part (ii) displays the job-finding rate differential. It is equal to the ethnic-independent job offer rate λ multiplied by the difference in probability mass at the hiring threshold. As initial belief distributions and hiring thresholds differ across types, job-finding rates differ across types and the differential is generally nonzero. Intuition suggests this differential is positive. However, part (i) implies this is actually ambiguous. On the one hand, first-order stochastic dominance implies that $G_B^0(p) > G_W^0(p)$ at given p. On the other hand, Blacks are negatively selected than Whites and the threshold $\underline{p}_W > \underline{p}_B$.

Part (iii) features a similar result. The job loss differential is non-trivial because there is endogenous separation. On the one hand, blacks are less selected, which implies they are more often employed in low quality matches. As these jobs must be destroyed at some finite time, Blacks tend to lose their jobs faster. On the other hand, the match surplus with such workers is larger. This lengthens expected job duration at any belief. The learning process implies that at any time there is a flow of employees who cross the belief threshold $\underline{p}_i$ and lose their job. The actual measure of this flow depends on (half of) the variance of the learning process σ_p^2 and on the slope of the density of the ergodic belief distribution, both evaluated at the belief threshold. Both components can differ across groups, and so the model can predict heterogenous job loss rates.

We now turn to wage discrimination. We distinguish two wage differentials. Let $\omega_i = w_{\alpha i}/\alpha$ be the efficient wage, i.e. the wage by type unit. We define $\Delta\omega(p) \equiv \omega_W(p) - \omega_B(p)$ as the ethnic efficient wage differential conditional on belief p. Let also F_i denote the group-i-specific efficient wage distribution, whereas G_i is the distribution of beliefs. We have

$$\Pr[\omega_i \leq \omega \mid i] = F_i(\omega). \tag{19}$$

Finally, let $\omega_{iq} = F_i^{-1}(q)$ be the q-th quantile of this distribution, whereas $z(q) \equiv \omega_{Wq} - \omega_{Bq}$ is the difference in quantile between Whites and Blacks.

Proposition 3 (Wage Discrimination I). *Under Assumption 1, the following statements hold:*

(i) for all $p \in [0, 1]$,

$$\omega_W(p) - \omega_B(p) = (1 - \beta)r(U_W - U_B) > 0, \tag{20}$$

(ii) for all $q \in [0, 1]$,

$$z(q) = \beta(\mu_H - \mu_L)\left[G_W^{-1}(q) - G_B^{-1}(q)\right] + (1 - \beta)r\left(U_W - U_B\right), \tag{21}$$

is such that $z(0) > z(1) > 0$.

Part (i) refers to the theoretical definition of discrimination. In our environment, two workers with the same type are similar if they hold a job with the same probability of being of good quality. Thus the (unobserved) p must be the same. Given Blacks and Whites only differ in G_0, the wage differential at given belief p simply reveals the outside option differential. Proposition 1 shows that the return to unemployment is higher for Whites than for Blacks. Thus, all things equal (i.e. at given type and belief), Whites are better paid than Blacks.

Part (ii) compares the two wage distributions by focusing on their different quantiles. The difference in quantile is composed of two terms. The first term is the difference in belief quantiles. It captures the difference in belief distributions. The second term is the difference in outside options. The first term tends to decrease with q. High quantiles of the distribution correspond to high-quality matches. Thus p is close to one in both groups, and the wage differential simply reflects the outside option differential. At low quantiles, we find workers who were just hired with an initial belief close to the hiring threshold, and more generally workers who are about to be dimissed. As the threshold is lower for Blacks, the difference in belief quantiles must be positive. Thus $z(0) > z(1)$.

The results of Proposition 3 hold for the efficient wage distribution. These wage differences are compatible with additional between-group heterogeneity due to different distributions of type in each group.

Overall ex-ante differences in prior distributions can predict that Blacks spend more time in unemployment, less time in employment, and that they earn lower wages. They also predict that the racial wage differential is larger at the bottom of the wage distribution than at its top.

2.5 Screening discrimination

Beyond type heterogeneity, we now suppose the only difference between Blacks and Whites is due output signals are more noisy for Blacks than for Whites. Meanwhile prior distributions coincide.

Assumption 2 (Restricted output volatility): *The standard deviation of output is larger for Blacks than for Whites, i.e.* $\sigma_{XB} > \sigma_{XW}$*, and* $G^0_B = G^0_W = G^0$*.*

Proposition 4 (Outside Option Differential II). *Under Assumption 2, the per-type value of unemployement of group-W workers is higher than that of group-B workers, that is,* $U_W > U_B$*.*

As information on match quality is poorer for Blacks than for Whites, the match surplus is lower for Blacks at given belief on match quality. Consequently, and like the previous case, Blacks have a lower return to search.

Proposition 5 (Employment Discrimination II). *Under Assumption 2, the following statements hold:*

(i) $\underline{p}_W < \underline{p}_B$,

(ii) the job-finding rate differential is

$$jfr_W - jfr_B = \lambda \left[G^0(\underline{p}_B) - G^0(\underline{p}_W) \right] > 0, \tag{22}$$

(iii) the job-loss rate differential is

$$jlr_B - jlr_W = \frac{1}{2} \left[\frac{\sigma^2_{pB}(\underline{p}_B) g'_B(\underline{p}_B)}{\int g_B(p)dp} - \frac{\sigma^2_{pW}(\underline{p}_W) g'_W(\underline{p}_W)}{\int g_W(p)dp} \right]. \tag{23}$$

Part (i) shows that, unlike the previous case, in the presence of screening discrimination Blacks are positively selected with respect to the Whites. The job tenure does not provide much information on the match quality when the job is occupied by a black worker. Thus employers rely more on the initial signal. This result also implies part (ii): as Blacks are more selected, they are less likely to form matches.

As for part (iii), the resulting job-loss rate differential has an ambiguous sign. Like the previous case, this sign depends on the variance of the learning process σ_p^2 and on the slope of the density of the ergodic belief distribution at the belief threshold. Note, however, that Assumption 2 implies that the former is larger for Whites at any given p, $\sigma_{pB}^2(p) < \sigma_{pW}^2(p)$. This tends to increase *ceteris paribus* the separation rate of Whites. The overall sign of the differential is ambiguous.

Proposition 6 (Wage Discrimination II). *Under Assumption 2, the following statements hold:*

(i) for all $p \in [0,1]$,

$$\omega_W(p) - \omega_B(p) = (1-\beta)r(U_W - U_B) > 0, \tag{24}$$

(ii) for all $q \in [0,1]$,

$$z(q) = \beta(\mu_H - \mu_L)\left[G_W^{-1}(q) - G_B^{-1}(q)\right] + (1-\beta)r\left(U_W - U_B\right) \tag{25}$$

is such that $z(1) > 0$ *and* $z(0) < z(1)$.

Part (i) shows that the wage differential conditional on belief p is positive again. It reflects the outside option differential, which is positive by Part (i) of Proposition 4. At all match quality, black workers pay the price of lower output observability. Part (ii) reproduces the formula of the wage quantile differential. Unlike the previous case, it tends to increase with quantile. The difference in wage quantile is positive at the top, whereas its sign is ambiguous at the bottom. Blacks are more selected than Whites. The min probability of having a good match is thus higher for Blacks. This is reflected by the min of the two wage distributions.

To summarize, differences in output observability can explain racial wage and employment disparities. They also predict that wage differentials are higher at the top than at the bottom of the wage distribution.

Illustration.—We suppose that $\sigma_{XB} = \infty$ and job tenure does not provide any information. Meanwhile, $\sigma_{XW} = 0$, and match quality is revealed right after hiring.

When $\sigma_{XB} = \infty$, the diffusion parameter $s_B = 0$. Thus the standard deviation $\sigma_{pB}(p) = 0$ for all $p \in [0,1]$. It follows that the belief on match quality does not change over tenure. The initial wage does not change with tenure and job separation only occurs for exogenous reasons.

The match surplus is equal to

$$S_B(p) = \frac{\bar{\mu}(p) - rU_B}{r+\delta}, \tag{26}$$

whereas the return to search is

$$rU_B = b + \beta\lambda \int_0^1 \max\left\{\frac{\bar{\mu}(p) - rU_B}{r+\delta}, 0\right\} g^0(p)dp. \tag{27}$$

When $\sigma_{XW} = 0$, match quality is immediately revealed. Thus the belief jumps to $p = 1$ if $\mu = \mu_H$ or $p = 0$ if $\mu = \mu_L$. In the former case, the worker keeps the job until exogenous separation occurs. In the latter case, the worker immediately quits the job and searches for another one.

As information acquisition is instantaneous, all contacts lead to matches. Thus $\underline{p}_W = 0 < \underline{p}_B$. The return to search is

$$rU_W = b + \beta\lambda \int_0^1 pS_W(1)g^0(p)dp. \tag{28}$$

For a given rU, we have $\bar{\mu}(p) - rU < p\mu_H - rU$. It follows that $rU_W > rU_B$.

The transition rates are

$$jfr_W = \lambda > jfr_B = \lambda[1 - G^0(p_B)], \tag{29}$$

$$jlr_W = \lambda \int_0^1 (1-p)g^0(p)dp + \delta > jlr_B = \delta. \tag{30}$$

In this particular case, jobs last longer for Blacks.

Wages are constant for Blacks. Thus the wage distribution coincides with the initial wage distribution. As for Whites, the initial wage distribution covers the unconditional prior distribution. However, such wages do not last: right after hiring only good matches survive and pay $\beta\mu_H + (1-\beta)rU_W$.

3 Estimation

We use worker-level data from the 2010 Current Population Survey (CPS). We merge the Monthly Outgoing Rotation Groups (MORG) with the Basic Monthly (BM) extracts, thus obtaining information on individual wages and transition rates across employment and unemployment. Our sample consists of individuals between 18 and 65 who declare themselves to be either black or white only. We restrict our sample to males without a college education.

We use simulation methods to estimate the structural parameters of the model. Following Gourieroux et al. (1993), we estimate the model by indirect inference. It consists of a simulated minimum distance estimator, in which some of the moments the estimator seeks to match are parameters from reduced-form auxiliary models that capture important aspects of the data-generating process. In the following paragraph we describe the mechanics of the estimator.

Let $\boldsymbol{\theta}$ denote the vector of structural parameters (to be specified in the next subsection), $\hat{m}^S(\boldsymbol{\theta})$ denotes the model-generated vector of parameters of the auxiliary models and $\hat{m}$ its empirical counterpart. The estimation procedure finds $\boldsymbol{\theta}$ such that the distance between the model-generated moments and their empirical counterparts is as small as possible, according to the following criterion function:

$$L_N(\boldsymbol{\theta}) = \frac{1}{2}\left(\hat{m} - \hat{m}^S(\boldsymbol{\theta})\right)'^{-1}\left(\hat{m} - \hat{m}^S(\boldsymbol{\theta})\right). \tag{31}$$

Theoretical moments are obtained by simulating the model. In this procedure, we give equal weight to all moments.

3.1 Econometric specification

To estimate the model developed in Section 2, we make parametric assumptions and calibrate a number of parameters. The population shares of worker types, m_W and m_B, are observed (92% and 8%, respectively). The monthly discount rate, ρ, is set at 0.0043 (equivalent to 5% per annum). Workers' bargaining power, β, is set to 1/2.

We assume that $G_B^0(p)$ and $G_W^0(p)$ have a truncated log-normal distribution with bounds $[0,1]$. For all $p \in [0,1]$ and $i = B, W$,

$$G_i^0(p) = \frac{\frac{1}{p\eta_i\sqrt{2\pi}} e^{-\frac{(\ln p - \gamma_i)^2}{2\eta^2}}}{\Phi\left(\frac{-\gamma_i}{\eta_i}\right)}, \tag{32}$$

where γ_i and η_i are, respectively, the log-scale and shape parameters of the distribution, and ϕ and Φ are, respectively, the pdf and the cdf of the standard normal distribution.

We are left with the following vector of eleven parameters to estimate:

$$\boldsymbol{\theta} = \{\gamma_W, \gamma_B, \eta_W, \eta_B, \mu_H, \mu_L, \sigma_{XW}, \sigma_{XB}, b, \lambda, \delta\}. \tag{33}$$

3.2 Auxiliary models

Our model predicts ethnic-specific labor market transition probabilities and wage distributions. We use these different predictions as auxiliary models in our estimation procedure.

There are four different monthly probabilities of transition across labor market states: the average probability of transitions from job to unemployment, $JTU_i(\boldsymbol{\theta})$, and the average rates of transition from unemployment to job, $UTJ_i(\boldsymbol{\theta})$, $i = B, W$. As the job-finding and job destruction rates do not depend on worker type, we have:

$$JTU_i(\boldsymbol{\theta}) = 1 - \exp\left(-\delta - \frac{\frac{1}{2}\sigma_{pi}^2(\underline{p}_i)g_i'(\underline{p}_i)}{\int g_i(p)dp}\right) \quad (34)$$

$$UTJ_i(\boldsymbol{\theta}) = 1 - \exp\left(-\lambda[1 - G_i^0(\underline{p}_i)]\right) \quad (35)$$

The average probability that an employed worker becomes unemployed, $JTU_i(\boldsymbol{\theta})$, has an exogenous component, δ, which is assumed to be common to both races. It also has an endogenous component depending on the flow number of employed workers who reach the belief threshold $\underline{p}_i$ and subsequently lose their job. When the product of the variance of the belief process at belief threshold $\sigma_p^2(\underline{p}_i)$ by the normalized density $g_i'(\underline{p}_i)/(1 - u_i)$ is larger for Blacks than for Whites, this flow number is larger for Blacks and they are more likely to lose their jobs.

Similarly, the average probability that unemployed workers find jobs, $UTJ_i(\boldsymbol{\theta})$, is governed by the product of the exogenous job meeting rate λ common to both groups and the probability of drawing a prior associated with positive surplus, $1 - G_i^0(\underline{p}_i)$. The difference in initial belief distributions is the only source of racial gap in the job-finding rate.

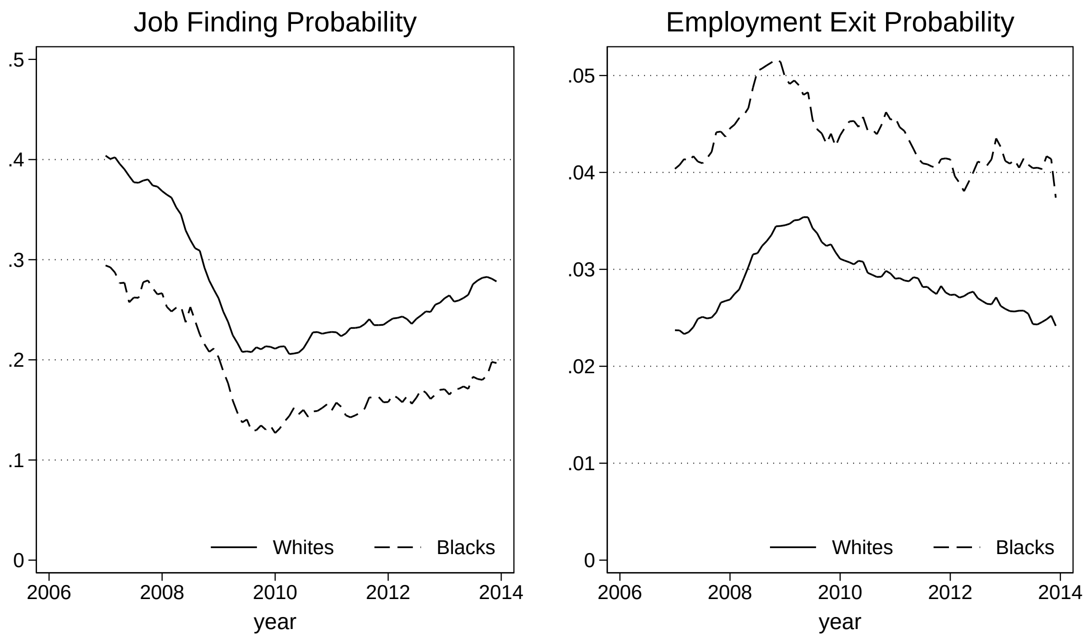
To measure these probabilities, we follow Shimer (2012). The details are explained in Appendix A. To make sure that the estimated differences in the job-finding and the employment exit probabilities are not due to the particular phase of the business cycle, we calculate them from January 2007 to December 2013.

The results, shown in Figure 1, portray striking racial gaps in both probabilities. The probability that an unemployed black worker finds a job is on average 30% lower compared to his white counterpart over the observed period. Similarly, black employees are 1.5 times more likely to lose their job. The magnitude of ethnic gaps in transition probabilities are consistent with longer unemployment duration and higher unemployment rates for blacks as reported in Lang and Lehmann (2012).

Figure 1 also show that racial gaps in transition probabilities are relatively stable over the business cycle. The only exception is the job-finding probability of black workers, which takes slightly longer to react to a change in business cycle conditions at the beginning of the Great Recession.

As for wages, equation (11) implies that there is a one-to-one mapping of posterior beliefs to equilibrium wages by type. Thus, the observed distribution of wages by

Figure 1: The ins and outs of unemployment, prime-age men with no college.



Notes: Prime-age men with no college, 2007m1-2013m12, 12 months moving average of monthly data.
Source: Current Population Survey and authors' calculations.

type reveals the equilibrium distribution of beliefs. In particular, the initial wage distribution reveals the prior distribution.

Following Lemma 1, we need to purge wage data from worker type heterogeneity α. We estimate Mincer wage regressions for the two groups. CPS data do not allow us to estimate worker fixed effects. Thus we estimate the following cross-sectional regression:

$$\ln w_j = A_j' d + a^0 + a^1 Black_j + a^2 \tau_j + a^3 Black_j . \tau_j + e_j, \tag{36}$$

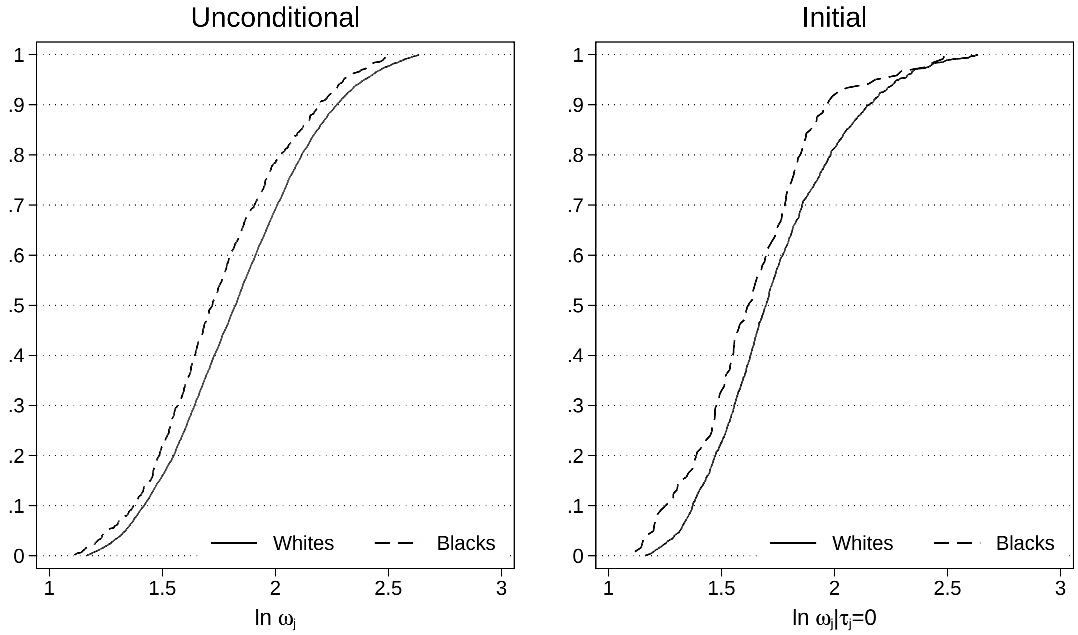
where A_j is a vector of individual characteristics, $Black$ is a dummy variable equal to one when the individual is black, τ_j is tenure, and e_j is the error term. The parameters a^2 and a^3 measure the apparent race-specific returns to tenure.

We then compute efficient wages as follows:

$$\ln \omega_j = \hat{a}^0 + \hat{a}^1 Black_j + \hat{a}^2 \ln \tau_j + \hat{a}^3 Black_j . \tau_j + \hat{e}_j. \tag{37}$$

Thus we remove observed characteristics and their return from log wage observations. Figure 2 depict the resulting initial and unconditional wage distributions, whereas Table 1 provides their main moments.

Figure 2: Residual uncoditional and initial log wage distributions.



Notes: Prime-age men with no college, 2010.
Source: Current Population Survey and authors' calculations.

Table 1: Summary Statistics of Residual Log Wages

	Mean	S.D.	Min	Max	Obs.
Unconditional					
Whites	1.835	0.319	1.164	2.635	6155
Blacks	1.753	0.309	1.112	2.503	571
Total	1.828	0.319	1.112	2.635	6726
Initial					
Whites	1.734	0.296	1.164	2.635	1325
Blacks	1.642	0.293	1.117	2.487	121
Total	1.727	0.297	1.117	2.635	1446

The mean wage differential is 8 log points for the unconditional wage distribution, and 7 log points for the initial wage distribution. The standard deviation is roughly the same for all wage distributions and amounts to 3 log points. The shape of unconditional wage distributions calls for a combination of differences in initial prior distributions and in output observability. The difference in unconditional wage quantile is indeed nonmonotonic in quantile.

In line with Propositions 3 and 6, we target fives quantiles of the unconditional wage distribution, and similar quantiles of the initial wage distribution. This leaves us with 14 moments for 9 parameters.

3.3 Model fit

Table 2 gives the parameter estimates. We only provide their means. In the future version of this paper, we will also provide their standard deviation obtained by bootstrap.

Table 2: Parameter Estimates

$\hat{\gamma}_W$	$\hat{\gamma}_B$	$\hat{\eta}_W$	$\hat{\eta}_B$	μ_H	μ_L	σ_{XW}	σ_{XB}	b	λ	δ
0.462	0.427	0.124	0.163	22.54	-7.01	395.30	554.67	-5.733	0.199	0.025

The means of the two initial belief distributions are 46.2% and 42.7% for Whites and Blacks respectively.[2] They differ by 3.5 percentage points. Their standard deviations are 12.4% and 16.3% respectively. The volatility of initial beliefs is thus large, meaning that we are far from Moscarini (2005) who assumes there is a single prior common to all new matches. The reason is that wage dispersion at the moment of hiring is as large as the unconditional wage dispersion as shown in Table 1.

The productivity of a good match is about 23, against -7 for the productivity of a bad match. This difference is quite large. To be compatible with non-trivial endogenous separation (i.e. people in a bad match are not immediately fired), output must be noisy. This is why the standard deviation of the output signal is 395 and 555 for Whites and Blacks respectively. Unemployment utility flow is $-5.7 < 0$. It is small compared with the highest wage. What this means is that unemployment income does not affect workers' and firms' hiring and separation decisions. The job-meeting rate common to both groups is close to 0.2, whereas exogenous separation occurs at rate 0.025.

Table 3 compares the model outcomes with the empirical moments chosen for the estimation. Given the small set of potential differences between Blacks and Whites, the model reproduces quite well the characteristics of empirical wage distributions. It is less successful in matching empirical transition rates that we document in Section 3. The job-finding rate differential is underestimated, and there is no difference in estimated race-specific job-separation rates.

Figure 3 depicts the initial and ergodic distributions of beliefs about match quality. The ergodic distributions are slightly more dispersed than the initial distributions. They feature a fat right tail. The threshold belief above which workers are hired and below which they quit the job, $p_W = 0.315$ and $p_B = 0.314$, are also shown. The two

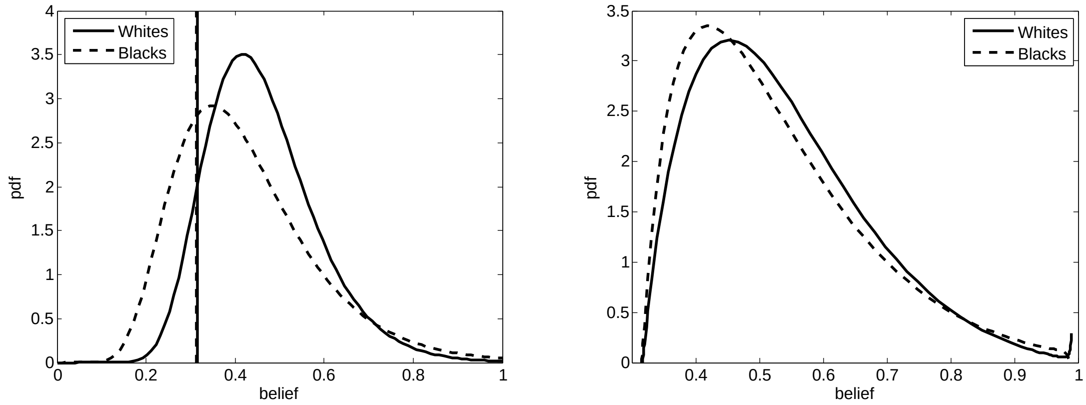
[2]The mean, $\hat{\gamma}_j$, and standard deviation, $\hat{\eta}_j$ of the lognormal distribution with log-scale and shape parameters γ_j and η_j are obtained using $\hat{\gamma}_j = \exp(\gamma_j + \eta_j^2/2)$ and $\hat{\eta}_j = \sqrt{\exp(2\gamma_j + \eta_j^2)(\exp(\eta^2) - 1)}$.

Table 3: Goodness of Fit

		Model	Data	
			Mean	Sd
		(1)	(2)	(3)
Transition Rates				
Unemployment to Job	White	0.165	0.190	–
	Black	0.138	0.125	–
Job to Unemployment	White	0.033	0.029	–
	Black	0.031	0.039	–
Unconditional Wage Distribution				
Min	White	3.335	3.202	–
	Black	2.995	3.040	–
25th percentile	White	5.068	4.944	–
	Black	4.506	4.614	–
Median	White	6.250	6.195	–
	Black	5.673	5.594	–
75th percentile	White	7.770	7.835	–
	Black	7.286	7.092	–
95th percentile	White	10.311	10.957	–
	Black	10.125	10.008	–
Initial Wage Distribution				
Min	White	3.335	3.203	–
	Black	2.995	3.054	–
25th percentile	White	4.509	4.581	–
	Black	3.943	4.301	–
Median	White	5.482	5.460	–
	Black	4.993	5.070	–
75th percentile	White	6.716	6.840	–
	Black	6.503	6.129	–
95th percentile	White	8.964	9.818	–
	Black	9.451	8.855	–

thresholds are almost identical. The two forces driving the differential in opposite directions compensate. On the one hand, negative stereotypes reduce selection for Blacks (Proposition 2 implies $\underline{p}_W > \underline{p}_B$). On the other hand, differential output observability, $\sigma_{XB} > \sigma_{XW}$, increases such selection (Proposition 3 implies $\underline{p}_W < \underline{p}_B$). Meanwhile, the probability mass located at the right side of the threshold is larger for Whites than for Blacks (90.6% versus 74.1%), which explains why the former find jobs faster.

The job loss rate differential is very small and of the opposite sign than in the data.



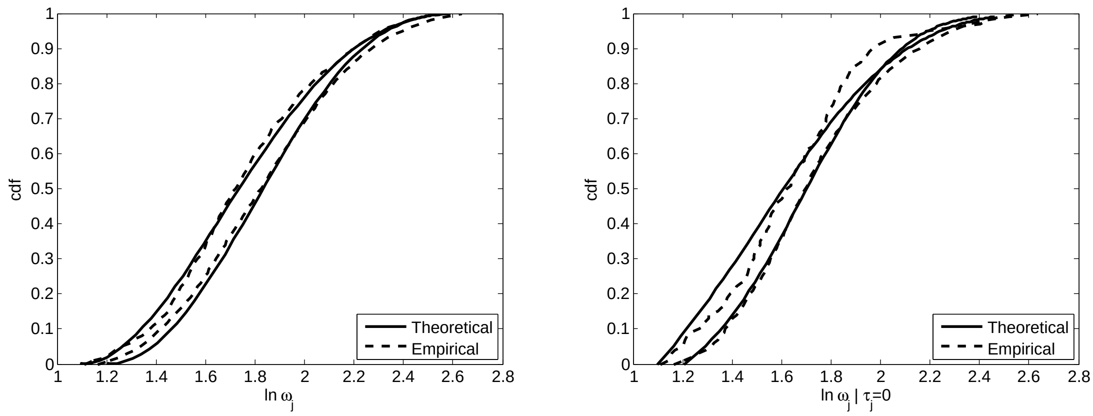
(a) Distributions of prior beliefs (b) Ergodic distributions of beliefs

Figure 3: Prior and Ergodic distributions of beliefs

The reason why Blacks have a smaller job-loss rate than Whites in the estimated model can be inferred from Proposition 2. Part (ii) states that the type-i specific endogenous component of the job destruction rate is $0.5\sigma_p^2(\underline{p}_i)g_i'(\underline{p}_i)/(1-u_i)$. On the one hand, the initial belief distribution is quite concentrated and its mean is close to the threshold belief. Thus many Black workers are initially closed to the threshold belief, which implies that the normalized density of the ergodic distribution is very steep at this threshold, i.e. $g_B'(\underline{p}_B)/(1-u_B)$ is large. On the other hand, a higher output observability noise for Blacks lowers the volatility of the belief updating process. Given that the limit thresholds are almost identical to $\underline{p}_B = \underline{p}_W$, equation (3) implies that $\sigma_p^2(\underline{p}_B) < \sigma_p^2(\underline{p}_W)$. The latter effect appears to be quantitatively more important.

There are two ways to alter this property. First, we can try to modify the different parameters until the endogenous slope $g'(\underline{p})/(1-u)$ becomes much larger for Blacks than for Whites. Second, we can consider race-specific μ_H and μ_L. This would allow us to increase the volatility $\sigma_p^2(\underline{p}_B)$, so that more Blacks would cross the threshold $\underline{p}_B$ and job separation would be larger for them.

Figure 4 shows the predicted and empirical wage quantiles of the different distributions. The figure displays the two cumulative functions for Black and White workers. Beyond the general shape of the different curves, the main comment to make is about the wage differential by wage level. In the raw data that we use, the wage differential is relatively small at the top and at the bottom of wage scale, and it is relatively large in the middle of the distribution. Our model is able to accommodate such regularity using both mechanisms, differences in prior belief distributions and differences in output observability. In the following section we provide more intuition on the mechanisms at work.



(a) Unconditional (b) Initial

Figure 4: Fit of the Wages

3.4 Counterfactuals

We examine two counterfactual scenarios. In the first one, output noise is the same for Blacks and Whites. Thus we keep all parameter estimates but one: $\sigma_{XW} = \hat{\sigma}_{XB}$. In the second one, the initial prior distribution is the same for Blacks and Whites. Thus we impose $\gamma_W = \gamma_B$ and $\eta_W = \eta_B$. In both cases we neglect equilibrium effects that transit through a change in market tightness and associated job contact rate (see Appendix D).

Table 4 shows the results of the simulations. Columns 1 and 2 display key moments of endogenous objects in the unrestricted estimation. Column 3 displays similar moments when we allow for differences in prior beliefs only. In column 4, differences in output observability is the only source of differences between the two groups. Changing black-specific deep parameters does not affect Whites' outcomes because of the assumption of constant exogenous contact rates. Finally, Figure 5 depicts the simulated unconditional log wage distributions for the unrestricted case and both counterfactual scenarios.

Same output observability. The results in column 3 of Table 4 show that differences in output observability contribute to generating low job-finding rates of Blacks. This is due to the selection effect established in Proposition 5. Also, higher output noise for Blacks result in a lower variance of belief updating process, which dampens the amount of endogenous separations. Shutting it down increases the Black-specific

	Whites	Blacks		
		Unrestricted	$\sigma_{XW} = \sigma_{XB}$	$\eta_W = \eta_B$ $\gamma_W = \gamma_B$
	(1)	(2)	(3)	(4)
Transition Rates				
Unemployment to Job	0.165	0.138	0.145	0.160
Job to Unemployment	0.033	0.031	0.034	0.031
Wage Distribution				
Unconditional				
Mean	6.575	6.088	6.284	6.373
S.D.	1.941	2.026	2.202	1.764
Initial				
Mean	5.667	5.350	5.261	5.706
S.D.	1.562	1.850	1.885	1.529

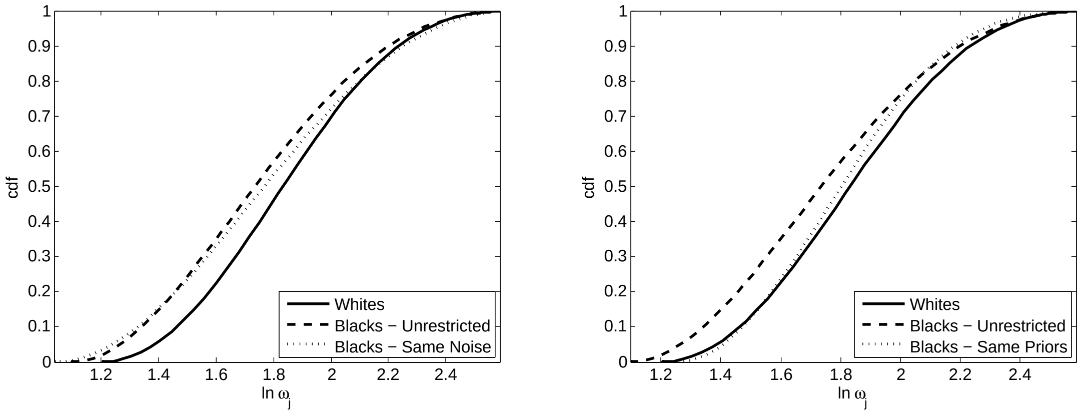
Table 4: Counterfactual Scenarios

separation rate from 0.31 to 0.34.

The bottom panel of Table 4 shows that differences in output observability explains 40% of the total average racial wage gap. This result is in line with Proposition 6. Figure 5a reveals more details of the contribution of this channel to wage differentials. Shutting down the difference in output observability increases differences in quantiles between Whites and Blacks at the bottom of the distribution.

Same initial prior distribution. The upper panel of column 4 of Table 4 reveals that differences in job-finding rates are almost entirely driven by the differences in initial prior distribution. Switching off this channel almost equalizes the two rates. Also, stereotypes explain about 60% of the unconditional average racial wage gap.

Figure 5a portrays the impact of differential output observability on the unconditional wage distribution. In essence, higher output noise for Blacks results in employers valuing significantly less individual signals. This translates into anti-clockwise rotation of the unconditional wage distribution. As a consequence, the difference in quantiles between Whites and Blacks at the bottom of the distribution becomes slightly negative, wheras the gap tends to increase with quantile.



(a) Same Output Noise (b) Same Initial Priors

Figure 5: Simulated Log Wages

3.5 Returns to tenure

We conclude this Section by discussing the predicted returns to tenure for both ethnic groups. Our discussion follows Fryer Jr et al. (2013). They develop a stylized three-period model of statistical discrimination. The model predicts that Black and White wages converge with tenure. They also provide evidence in favour of this thesis. They use a recent dataset from the Princeton University Survey Research Center and estimate that the return to tenure is larger for Blacks by 1.1 log points.

Figure 6 reports the predicted log mean wage differential by tenure. It is built from the simulation of 10,000 individual trajectories, with a 50-50 Black-White divide. The curve is smoothed to avoid artificial fluctuations due to sample attrition. The horizontal line is the minimum log wage differential between the two groups. It reflects the racial differential in the return to search.

The mean wage differential is nonmonotonic with tenure. It initiallly amounts to 4 log points, 1.5 log points above the minimum log wage differential. It then increases over tenure, reaches 7 log points slightly before 15 years, and then falls. As in Fryer Jr et al. (2013), Black and White wages do converge, but the process is nonmonotonic and excessively slow.

Our model with initial prior differences broadly capture Fryer Jr et al. (2013) argument. Proposition 3 shows that the wage quantile decreases with quantile, which is mostly due to the fact that Blacks have larger returns to tenure through selection. However, CPS data do not feature such a monotonic decline in wage quantiles. This is why we estimate that output observability differs across groups. Blacks become as selected as Whites, learning on match quality is very slow when the job is occupied

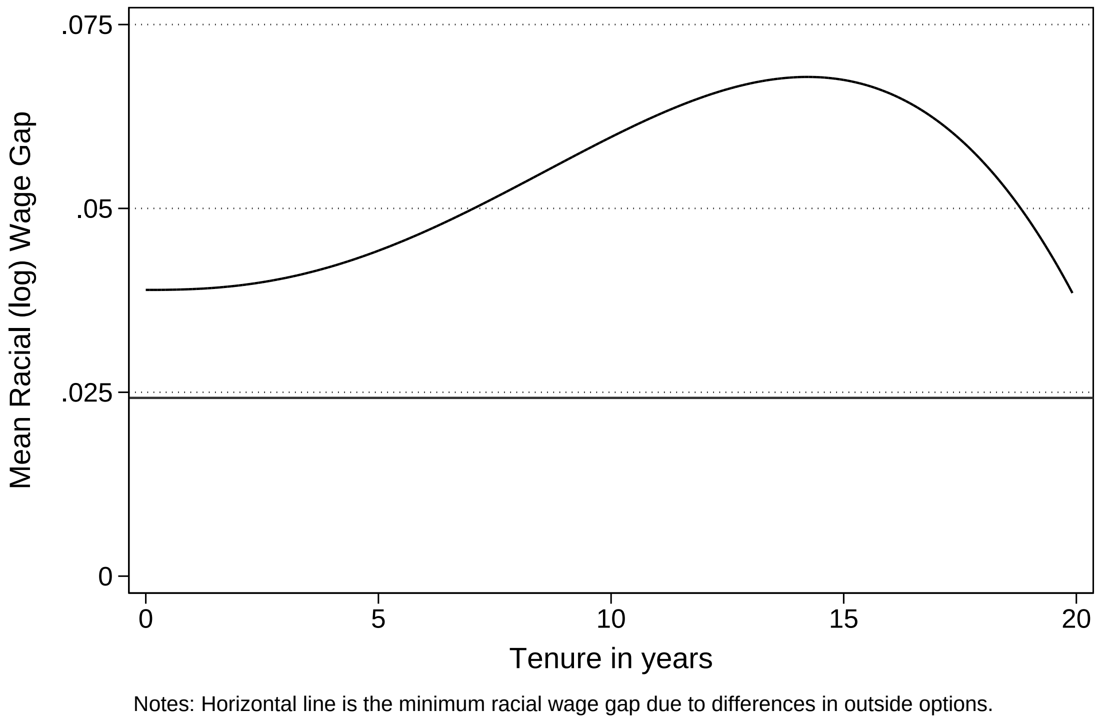


Figure 6: Racial Gap - Tenure profile

by a Black worker. Both properties imply that the wage gap is initially positive and increasing over tenure. After 15 years, most of the possible gains have been achieved for Whites and wages do not increase much more after this date. The learning process being slower for Blacks, such workers still gain more with tenure and the wage gap declines.

4 Conclusion

This paper is the first one to estimate the structural parameters of an equilibrium search model of statistical discrimination. We consider a variant of Moscarini (2005) who embeds a learning model à la Jovanovic (1984) into the canonical DMP model of equilibrium search unemployment. In this model firms have a priori beliefs on match quality, and output observability is complicated by a noise that follows a Brownian motion. The model generates a stationary distribution of beliefs on match quality. This distribution translates into a stationary distribution of wages with empirically appealing properties. We extend this model to two sub-populations, Blacks and Whites, and we add more initial belief heterogeneity to account for hiring discrimination. We show how the model can predict wage and employment outcomes for both groups. We then estimate it on CPS data by indirect inference.

We see several avenues for research. A first extension would introduce on-the-job search. Such a model would enrich our discussion of the returns to tenure and experience. In the same vein, a second extension would focus on human capital investments. Finally, different employers might have different abilities in signal extraction. For instance, Black employers could be more able than Whites to detect good matches when the employee is black.

A Measuring the ins and outs of unemployment

To measure the probability that an employed worker becomes unemployed, JTU_t, and the probability that an unemployed worker finds a job, UTJ_t, we follow Shimer (2012). Let u_t denote the number of unemployed workers at the end of month t and u_t^s the number of workers who at time t have been unemployed for less then a month, the so-called short-term unemployed; then UTJ_t can be backed out from the data using:

$$u_{t+1} = (1 - UTJ_t)u_t + u_{t+1}^s, \tag{38}$$

which implies

$$UTJ_t = 1 - \frac{u_{t+1} - u_{t+1}^s}{u_t}. \tag{39}$$

To calculate the employment exit probabilities, we assume that there is no possibility of both finding and losing a job within one month, which yields the following expression:

$$JTU_t = \frac{u_{t+1}^s}{l_t}, \tag{40}$$

where l_t is the stock employed persons. We are aware that time aggregation may cause a bias in our employment exit probabilities, however there is no reason to believe that the bias may differ across races.

B Hamilton-Jacobi-Bellman equations

The derivation is based on Stokey (2008).

Let $w(p)$ a return function and r is the interest rate. Ignore the exogenous job destruction shock, δ, and consider an infinite stream of returns, $w(p)$, where p is a diffusion with infinitesimal parameters $\mu(p)$ and $\sigma(p)$. Define $W(p_0)$ to be the expected discounted value of the stream of returns given the initial state $p(0) = p_0$:

$$W(p_0) \equiv E\left[\int_0^\infty e^{-rt} w(p(t,\omega))dt \,|\, p(0) = p_0\right], \forall p_0. \tag{41}$$

For any small interval of time Δt, (41) has the Bellman-type property:

$$W(p_0) \approx w(p_0)\Delta t + \frac{1}{1 + r\Delta t} E\left[W(p(0 + \Delta t)) \,|\, p(0) = p_0\right]. \tag{42}$$

Multiply this equation by $(1 + r\Delta t)$, subtract $W(p_0)$ and divide by Δt to get

$$rW(p_0) \approx w(p_0)(1 + r\Delta t) + \frac{1}{\Delta t} E\left[W(p(0 + \Delta t)) - W(p_0) \,|\, p(0) = p_0\right]. \tag{43}$$

Let $\Delta t \to 0$ to find that

$$rW(p_0) \approx w(p_0) + \frac{1}{dt} E\left[dW \,|\, p(0) = p_0\right]. \tag{44}$$

Next, we approximate the total differential of $W(t,p)$, called dW, with a Taylor series expansion. Let $W_t \equiv \partial W/\partial t$, $W_p \equiv \partial W/\partial p$, and so on, to denote the partial derivatives of W. We have

$$dW = W_t dt + W_p dp + \frac{1}{2} W_{pp}(dp)^2 + \ldots \tag{45}$$

Equation (2) established that p is a diffusion with initial value $p(0) = p_0$, zero drift $\mu_p(p) = 0$ and a diffusion parameter $\sigma_p(p) = p(1-p)\frac{\mu_H - \mu_L}{\sigma_X}$. Plug it into (45) to find

$$dW = W_t dt + W_p \sigma_p(p) d\overline{Z} + \frac{1}{2} W_{pp} \sigma_p(p)^2 (d\overline{Z})^2.$$

Since $\overline{Z}$ is a standard Wiener process, $E[d\overline{Z}] = 0$ and $E[(d\overline{Z})^2] = dt$. Hence,

$$E[dW] = [W_t + \frac{1}{2}\sigma_p(p)^2 W_{pp}]dt. \tag{46}$$

The function W does not directly depend on time, so using (46) in (44) and dropping the subscript on the initial condition produces the HJB equation:

$$rW(p) = w(p) + \frac{1}{2}\sigma_p(t)^2 W''(p). \tag{47}$$

C The solution of the HJB equation

The solution is a slight amendment to Moscarini (2005). It can be verified directly. Define $n = \sqrt{1/4 + 2(r+\delta)/s_i^2}$. Thus,

$$\begin{aligned} S_i'(p) &= c_i(1/2 - n)p^{-1/2-n}(1-p)^{1/2+n} \\ &\quad - c_i(1/2+n)p^{1/2-n}(1-p)^{-1/2+n} \\ &\quad + \frac{\overline{\mu}'(p)}{r+\delta} \\ &= c_i p^{-1/2-n}(1-p)^{-1/2+n}(1/2 - n - p) + \frac{\mu_H - \mu_L}{r+\delta}. \end{aligned} \tag{48}$$

and

$$\begin{aligned} S_i''(p) &= c_i(1/2-n)(-1/2-n)p^{-3/2-n}(1-p)^{1/2+n} \\ &\quad - c_i(1/2-n)(1/2+n)p^{-1/2-n}(1-p)^{-1/2+n} \\ &\quad - c_i(1/2+n)(1/2-n)p^{-1/2-n}(1-p)^{-1/2+n} \\ &\quad + c_i(1/2+n)(-1/2+n)p^{1/2-n}(1-p)^{-3/2+n} \\ &= -c_i(1/4-n^2)p^{-3/2-n}(1-p)^{-3/2+n}. \end{aligned} \tag{49}$$

Finally, plugging (49) into (9) verifies the solution (10).

Also, note that c_i and $\underline{p}_i$ solve the following system of equations:

$$\begin{aligned} S_i'(\underline{p}_i) &= c_i\underline{p}_i^{-1/2-n}(1-\underline{p}_i)^{-1/2+n}(1/2-n-\underline{p}_i)+\frac{\mu_H-\mu_L}{r+\delta}=0, \\ S_i(\underline{p}_i) &= c_i\underline{p}_i^{\frac{1}{2}-\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}}(1-\underline{p}_i)^{\frac{1}{2}+\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}}+\frac{\overline{\mu}(\underline{p}_i)-rU_i}{r+\delta}=0. \end{aligned}$$

D Endogenous contact rate

In the model the value of a vacancy is arbitrarily set to 0, whereas the contact rate λ is exogenous. As in Papageorgiou (2014), this assumption is innocuous. Suppose there is a constant-return to scale Cobb-Douglas matching function that sets the number of meets. Then the contact rate is $\lambda(\theta)=A\theta^{\alpha}$, where θ is the vacancy-to-unemployed ratio. Moreover, suppose that holding a vacancy involves paying the flow cost κ. The value of a vacancy V solves

$$rV=-\kappa+\frac{\lambda(\theta)}{\theta}\sum_i m_i\int_{\underline{p}_i}^1[J_i(p)-V]g_i^0(p)dp. \tag{50}$$

Assuming free entry of new firms leads to $V=0$ and so

$$\kappa=\frac{\lambda(\theta)}{\theta}\sum_i m_i\int_{\underline{p}_i}^1 J_i(p)g_i^0(p)dp. \tag{51}$$

Thus for a given set of parameter estimates and a given θ, we compute $\frac{\lambda}{\theta}\sum_i m_i\int_{\underline{p}_i}^1 J_i(p)g_i^0(p)dp$, we set $A=\lambda\theta^{-\alpha}$ and we finally choose κ as in equation (51).

E Proofs

Proof of Proposition 1. **Outside Option Effects**

Proof by contradiction. Assume $U_W\le U_B$. Using (5) and (8), the value of unemployment can be rewritten as

$$rU_W=b+\beta\lambda\int_0^1\max\{S_W(p),0\}g_W^0(p)dp. \tag{52}$$

The surplus equation (9) implies that $S_W''(p)=S_B''(p)$ and thus $S_W(p)\ge S_B(p)$, for all $p\in[0,1]$. Using this fact

$$rU_W\ge b+\beta\lambda\int_0^1\max\{S_B(p),0\}g_W^0(p)dp. \tag{53}$$

Since $S_i(p)$ is strictly increasing in $p \in [\underline{p}_i, 1]$, Assumption 1 also implies

$$b + \beta\lambda \int_0^1 \max\{S_B(p), 0\} g_W^0(p) dp > b + \beta\lambda \int_0^1 \max\{S_B(p), 0\} g_B^0(p) dp = rU_B. \quad (54)$$

And so we have proved $U_W > U_B$. This contradicts the assumption that $U_W \leq U_B$. Thus, $U_W > U_B$.

Proof of Proposition 2. **Employment discrimination I.**

Part (i). Forthcoming.

Part(ii). The group-i specific job-finding rate is $jfr_i = \lambda[1 - G_i^0(\underline{p}_i)]$, $i = B, W$. The result follows from part (i) and Assumption 1.

Part (iii). At any time, the flow number of group-i workers who lose their job is $\delta(1 - u_i) + .5\sigma^2(\underline{p}_i) g_i'(\underline{p}_i)$. Thus the group-$i$ specific job-loss rate is $jlr_i = \delta + .5\sigma^2(\underline{p}_i) g_i'(\underline{p}_i)/(1 - u_i)$, $i = B, W$. The result follows.

Proof of Proposition 3. **Wage discrimination I.**

Part (i). We have $\omega_i(p) = \beta\bar{\mu}(p) + (1 - \beta) rU_i$ for $i = B, W$ and $p \in [\underline{p}_i, 1]$. This proves the result.

Part (ii). We have

$$\begin{aligned}
\Pr[\omega_i &\leq \omega \mid i] = \Pr[\beta\bar{\mu}(p) + (1 - \beta) rU_i \leq \omega \mid i] \\
&= \Pr[p \leq \beta^{-1}(\omega - (1 - \beta) rU_i - \beta\mu_L)/(\mu_H - \mu_L) \mid i] \\
&= \tilde{G}_i[\beta^{-1}(\omega - (1 - \beta) rU_i - \beta\mu_L)/(\mu_H - \mu_L)] \\
&= F_i(\omega). \quad (55)
\end{aligned}$$

Thus

$$\begin{aligned}
\omega_{iq} &= F_i^{-1}(q) \\
&= \beta(\mu_H - \mu_L)\tilde{G}_i^{-1}(q) + (1 - \beta) rU_i + \beta\mu_L. \quad (56)
\end{aligned}$$

The result follows.

Proof of Proposition 4. **Outside option differential II.**

Forthcoming.

Proof of Proposition 5. **Employment discrimination II**

Part (i). Forthcoming.

Part (ii). The group-i specific job-finding rate is $jfr_i = \lambda[1 - G^0(\underline{p}_i)]$, $i = B, W$. The result follows from part (i).

Part (iii). See part (iii) of Proposition 2.

Proof of Proposition 6. **Wage discrimination II.**

Part (i). The result follows from Proposition 4.

Part (ii). We have $z(0) = \beta(\mu_H - \mu_L)\left[\underline{p}_W - \underline{p}_B\right] + (1-\beta)r\left(U_W - U_B\right) < (1 - \beta)r\left(U_W - U_B\right) = z(1)$.

References

Arrow, Kenneth, "The theory of discrimination," *Discrimination in labor markets*, 1973, *3* (10), 3–33.

Becker, Gary S, "The Economics of Discrimination," *University of Chicago Press Economics Books*, 1971.

Black, Dan A, "Discrimination in an equilibrium search model," *Journal of labor Economics*, 1995, pp. 309–334.

Borowczyk-Martins, Daniel, Jake Bradley, and Linas Tarasonis, "Racial Discrimination in the U.S. Labor Market: Employment and Wage Differentials by Skill," IZA Discussion Papers 8176, Institute for the Study of Labor (IZA) May 2014.

Bowlus, Audra J and Zvi Eckstein, "Discrimination and Skill Differences in an Equilibrium Search Model*," *International Economic Review*, 2002, *43* (4), 1309–1345.

Eckstein, Zvi and Kenneth I Wolpin, "Estimating the effect of racial discrimination on first job wage offers," *Review of Economics and Statistics*, 1999, *81* (3), 384–392.

Flabbi, Luca, "Gender Discrimination Estimation in a Search Model with Matching and Bargaining," *International Economic Review*, 2010, *51* (3), 745–783.

_ , "Prejudice and gender differentials in the US labor market in the last twenty years," *Journal of Econometrics*, 2010, *156* (1), 190–200.

Gayle, George-Levi and Limor Golan, "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997," *The Review of Economic Studies*, 2011, p. rdr019.

Gourieroux, Christian, Alain Monfort, and Eric Renault, "Indirect inference," *Journal of applied econometrics*, 1993, *8* (S1), S85–S118.

Holden, Steinar and Åsa Rosén, "Discrimination and employment protection," *Journal of the European Economic Association*, 2014, *forthcoming.*

Jovanovic, Boyan, "Matching, turnover, and unemployment," *The Journal of Political Economy*, 1984, pp. 108–122.

Jr, Roland G Fryer, Devah Pager, and Jörg L Spenkuch, "Racial Disparities in Job Finding and Offered Wages," *Journal of Law and Economics*, 2013, *56* (3), 633–689.

Lang, Kevin and Jee-Yeon K Lehmann, "Racial Discrimination in the Labor Market: Theory and Empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006.

Mailath, George J, Larry Samuelson, and Avner Shaked, "Endogenous inequality in integrated labor markets with two-sided search," *American Economic Review*, 2000, pp. 46–72.

Moro, Andrea, "The Effect Of Statistical Discrimination On Black–White Wage Inequality: Estimating A Model With Multiple Equilibria," *International Economic Review*, 2003, *44* (2), 467–500.

Mortensen, Dale T and Christopher A Pissarides, "Job creation and job destruction in the theory of unemployment," *The review of economic studies*, 1994, *61* (3), 397–415.

Moscarini, Giuseppe, "Job matching and the wage distribution," *Econometrica*, 2005, *73* (2), 481–516.

Papageorgiou, Theodore, "Learning Your Comparative Advantages," *The Review of Economic Studies*, 2014, *81* (3), 1263–1295.

Phelps, Edmund S, "The statistical theory of racism and sexism," *The american economic review*, 1972, pp. 659–661.

Rosén, Åsa, "An equilibrium search-matching model of discrimination," *European Economic Review*, 1997, *41* (8), 1589–1613.

Shimer, Robert, "Reassessing the ins and outs of unemployment," *Review of Economic Dynamics*, 2012, *15* (2), 127–148.

Stokey, Nancy L, *The Economics of Inaction: Stochastic Control models with fixed costs*, Princeton University Press, 2008.

Wonham, W Murray, "Some applications of stochastic differential equations to optimal nonlinear filtering," *Journal of the Society for Industrial & Applied Mathematics, Series A: Control*, 1964, *2* (3), 347–369.