LIETUVOS BANKAS
EUROSISTEMA

Statistical Discrimination in a Search Equilibrium Model: Racial Wage and Employment Disparities in the US

Working Papers Series

No 82 / 2020

ISSN 2029-0446 (ONLINE)

Working Papers Series
No 82 / 2020

Statistical Discrimination in a Search Equilibrium Model: Racial Wage and Employment Disparities in the US[*]

Bruno Decreuse[†]

Linas Tarasonis[‡]

[*] We thank Katarína Borovičková, Jake Bradley, Denis Fougère, Grégory Jolivet, Andrea Moro, Grigorios Spanos, 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, the EEA annual congress in Mannheim, various meetings of the Search and Matching (SaM) network, the annual Lithuanian Conference on Economic Research, the annual conference of the Spanish Economic Association, the China and Asia meetings of the Econometric Society. Bruno Decreuse thanks the Institut Universitaire de France for financial support. This work was supported by French National Research Agency Grants ANR-17-EURE-0020.

[†] Aix-Marseille University, CNRS, AMSE, Marseille, France, bruno.decreuse@univ-amu.fr, http://sites.google.com/site/brunodecreuseecon/

[‡] Vilnius University & Bank of Lithuania, Vilnius, Lithuania, linas.tarasonis@evaf.vu.lt, http://linas.econ.lt/

Gedimino pr. 6, LT-01103 Vilnius

www.lb.lt

Working Papers describe research in progress by the author(s) and are published to stimulate discussion and critical comments.

The series is managed by Applied Macroeconomic Research Division of Economics Department and the Center for Excellence in Finance and Economic Research.

All Working Papers of the Series are refereed by internal and external experts.

ABSTRACT

In the US, black workers spend more time in unemployment, lose their jobs more rapidly, and earn lower wages than white workers. This paper quantifies the contributions of statistical discrimination, as portrayed by negative stereotyping and screening discrimination, to such employment and wage disparities. We develop an equilibrium search model of statistical discrimination with learning based on Moscarini (2005) and estimate it by indirect inference. We show that statistical discrimination alone cannot simultaneously explain the observed differences in residual wages and monthly job loss probabilities between black and white workers. However, a model with negative stereotyping, larger unemployment valuation and faster learning about the quality of matches for black workers can account for these facts. One implication of our findings is that black workers have larger returns to tenure.

Keywords: Learning; Screening discrimination; Job search; Indirect inference.

JEL codes: J31; J64; J71.

1 Introduction

Compared to white male workers, black male workers earn lower wages, stay unemployed longer and lose their jobs more rapidly. Though a substantial part of such disparities can be explained by differences in observable characteristics such as age, education or location, residual wages and employment differentials are large and persist over time (see, e.g., Lang and Lehmann (2012)). These findings raise questions on the roles played by unobserved skills, i.e., skills that are undocumented in survey data but have a crucial impact on workers' performance, and on their learning by employers and workers. How are such skills distributed among blacks and whites? How much time do employers need to figure out the true productivity of a worker-job pair? Is this learning process faster for whites than for blacks?

The branch of economics addressing these questions is referred to as statistical discrimination. This type of discrimination arises when employers imperfectly observe the productivity of workers, while the distribution of productive outcomes varies across race. Since Phelps (1972), statistical discrimination takes two forms: negative stereotyping and screening discrimination. Negative stereotyping (hereafter, NS) happens when employers believe that jobs occupied by black workers are on average less productive. All blacks are attributed the mean productivity of their group, which generates wage redistribution among black employees, from the most productive workers to the least ones. Screening discrimination (hereafter, SD) occurs when employers need more time to learn the productivity of jobs occupied by black workers. These workers are seen as less employable and experience slower wage growth.

Statistical discrimination has never been evaluated within the context of a formal model predicting racial wage gaps as well as differences in the probability that an unemployed worker finds a job and the probability that an employed worker becomes unemployed. This paper aims to fill this gap. We provide a dynamic model of statistical discrimination with search frictions and employer-employee learning and then estimate its structural parameters with indirect inference. Our results shed light on a fundamental trade-off between fitting wage disparities and fitting employment ones. This leads us to a different perspective on SD: the learning process is likely faster with blacks than with whites, a phenomenon we refer to as anti-screening discrimination (hereafter, anti-SD).

We first begin our analysis by presenting several empirical regularities in Section 2. Specifically, we use the Current Population Survey and focus on prime-aged low-skilled male workers to describe the black-white wage and employment disparities. We compute the job-finding and job-separation rates of blacks and whites following Shimer (2012). The job-finding and job loss rate differentials are above 30%, i.e., blacks spend 30% more time in unemployment and are 50% more likely to lose their job in the following month. We then compute residual wages using a Mincer wage regression. We find the quantile differentials of the residual wage distributions are large – the wage gap amounts to 14 percent for both entry wages and unconditional wages – and increasing, i.e., residual wage disparities are larger in levels at the top than at the bottom of the wage distribution.

We then proceed to our theoretical model, presented in Section 3, which draws from Moscarini (2005) who introduces job turnover in the spirit of Jovanovic (1984) in an equilibrium search unemployment framework. Each match between ex-ante identical workers and firms is characterized by an unobserved match quality that can be high or low. All worker-firm pairs start with a probability of being in a high-quality match and the true match quality is gradually learnt over time by observing output realizations. Job loss occurs when a worker-firm pair learns their match is sufficiently likely to be of low quality. Wage bargaining over the match surplus implies there is a mapping from the ergodic distribution of posterior beliefs about match quality to the stationary wage distribution.

In our model we introduce two groups of workers, blacks and whites, and group-specific distributions of observed and unobserved skills. To account for hiring discrimination, prior beliefs about the quality of matches are drawn from a distribution with a continuous support. This distribution is allowed to be different between blacks and whites to reflect differences in unobserved heterogeneity between the two groups. As a consequence, the model predicts job-finding, job loss and the wage distribution for both groups of workers.

In Section 4, we use the simulated method of moments to estimate our model. We target moments characterizing labor market outcomes for the two groups, mean monthly job-finding and job loss probabilities, quantiles of the unconditional wage distribution, and quantiles of the entry wage distribution, and obtain two main results.

On the one hand, statistical discrimination, as portrayed by NS and SD, fails to match simultaneously the properties of the quantile differentials and those of the job-finding and job loss probability differentials. In particular, NS predicts globally decreasing quantile differentials and small job loss differential, whereas SD predicts increasing quantile differentials and higher job loss for whites. The intuition for this result is straightforward and goes beyond the particularities of our model: when learning about match quality is faster for whites, these workers benefit from higher wage growth (high-quality matches are rapidly revealed), but also experience shorter job durations (low-quality matches are also rapidly revealed).

On the other hand, we show that together with differences in the valuation of unemployment between blacks and whites, statistical discrimination can explain all of the observed empirical regularities. The resulting estimation involves NS, anti-SD and higher utility when unemployed for blacks. Anti-SD means that output signals occurring during employment are more accurate when the worker is black. Following the previous reasoning, blacks lose their jobs faster. Then, NS guarantees that the black-white differential remains large, whereas higher unemployment valuation for blacks ensures the quantile differentials of the wage distributions are increasing.

Given the focus of the literature, anti-SD seems counter-intuitive. The screening discrimination literature mainly emphasizes the opposite.[1] One recent paper, however, offers a micro foundation of anti-SD. Cavounidis and Lang (2015) study managers' incentive to monitor the different groups of workers. They show that when blacks are more often in unproductive jobs, employers have stronger incentive to monitor them.

Furthermore, anti-SD offers a new perspective on racial returns to tenure. Fryer Jr et al. (2013) estimate blacks have a return-to-tenure rate that is 1.1 percentage points higher than for whites. They explain this result with a stylized three-period model of statistical discrimination where productivity is revealed after one period. NS implies that black workers will, on average, receive lower wages than whites, which leads to larger scope for wage improvement. Our estimate with anti-SD also concludes the return to tenure is larger for blacks than for whites (by 1.7 percentage points), but for a completely different reason: anti-SD implies learning is faster for blacks.

There already exist estimates of models of statistical discrimination for the labor market, but they do not feature search unemployment. Moro (2003) develops and estimates a model of racial discrimination with complementarities between skilled and unskilled workers, whereas Gayle and Golan (2012) focus on gender gaps.

[1] For example, Ritter and Taylor (2011) have an efficiency wage model rationalizing racial unemployment disparities in which performance observability during employment is better for whites than for blacks. They relate this assumption to the theory of language discrimination (Lang, 1986).

Both papers abstract from search frictions and do not account for racial differences in unemployment duration and job separation.

By contrast, there is a substantial literature offering estimates of taste-based discrimination in search unemployment models. These models follow Becker (1971) and aim to disentangle the respective roles played by racial prejudice and unobserved worker heterogeneity in the labor market outcomes of different demographic groups: Black (1995), Eckstein and Wolpin (1999), Bowlus and Eckstein (2002), Borowczyk-Martins et al. (2017), for racial discrimination, but also Flabbi (2010a,b) for gender discrimination. Our approach does not aim to disentangle prejudice from unobserved worker heterogeneity, but we are not necessarily inconsistent with this literature. Indeed, NS may reflect, in addition to differences in productivity of the members of a particular group, existence of employers with discriminatory tastes.

More generally, search and matching models provide an interesting framework to study discrimination. In the spirit of Arrow (1973), several papers show discrimination can arise in equilibrium despite employers having no taste for discrimination and blacks and whites having similar characteristics.[2] We do not explore the rich possibilities offered by such models. Instead, we draw from the framework of Phelps (1972) where skills are exogenously different between blacks and whites, and output observability varies across ethnic groups.

2 Evidence

The purpose of this section is twofold. First, we summarize key differences in labor market outcomes of blacks and whites. Our analysis focuses on job-finding and job loss probabilities, as well as residual wages. Second, we provide a number of empirical moments that allow us to estimate the parameters of a theoretical model presented in Section 3 and estimated in Section 4.

Data.—We use Basic Monthly Data of the Current Population Survey (CPS) from January 2003 to December 2008 and limit the sample to individuals who declare

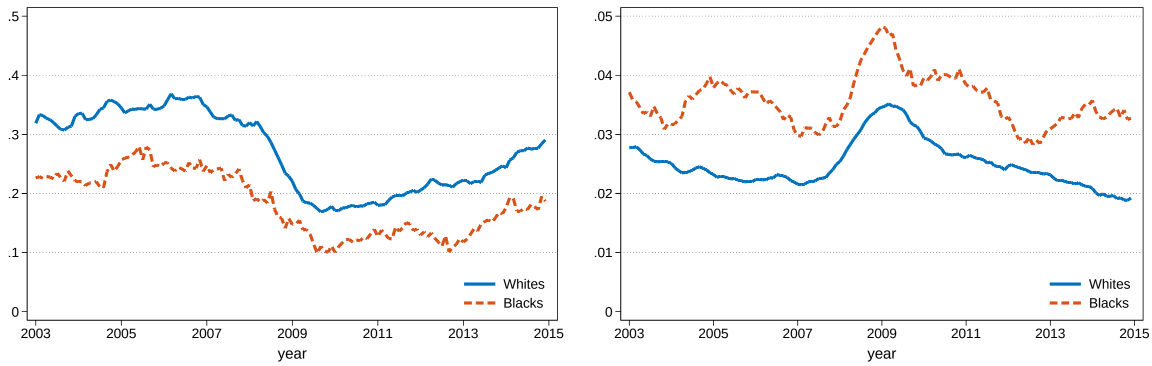
[2]In Rosén (1997), employers have private information on match-specific productivity. Discriminated Blacks apply for low-quality matches, thereby creating the type of belief that leads employers to discriminate them. In Mailath et al. (2000), employers can direct their search towards Blacks or Whites, whereas workers make a pre-market investment in skills. If employers do not send offers to Blacks, then these workers invest less in human capital, justifying employers' behavior. In Holden and Rosén (2014), match quality is random and workers in bad matches 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.

themselves to be either black or white males. For homogeneity, we only consider individuals without college education between the ages of 25 and 55 and focus on full-time workers in non-agriculture private sectors, and exclude self-employed workers. We do so because our model implies that wages depend on individual characteristics, whereas employment to unemployment and unemployment to employment transition rates do not. This implication forces us to focus on a group of relatively homogenous workers.[3]

Transition probabilities.—To measure the average monthly job loss probability and the average monthly job-finding probability, we follow Shimer (2012) and suppose that all workers of a given group have the same job-finding and job separation rates and ignore movements in and out of the labor force. The method uses monthly measures of the number of employed and unemployed workers as well as the number of unemployed workers with zero to four weeks duration of each group. The details are explained in Appendix A.

Figure 1: The ins and outs of unemployment

(a) Job-finding probability (b) Job loss probability



Notes.— Prime-age men with no college, 2003m1-2014m12, 12 month moving average of monthly data. Source: Current Population Survey and authors' calculations.

Figure 1 shows the monthly job-finding and employment exit probabilities over an enlarged period that also includes the Great Recession. Several facts about racial differences in labor market transition probabilities stand out. First, a typical black unemployed worker is on average 30% less likely to find a job in a given month over the observed period. The racial gap in the job-finding probability is relatively stable over the business cycle. Second, black workers are 50% more

[3]We choose a group of low educated workers because of the public debate about Black labor market outcomes and Black education. We acknowledge there is heterogeneity within this group. However, we do not wish to restrict the analysis further (for instance to those with high school diploma only) because of data availability and lack of external relevance.

likely to become unemployed in a given month than white workers. The racial gap in separation rates appears to be less stable over time, however, it is mainly due to the less precise estimates of the separation rate of black workers. Finally, as can be observed in Figure 1, the transition rates are relatively stable during the pre-crisis period that we use for our estimation.

Residual wages.—To construct residual wages, we use the Merged Outgoing Rotation Groups when information on usual weekly hours/earnings is recorded. Specifically, this information is measured at the household's fourth and eighth month in the survey. To obtain hourly wages we use reported usual hourly wages when a worker is paid hourly, or usual weekly earnings divided by usual weekly hours worked otherwise. Since wages are top-coded, we only consider observations with hourly wages above $1 and below $100 when estimating returns to observable characteristics and we trim the top and bottom 2% of the residual hourly wage distribution for both groups. The resulting sample contains about 118,000 individual-year observations, of which nearly 9% correspond to blacks. We define newly hired workers as those employed during the fourth or the eighth month in the survey and nonemployed at any point previously.

To account for characteristics that are not modeled by our theory, we omit black workers and workers with tenure because their effects will be precisely modeled in the next section. To obtain the returns to explanatory variables other than race and tenure, we estimate a reduced-form wage regression on a subset of newly hired white workers:

$$\ln w_j = A_j \Gamma + \epsilon_j, \tag{1}$$

where w_j is the real hourly wage of a newly hired white individual j, ε_j is the error term and A_j is a vector of individual characteristics including a constant term, years of schooling, age, age squared, marital status, state, occupation and industry dummies. Considering age and age squared allows us to account for the effects of experience and general human capital accumulation that our model neglects. The underlying assumption is that workers accumulate general human capital whether they are employed or not.

We use the estimated returns to characteristics, $\hat{\Gamma}$, to obtain residual wages of all workers in our sample including blacks and those with positive tenure within the firm. For individual i, we define the normalized residual wage, or, more simply, the residual wage, as follows

$$\omega_i = \frac{w_i \exp(-A_i \hat{\Gamma})}{\max_j \{w_j \exp(-A_j \hat{\Gamma})\}}, \quad (2)$$

where w_i is the observed hourly wage, $\hat{\Gamma}$ is the vector of OLS estimates of equation (1). The normalization implies that the maximum residual wage is equal to one. This procedure leaves us with two residual wage distributions, one for each group.

Table 1: Summary statistics of residual wages

	All jobs				New jobs			
	Whites	Blacks	Diff.	Log-Diff.	Whites	Blacks	Diff.	Log-Diff.
Mean	.444	.394	.051	.121	.411	.365	.045	.117
St.-dev.	.161	.148	.013	.086	.155	.144	.011	.072
Min	.181	.161	.019	.113	.181	.169	.012	.070
5th perc.	.232	.204	.027	.126	.218	.199	.019	.091
25th perc.	.323	.286	.037	.122	.296	.261	.035	.126
50th perc.	.415	.361	.054	.139	.377	.328	.049	.139
75th perc.	.536	.473	.064	.126	.492	.439	.053	.115
95th perc.	.761	.701	.060	.082	.727	.677	.051	.072
Max	1.000	.911	.089	.093	.999	.890	.109	.115
N	107,223	10,698			5,146	633		

Notes.—Residual wages are defined in (2). The first four columns correspond to the unconditional residual wage distribution, the last four ones to the entry wage distribution. The third and fourth columns for both distributions, *Diff.* and *Log-Diff.*, report the black-white difference and log-difference, respectively.

Table 1 provides the main moments of residual wage distributions of both groups. Again, several facts stand out. First, the median black-white wage gap is around 14 percent for both the unconditional and the entry wage distributions. Second, as can be seen in column Diff., quantile differentials of both distributions are strongly increasing. Having increasing quantile differentials means that wage disparities are larger in levels at the right of the distribution than at its left. Third, the unconditional distribution of wages stochastically dominates the entry wage distribution and the difference is substantial. At the median, wages in all jobs are around 10 percent lower than in new jobs for both groups of workers.

The facts we describe here are in line with the literature. Elsby et al. (2010) find quantitatively similar aggregate racial differences in unemployment inflow and outflow rates using the CPS data, whereas DellaVigna and Paserman (2005) document the job-finding rate from unemployment is about 20% lower for blacks than for whites using the National Longitudinal Survey of Youth. Wage gaps are slightly higher than usual: Lang and Lehmann (2012) summarize the evidence by stating the unexplained wage gap is in the order of 10 percentage points. However, by design our measure of residual wage dispersion does not correct for differential tenures and returns to tenure between the two groups.

To summarize, blacks find jobs less rapidly, their jobs last shorter and differential quantiles are increasing. The rest of the paper is devoted to explaining these facts. Section 3 presents a dynamic model of statistical discrimination with employer-employee learning and search frictions, whereas Section 4 describes its estimation.

3 Theory

Our model draws from Moscarini (2005) who introduces job turnover in the spirit of Jovanovic (1984) in a Mortensen and Pissarides (1994) equilibrium search unemployment framework. We add two groups of workers with different productive abilities, different prior beliefs on the quality of matches, and different output observability. We first present the model and then focus on mechanisms and outcomes of discrimination.

All value functions, solutions and proofs lie in the Appendices B and E.

3.1 Model

Assumptions.—The labor market is populated by a continuum of risk-neutral workers of measure one and a large mass of firms, ensuring free entry. The labor market is characterized by random search frictions. Firms are ex-ante identical, whereas workers differ in observable type α and demographic group $i = B, W$, where B stands for black and W for white. The measure of each group is m_i, such that $m_B + m_W = 1$, and the distribution of type is group-specific: the cumulative distribution function (cdf) is Ψ_i and the associated probability density function (pdf) is $\psi_i \equiv \Psi_i'$.

The endogenous measure of unemployed workers in group i is u_i. When unemployed workers of type α obtain utility flow $b_i\alpha$ and receive job offers at rate λ, irrespective of their type.[4] Then the firm and the worker decide if they form

[4]The assumption that unemployment valuation is proportional to type is made for two reasons. First, this is for simplicity. This assumption leads to Lemma 1 in Appendix B whereby (i) the wage is proportional to type, which justifies the auxiliary regression on entry wages to purge data from observable characteristics, and (ii) the belief threshold $p_{\alpha i}$ does not depend on α, which allows us to measure transition rates in a simple and aggregate way. Second, unemployment valuation is influenced by unemployment income, which is tied to former wage. Having unemployment

a match. Employed workers lose their job at exogenous rate δ and also when the match surplus falls below zero. Though Moscarini (2005) extends his model to on-the-job search, we do not allow for it. The idea of our paper is to isolate learning as the only factor of wage growth and see how far statistical discrimination can go to explain employment and wage disparities between blacks and whites. Therefore we do not consider alternative mechanisms on the premise that they are similar for blacks and whites.[5]

The output of a firm-worker pair depends on workers' type and match quality μ according to $y_{\alpha\mu} = \alpha\mu$. Match quality can take two values: the match is good when $\mu = \mu_H$ and bad when $\mu = \mu_L < \mu_H$. Match quality is imperfectly observed at hiring and gradually learnt with tenure. When a firm and a worker meet, they draw a common signal $p_0 \in [0, 1]$ about the average productivity of the match. The signal is such that $p_0 = \Pr(\mu = \mu_H) = 1 - \Pr(\mu = \mu_L)$. In Moscarini (2005), this signal takes a single value. To account for hiring discrimination and for a non-degenerate wage distribution of those newly hired, we assume the signal is drawn from the group-specific cdf G_i^0 with associated pdf $g_i^0 \equiv G_i^{0\prime}$.

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 $\alpha^2\sigma_{Xi}^2$:

$$X_{\alpha it} = \alpha\left(\mu t + \sigma_{Xi} Z_t\right) \sim \mathcal{N}(\alpha\mu t, \alpha^2\sigma_{Xi}^2 t), \tag{3}$$

where Z_t is a Wiener process that keeps μ hidden. Given log-linearity in α, the variance-to-output ratio is type-independent. Hereafter, $s_i \equiv (\mu_H - \mu_L)/\sigma_{Xi}$ is the signal-to-noise ratio.

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. Wonham (1964) shows that $p_{\alpha it}$ follows a diffusion process:

$$dp_{\alpha it} = \sigma_{pi}(p_{\alpha it}) d\overline{Z}_{\alpha it}, \tag{4}$$

where

$$\sigma_{pi}(p) = p(1-p)\frac{\mu_H - \mu_L}{\sigma_{Xi}} \tag{5}$$

valuation proportional to type is a frequent assumption made to account for this wage dependence of unemployment valuation without having to deal with tricky computational implications when this dependence is explicitly modeled.

[5] The computation of residual wages provides a good illustration of this idea. Entry wages are regressed on age and age squared to capture the effects of labor market experience. Such effects, by assumption, are the same for blacks and whites.

is the diffusion parameter and

$$d\overline{Z}_{\alpha it} = \frac{dX_{\alpha it} - p_{\alpha it}\alpha\mu_H dt - (1 - p_{\alpha it})\alpha\mu_L dt}{\alpha\sigma_{Xi}} \tag{6}$$

is the innovation process, i.e., the normalized difference between realized and unconditionally expected flow output. The variable $\overline{Z}_{\alpha it}$ follows a standard Wiener process. Note that $dX_{\alpha it}$ is log-linear in worker type α and so equations (4)-(6) imply that beliefs depend on worker group and job tenure, but not on worker type.

Match formation and dissolution obey the same optimal stopping strategy. The match is stopped when the posterior belief of a good match falls to a belief, $\underline{p}_{\alpha i}$, where the firm-worker pair separate endogenously and restart searching on their own. To ensure the threshold belief $\underline{p}_{\alpha i}$ is nontrivial, we assume the following parametric restrictions hold:

$$\mu_H > b > \mu_L - \beta\lambda\frac{\mu_H - \mu_L}{r + \delta}\int_0^1 p dG_i^0(p). \tag{7}$$

The first inequality states that flow output in a good match must be larger than the utility flow derived from unemployment. If not, the match surplus is negative. The second inequality states that the utility flow derived from unemployment must be sufficiently large. Otherwise, all meetings give birth to employment relationships and hiring discrimination does not take place.

Let $w_{\alpha i}(p)$ be the wage and $U_{\alpha i}$ the expected utility of an unemployed worker. Nash bargaining implies

$$w_{\alpha i}(p) = \beta\alpha\overline{\mu}(p) + (1 - \beta)rU_{i\alpha}, \tag{8}$$

where $\beta \in [0, 1]$ is the worker's bargaining power and $\overline{\mu}(p) = p\mu_H + (1 - p)\mu_L$ is the expected match quality.

In the Appendix B, Lemma 1 shows that the value of unemployment and the wage are log-linear in α, i.e., $U_{\alpha i} = \alpha U_i$ and $w_{\alpha i} = \alpha w_i$, whereas the optimal stopping belief does not depend on α, i.e., $p_{\alpha i} = p_i$. These properties justify our procedure to construct residual wages. Wage log-linearity implies we can isolate residual wages by mean of a Mincer regression on newly hired whites, whereas the independence of the threshold belief vis-à-vis α guarantees the job-finding and job loss rates do not vary within groups[6].

[6]The distributions of observed individual characteristics do not differ much by employment status in our sample of low-skilled workers. In particular, education is roughly the same (the mean schooling duration is 11.0 years for the unemployed and 11.2 years for the employed). Employed workers are older by one year, which is compatible with the theory. There is one exception, the marital status: employed workers are more frequently married (65.5% against 46.9%).

Ergodic belief distribution.—Let $g_i(p)$ be the unnormalized pdf of the ergodic belief distribution among workers of group i. For beliefs below the threshold $\underline{p}_i$, this density is $g_i(p) = 0$. For beliefs above the threshold, the Kolmogorov forward equation describes its motion. Imposing stationarity we obtain:

$$0 = \partial_t g_i(p) = \partial_{pp}\left(\frac{1}{2}\sigma_{pi}^2(p) g_i(p)\right) + \lambda u_i g_i^0(p) - \delta g_i(p), \tag{9}$$

where u_i is the measure of unemployed workers. The first term balances all flows due to learning. The second term is the flow of workers at p from unemployment. The last term captures the attrition due to exogenous separation.

The forward equation is subject to two boundary conditions. The first condition states that the mass of workers above $\underline{p}_i$ is equal to the mass of employees, i.e., $\int g_i(p)dp = m_i - u_i \in [0, m_i]$. Moscarini (2005) names the second condition *no time spending at* $\underline{p}_i$ (NTS): $\frac{1}{2}\sigma_{pi}^2(\underline{p}_i) g_i(\underline{p}_i) = 0$. As $\sigma_{pi}^2(\underline{p}_i) \neq 0$ for $\underline{p}_i > 0$, the NTS condition implies that $g_i(\underline{p}_i) = 0$. Therefore the density of the belief distribution must be zero at its lower bound.

Hereafter, we suppose there exist $A < \infty$ and $a > -1$ such that $\lim_{p \to 1} g_i^0(p)(1 - p)^{1/2 - (1/4 + 2\delta/s_i^2)^{1/2}} / [A(1-p)^a] < 1$, i.e., the density function g_i^0 is sufficiently small in the neighborhood of $p = 1$.

Proposition 1 (Ergodic belief distribution). *Let $\nu_i = (1/4 + 2\delta/s_i^2)^{1/2}$, $i = B, W$. For all $p \geq \underline{p}_i$,*

$$\begin{aligned} g_i(p) &= \lambda u_i [p(1-p)]^{-3/2-\nu_i} / (\nu_i s_i^2) \times \\ & \left\{ p^{2\nu_i} \int_p^1 g_i^0(x) x^{1/2-\nu_i} (1-x)^{1/2+\nu_i} dx \right. \\ & \left. + (1-p)^{2\nu_i} \left(k_i - \int_p^1 g_i^0(x) x^{1/2+\nu_i} (1-x)^{1/2-\nu_i} dx \right) \right\}, \end{aligned} \tag{10}$$

with

$$k_i = \int_{\underline{p}_i}^1 g_i^0(x) x^{1/2+\nu_i} (1-x)^{1/2-\nu_i} dx - \left(\frac{\underline{p}_i}{1 - \underline{p}_i}\right)^{2\nu_i} \int_{\underline{p}_i}^1 g_i^0(x) x^{1/2-\nu_i} (1-x)^{1/2+\nu_i} dx. \tag{11}$$

This (unnormalized) pdf generalizes Moscarini (2005) to the case of any non-degenerate prior distribution. As in Moscarini, the density is such that $g_i(\underline{p}_i) = 0$ and $g_i(1) = 0$ when $\delta > s_i^2$. Otherwise, $\lim_{p \to 1} g_i(p) = \infty$. Thus the exogenous component of job destruction must be sufficiently large to avoid cases with a large concentration of workers around the highest possible beliefs. Lastly, $g_i(\underline{p}_i) = 0$ and

$g(1) = 0$ does not imply singlepeakedness, which depends on properties of the prior distribution.

Given g_i, the corresponding normalized pdf is obtained by dividing by the mass of workers in employment. Namely, let $\tilde{g}_i$ be the normalized pdf of the ergodic belief distribution and $\tilde{G}_i$ be the corresponding cdf. We have $\tilde{g}_i(p) = g_i(p)/\int g_i(a)da$ and $\tilde{G}_i(p) = \int_{a \leq p} \tilde{g}_i(a)da$.

Stationary wage distribution.—As explained above, wages are linear functions of beliefs, i.e., $w_{\alpha i}(p) = \beta\alpha\overline{\mu}(p) + (1-\beta)rU_{\alpha i}$. We define $\omega_i = w_{\alpha i}/\alpha$ as the *residual wage* and $\Delta\omega(p) \equiv \omega_W(p) - \omega_B(p)$ as the *black-white residual wage differential* conditional on belief p. The residual wage differential measures wage discrimination because it focuses on two seemingly identical workers who hold a job characterized by the same belief on match quality.

We also define F_i as the group-i-specific residual wage distribution. By definition, $F_i(\omega) = \Pr[\omega_i \leq \omega \mid i]$. We compute the different quantiles of the distribution as follows: for $q \in [0,1]$, $\omega_{iq} = F_i^{-1}(q)$ is the q-th quantile of the group-i-specific wage distribution. Lastly, $z(q) \equiv \omega_{Wq} - \omega_{Bq}$ is the *black-white quantile differential*, or quantile differential for short.

The quantiles of the wage distribution are such that $\omega_{iq} = \beta(\mu_H - \mu_L)\tilde{G}_i^{-1}(q) + (1-\beta)rU_i + \beta\mu_L$. Therefore the quantile differential is

$$z(q) = \beta(\mu_H - \mu_L)\left[\tilde{G}_W^{-1}(q) - \tilde{G}_B^{-1}(q)\right] + (1-\beta)r\left(U_W - U_B\right). \tag{12}$$

The first term is the difference in belief quantile. This term depends on the threshold beliefs of the two groups, the group-specific variances of output noise, the rates at which worker-firm pairs learn match quality, and the job separation rates. It is positive when the ergodic belief distribution of whites stochastically dominates at first order the distribution of blacks. The second term depends on the differential return to search. It is positive when whites fare better than blacks in the labor market.

By construction, $z(1) = (1-\beta)r\left(U_W - U_B\right)$ and $z(0) = \beta(\mu_H - \mu_L)(\underline{p}_W - \underline{p}_B) + (1-\beta)r\left(U_W - U_B\right)$. The top quantile differential mirrors the outside option differential. The bottom quantile differential also reflects differential selection through the differential belief threshold.

3.2 Discrimination

We now turn to potential differences between blacks and whites conditional on type α. Hereafter we refer to residual wages as being simply wages. We consider the two aspects of statistical discrimination, i.e., negative stereotyping (NS) and screening discrimination (SD). We also study unemployment valuation (UV) heterogeneity because this factor is important in the next section. In this presentation, blacks are supposedly exposed to discrimination and higher unemployment valuation.

Negative stereotyping.—In this case, employers hold negative (rational) beliefs about the ability of black workers to form a good match.

Assumption 1 (Stochastic dominance): *The distribution $G_W^0(p)$ stochastically dominates the distribution $G_B^0(p)$ at first order, i.e., $G_B^0(p) \geq G_W^0(p)$ for all $p \in [0,1]$ and there is $\tilde{p} \in [\underline{p}_W, 1]$ such that $G_B^0(\tilde{p}) > G_W^0(\tilde{p})$. Moreover, $\sigma_{XB} = \sigma_{XW} = \sigma_X$ and $b_B = b_W = b$.*

Stochastic dominance is a simple way to describe prior heterogeneity between groups. Black workers tend to draw lower initial beliefs on match quality. The origin of such a differential is not discussed here.

Proposition 2 (Negative stereotyping). *Under Assumption 1, the following statements hold:*

A. Return to search: $rU_W > rU_B$.

B. Employment disparities:

- *(i) $1 > \underline{p}_W > \underline{p}_B > 0$;*
- *(ii) the job-finding rate differential is $\Delta jfr = \lambda \left[G_B^0(\underline{p}_B) - G_W^0(\underline{p}_W) \right]$ and may be positive or negative;*
- *(iii) the job-loss rate differential is $\Delta jlr = \frac{1}{2} \left[\sigma_p^2(\underline{p}_B)\tilde{g}_B'(\underline{p}_B) - \sigma_p^2(\underline{p}_W)\tilde{g}_W'(\underline{p}_W) \right]$ and may be positive or negative.*

C. Wage disparities:

- *(i) for all $p \in [\underline{p}_W, 1]$, the wage differential is $\Delta\omega(p) = (1-\beta)r(U_W - U_B) > 0$;*
- *(ii) the quantile differential is such that $z(0) > z(1) > 0$.*

Part A shows that whites enjoy a larger return to search. Matches with whites are more productive on average. Therefore wage and employment expectations are better for these workers.

Part B describes the ambiguous impacts of NS on employment outcomes. (i) shows that whites are more selected than blacks into employment. That $U_W > U_B$ implies the match surplus conditional on belief p is always larger for blacks than for whites. Therefore the lowest belief compatible with nonnegative surplus is larger for whites. (ii) shows that the job-finding rate differential has ambiguous sign. Whites have a better prior distribution, which improves their job-finding rate, but are also more selected, which reduces their chance of finding a job. (iii) shows the job loss rate differential also has an ambiguous sign. The flow of employees who cross the belief threshold $\underline{p}_i$ depends on the variance of the learning process σ_p^2 and on the slope of the pdf of the belief distribution evaluated at the belief threshold. Both components differ across groups.

Part C features the non-ambiguous impacts of NS on wage outcomes. (i) shows that blacks are discriminated against: the wage differential reveals the outside option differential benefiting to whites. (ii) shows that the associated quantile differentials are positive. Whites have a better belief distribution and a higher return to search. The former effect is especially strong at the bottom of the distribution but disappears at its top where there is no uncertainty on match quality. Therefore NS is better at explaining wage disparities at the bottom of the wage distribution than at the top.

To summarize, ex-ante differences in prior distributions have ambiguous impacts on employment outcomes and unambiguous effects on wage outcomes. In particular, they predict positive but decreasing quantile differentials. This property is at odds with the evidence reported in Section 2.

Screening discrimination.—We now suppose the precision of output signals during employment differs between blacks and whites. This hypothesis is associated to Aigner and Cain (1977) and Cornell and Welch (1996) in the context of static models. Closer to us, Ritter and Taylor (2011) provide a model of SD with employer learning.

Assumption 2 (OUTPUT OBSERVABILITY). *The standard deviation of output is larger for blacks than for whites, i.e.,* $\sigma_{XB} > \sigma_{XW}$. *Moreover,* $G_B^0 = G_W^0 = G^0$ *and* $b_B = b_W = b$.

Employers have more difficulties to infer match quality from output signals when the worker is black. Therefore learning is faster with whites, which exposes blacks to hiring discrimination and lower wage growth.

Proposition 3 (SCREENING DISCRIMINATION). *Under Assumption 2, the following statements hold:*

A. *Return to search:* $U_W > U_B$.

B. *Employment disparities:*

   (i) $1 > \underline{p}_B > \underline{p}_W > 0$;

(ii) the job-finding rate differential is $\Delta jfr = \lambda \left[G^0(\underline{p}_B) - G^0(\underline{p}_W) \right] > 0$*;*

(iii) the job-loss rate differential is $\Delta jlr = \frac{1}{2} \left[\sigma^2_{pB}(\underline{p}_B)\tilde{g}'_B(\underline{p}_B) - \sigma^2_{pW}(\underline{p}_W)\tilde{g}'_W(\underline{p}_W) \right]$ *and may be positive or negative.*

C. Wage disparities:

(i) for all $p \in [\underline{p}_W, 1]$*, the wage differential* $\Delta\omega(p) = (1 - \beta)r(U_W - U_B) > 0$*;*

(ii) the quantile differential is such that $z(1) > \max\{z(0), 0\}$*.*

Part A shows that, like NS, blacks have a lower return to search. Learning has less value when the worker is black and the match surplus is smaller at given belief on match quality.

Part B (i) shows that, unlike NS, blacks are more selected than whites into employment. Job tenure provides less information on match quality when the job is occupied by a black worker. Thus employers have less incentive to hire blacks. This result implies (ii): blacks are less likely to form matches and their job-finding rate is smaller. (iii) shows a more intriguing result: whites may lose their jobs faster than blacks. Output signals convey more accurate information when the worker is white. Bad output signals, therefore, more often lead to match dissolution with such workers. Formally, the variance of the learning process σ^2_p is larger for whites, i.e., $\sigma^2_{pB}(p) < \sigma^2_{pW}(p)$. Note, however, that the job loss rate differential still has ambiguous sign because it depends on the respective numbers of blacks and whites at risk of being dismissed. These numbers are defined by the slopes of the pdf of the ergodic belief distributions evaluated at belief thresholds.

Part C underlines the effect of differential selection in employment on wage disparities. (i) shows that blacks receive lower wages conditional on belief on match quality. At given match quality, blacks pay the price of lower output observability. However, they are more selected than whites. This implies that the belief distribution may be better for blacks than for whites at its bottom. Therefore (ii) shows that, unlike NS, the quantile differential tends to increase with quantile.

To summarize, SD can explain why discriminated workers stay longer in unemployment and receive lower wage conditional on type and belief on match quality. It also predicts the quantile differential should be increasing in quantile. However, SD also allows for the possibility where black workers enjoy longer employment episodes, which is at odds with the empirical evidence reported in Section 2. We now illustrate this claim through an example.

Suppose σ_{XB} is arbitrarily large and $\sigma_{XW} = 0$ so that job tenure does not provide information for blacks, whereas match quality is revealed right after hiring for whites. When σ_{XB} is arbitrarily large, the standard deviation is $\sigma_{pB}(p) = 0$ for all $p \in [0,1]$. It follows that the belief on match quality does not change with tenure. Consequently, the wage does not change with tenure and job separation only occurs for exogenous reasons. When $\sigma_{XW} = 0$, the belief immediately jumps after hiring 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 white applicants are hired. Thus $\underline{p}_W = 0 < \underline{p}_B$. The job-finding rate differential is $\Delta jfr = \lambda G^0(p_B) > 0$, whereas the job loss rate differential is $\Delta jlr = \lambda \int_0^1 (1-p) dG^0(p) > 0$. Therefore jobs occupied by blacks last longer.

Unemployment valuation heterogeneity.—We finally assume the utility flow derived from unemployment differs between blacks and whites.

Assumption 3 (Unemployment valuation): *The utility flow derived from unemployment is larger for blacks than for whites, i.e., $b_B > b_W$. Moreover, $G^0_B = G^0_W = G^0$ and $\sigma_{XB} = \sigma_{XW} = \sigma_X$.*

As usual in the literature, a larger b can be associated with a higher preference for leisure, either because blacks are less willing to work, have a larger home production, or have a better access to the informal sector. A larger b can also be due to lower unemployment stigma, something understandable in a community of workers over-exposed to unemployment.

Proposition 4 (Unemployment valuation). *Under Assumption 3, the following statements hold:*

A. Return to search: $U_B > U_W$.

B. Employment disparities:

- *(i)* $1 > \underline{p}_B > \underline{p}_W > 0$;
- *(ii) the job-finding rate differential is* $\Delta jfr = \lambda \left[G^0(\underline{p}_B) - G^0(\underline{p}_W) \right] > 0$;
- *(iii) the job-loss rate differential is* $\Delta jlr = \frac{1}{2} \left[\sigma_p^2(\underline{p}_B) \tilde{g}'_B(\underline{p}_B) - \sigma_p^2(\underline{p}_W) \tilde{g}'_W(\underline{p}_W) \right]$ *and may be positive or negative.*

C. Wage disparities:

- *(i) for all $p \in [\underline{p}_B, 1]$, the wage differential is $\Delta\omega(p) = (1-\beta) r (U_W - U_B) < 0$;*
- *(ii) the quantile differential is such that $z(0) < z(1) < 1$.*

Part A shows that blacks have a larger return to search. They enjoy higher utility flows in unemployment, which has permanent positive impacts on the value of being unemployed.

Part B describes the implications of Assumption 3 for employment differentials. (i) shows that blacks are more selected than whites. Matching with blacks generates lower match surplus. This implies (ii): blacks are less likely to form matches and their job-finding rate is smaller. Less can be said for the job separation rate differential because, here again, it depends on the derivative of the pdf of the belief distribution at belief threshold $\tilde{g}_i'(\underline{p}_i)$.

Part C (i) shows that blacks bargain higher wages conditional on match quality. This effect combined with the fact that blacks are more selected implies (ii): the quantile differential tends to be negative and increasing in quantile. Selection does not play any role at the top quantiles where all matches are good. Therefore the wage quantile differential tends to increase with quantile.

To summarize, higher UV for blacks can explain differences in employment outcomes. However, it also makes counterfactual predictions for wage outcomes, implying that blacks are paid more at given belief on match quality as well as on average. From an empirical perspective, UV heterogeneity is useful because it implies that the wage quantile differential increases with quantile.

4 Structural estimation

Our goal in this section is twofold. From a theoretical perspective, we want to quantify the internal trade-offs of our model when brought to data. On the empirical side, we want to measure the contribution of statistical discrimination to racial employment and wage disparities. To achieve both goals, we try to replicate a number of moments that exceeds the number of parameters. Therefore, our model is overidentified and we cannot simply calibrate it. Instead, we estimate its parameters by indirect inference.

We first present the estimation methodology, then turn to estimation results and discuss the implications of our estimates for the black-white differential return to tenure.

4.1 Econometric methodology

Indirect inference.—Following Gourieroux et al. (1993), we estimate the model by indirect inference. It consists of a simulated method of moments (SMM) estimator, in which some of the moments are estimated from reduced-form auxiliary models. Let $\boldsymbol{\theta}$ denote the vector of structural parameters, $\boldsymbol{m}^S(\boldsymbol{\theta})$ be the model-generated vector of parameters of the auxiliary models and $\boldsymbol{m}^D$ the corresponding empirical vector. The estimation procedure finds $\boldsymbol{\theta}$ such that the distance between the model-generated moments and their empirical counterparts is as small as possible.

Specifically, the set of estimated parameters minimizes the following function:

$$L(\boldsymbol{\theta}) = \left(\boldsymbol{m}^D - \boldsymbol{m}^S(\boldsymbol{\theta})\right)^T \boldsymbol{W} \left(\boldsymbol{m}^D - \boldsymbol{m}^S(\boldsymbol{\theta})\right), \tag{13}$$

where $\boldsymbol{W}$ is a weighting matrix. Assuming $L_N(\boldsymbol{\theta})$ is differentiable and attains its global minimum at the true parameter vector $\boldsymbol{\theta_0}$, a minimum verifies the following first-order condition:

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta_0}) = -2\frac{\partial \boldsymbol{m}^{S^T}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta_0})\boldsymbol{W}\left(\boldsymbol{m}^D - \boldsymbol{m}^S(\boldsymbol{\theta_0})\right) = 0.$$

Furthermore, assuming each moment in $\boldsymbol{m}^D$ is asymptotically normally distributed yields the following asymptotic distribution for $\hat{\boldsymbol{\theta}}$:[7]

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0}) \sim \mathcal{N}\left(0, \left(\boldsymbol{M}^T\boldsymbol{W}\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{W}\boldsymbol{S}\boldsymbol{W}\boldsymbol{M}\left(\boldsymbol{M}^T\boldsymbol{W}\boldsymbol{M}\right)^{-1}\right)$$

where N is the sample size, $\boldsymbol{M}$ the Jacobian matrix of the moment conditions with respect to the parameters, $\boldsymbol{M} = \frac{\partial \boldsymbol{m}^S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}(\hat{\boldsymbol{\theta}})$, and $\boldsymbol{S}$ the variance-covariance matrix of the empirical moments, $\boldsymbol{S} = V\left(\sqrt{N}(\boldsymbol{m}^D - \boldsymbol{m}^S(\boldsymbol{\theta_0}))\right)$.

We approximate $\boldsymbol{M}$ using two-sided finite differences, $\boldsymbol{S}$ is obtained by bootstrapping sample moments with 500 replications, and $\boldsymbol{W}$ is the estimated covariance matrix of the moments, $\boldsymbol{W} = \boldsymbol{S}$.

Parametric assumptions.—We fix some of the parameters to standard values. The monthly discount rate r is set to 0.0043, which is equivalent to 5% per annum. Workers' bargaining power β is arbitrarily set to 1/2. When parameters μ_H, μ_L, b_W and b_B are increased by a constant term, wages increase by the same constant and labor market transitions remain unchanged. Therefore they cannot be separately identified and we need to fix at least one of them. We choose $\mu_L = -\mu_H$ and leave b_W and b_B free. In practice, the productivity parameter differential $\mu_H - \mu_L$ must be sufficiently large so that the model can replicate the support of the empirical (residual) wage distribution.

[7] By the mean value theorem for some $\overline{\boldsymbol{\theta}}$ between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta_0}$, we have $\boldsymbol{m}^D - \boldsymbol{m}^S(\hat{\boldsymbol{\theta}}) = \boldsymbol{m}^D - \boldsymbol{m}^S(\boldsymbol{\theta_0}) + \frac{\partial \boldsymbol{m}^S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}(\overline{\boldsymbol{\theta}}) \times (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0})$, which is substituted into the first-order condition to obtain this result.

We also make parametric assumptions on the prior belief distributions. We suppose G_B^0 and G_W^0 have truncated log-normal distributions on the support $[0, p_{\max}]$ with $p_{\max} < 1$. For all $p \in [0, p_{\max}]$ and $i = B, W$,

$$G_i^0(p) = \frac{1}{p\eta_i\sqrt{2\pi}} \frac{\exp[-(\ln p - \gamma_i)^2/(2\eta_i^2)]}{\Phi[(\ln p_{\max} - \gamma_i)/\eta_i]}, \tag{14}$$

where γ_i and η_i are, respectively, the location and scale parameters of the distribution, and ϕ and Φ are, respectively, the pdf and the cdf of the standard normal distribution. The choice of log-normal functional forms is motivated by the fact that the distributions of entry wages are actually close to log-normal. The restriction $p_{\max} < 1$ guarantees that the assumption made right before Proposition 1 holds. Namely, we have $\lim_{p\to 1} g_i^0(p)(1-p)^{1/2-(1/4+2\delta/s_i^2)^{1/2}}/[A(1-p)^a] < 1$, which ensures that the top quantiles of the unconditional wage distribution are larger than the corresponding quantiles of the entry wage distribution. In practice, we set $p_{\max} = 0.9$ and check that the mass of the log-normal distribution above $p_{\max}$ is negligible.

We are left with the following vector of ten parameters to estimate $\boldsymbol{\theta} = \{\gamma_W, \eta_W, \gamma_B, \eta_B, \sigma_{XW}, \sigma_{XB}, b_W, b_B, \lambda, \delta\}$. The first four relate to the distributions of prior beliefs about match quality, σ_{XW} and σ_{XB} determine the group-specific standard deviations of output, b_W and b_B are the utility flows derived from unemployment, λ is the job offer rate common to both groups, and δ is the exogenous component of job separation.

Choice of moments and identification.—Here we discuss the different moments we use and how they contribute to the identification of the different parameters.

We emphasize employment and wage differentials between blacks and whites. Therefore we give as much weight to blacks as a group as to whites in the estimation procedure. We target 12 average labor market outcomes achieved by white workers and consider the 12 associated ethnic differentials. The 12 moments are described in Section 2 and follow the theoretical discussions in Section 3.2. As for wage outcomes, we consider five quantiles of the (residual) entry wage distribution and five quantiles of the (residual) overall wage distribution. As for employment outcomes, we consider the job-finding and job loss rates.

The quantiles and quantile differentials of the entry wage distributions are key to identify the parameters of the prior distributions G_B^0 and G_W^0. The quantiles of the unconditional wage distribution for whites are crucial to identify the standard deviation of output σ_{XW}: the difference between the entry wage distribution and the overall distribution depends on the speed of the learning process, which is inversely related to this variance. Quantile differentials give a first piece of information with regard to the respective magnitudes of NS, SD and UV heterogeneity. Proposition 2 shows NS implies quantile differentials tend to decrease with quantiles, whereas Propositions 3 and 4 suggest the opposite pattern when there is SD or UV heterogeneity.

The job-finding rate and job-finding rate differentials allow us to identify the job offer rate λ and the difference in probability of forming a match $G_B^0(\underline{p}_B) - G_B^0(\underline{p}_W)$. Once combined with the information derived from quantiles of entry wage distributions, this probability difference helps us to identify the threshold beliefs $\underline{p}_B$ and $\underline{p}_W$. Lastly the job separation rate and the job separation rate differential allow us to disentangle the exogenous and endogenous components of job loss. This procedure easily provides a value to the exogenous separation rate δ. The difference in normalized flows of employees who cross the threshold beliefs is $\sigma_p^2(\underline{p}_B)\tilde{g}_B'(\underline{p}_B)$ $-\sigma_p^2(\underline{p}_W)\tilde{g}_W'(\underline{p}_W)$. The derivative of the pdf of the belief distribution $\tilde{g}_i'(\underline{p}_i)$ depends, among other things, on the derivative of the pdf of the prior distribution evaluated in the belief threshold. The variance $\sigma_p^2(\underline{p}_i) = \underline{p}_i(1-\underline{p}_i)\frac{\mu_H-\mu_L}{\sigma_{Xi}}$ is inversely related to the standard deviation of output σ_{Xi}. Therefore we have additional information to identify the parameters of the prior distribution and the standard deviation of output.

Blacks and whites can differ in three ways, but the estimation forces many parameters to be the same across groups: the exogenous separation rate δ, the job offer rate λ, discount rate r, bargaining power β, output levels in a good match μ_H and in a bad one μ_L. Moreover, there are trade-offs between the different moments. In particular, fitting quantiles of the unconditional wage distribution, quantiles of the entry wage distribution and job separation is challenging. The learning process determines wage growth over tenure but also affects the stationary flow of employees who lose their job through the term σ_p^2. Similarly, the prior belief distribution shapes the entry wage distribution but also impacts the latter flow through the term $\tilde{g}_i'(\underline{p}_i)$.

4.2 Estimation results

Fit of the moments.—Table 2 compares the model outcomes with the empirical moments chosen for estimation. We run five specifications. The first three columns display the results when a single mechanism is at play, i.e., NS in column (1), SD in column (2) and differences in UV in column (3). Column (4) allows for both types of statistical discrimination (NS-SD). Finally, column (5) combines the three mechanisms (NS-SD-UV). Columns (6) and (7) contain the means and standard deviations of the bootstrapped moments that our estimation procedure matches. The goodness of fit of each specification is summarized by the maximized value of the criterion displayed by equation (13).

Table 2 shows two important results. On the one hand, models based on statistical discrimination alone (NS and SD) face a fundamental trade-off between fitting the quantile differentials of wage distributions and fitting the job loss differential. On the other hand, combining statistical discrimination with UV allows us to escape this trade-off.

All models fit reasonably well the different quantiles of the white wage distributions, the white job-finding probability and the white job loss probability. However, all models except the NS-SD-UV fail to reproduce the large positive and increasing quantile differentials of the unconditional wage distributions and the large job-finding and job loss rate differentials. In columns 1 to 4, the quantile differentials are far from the empirical ones, the job-finding rate differential is modestly positive and the job-loss rate differential is zero or negative. By contrast, the model combining NS-SD-UV in column 5 correctly fits the quantile and transition rate differentials. It slightly overestimates the job-finding rate differential (11.4 percentage points against 10.2 percentage points in the data) and accounts for 65% of the job loss rate differential.

We now describe parameter estimates before explaining these results in detail.

Parameter Estimates.—Tables 3 and 4 show the parameter estimates and some of the endogenous variables. For expositional purposes, we report the mean and standard deviation of the prior distribution, as opposed to the location and scale parameters.[8]

In columns 1, 2 and 4, only statistical discrimination is taken into account. The estimated parameters feature the expected situation where blacks endure both NS and SD. When NS is involved as in columns 1 and 4, whites enjoy a better prior distribution.[9] When SD is involved as in columns 2 and 4, output signals are

[8]The mean $\overline{p_0}_i$ and standard deviation $\sigma_{p_0 i}$ of the log-normal distribution with location and scale parameters γ_i and η_i are $\overline{p_0}_i = \exp(\gamma_i + \eta_i^2/2)$ and $\sigma_{p_0 i} = \sqrt{\exp(2\gamma_i + \eta_i^2)(\exp(\eta_i^2) - 1)}$. In our case, the priors have truncated log-normal distribution, however the latter formulas are good proxies due to the fact that the probability mass above the truncation point $p_{\max}$ is negligible in all of our estimations.

[9]In column 1, the standard deviation is larger for Blacks, which implies that the white distri-

Table 2: Fit of the moments

		Model					Data	
		NS	SD	UV	NS-SD	NS-SD-UV	Mean	St.-dev.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Whites								
Transitions								
	JFP	.294	.294	.288	.307	.330	.331	(.003)
	JLP	.024	.025	.024	.025	.025	.025	(.000)
Wages								
All Jobs	Min	.188	.185	.187	.187	.189	.181	(.002)
	25th perc.	.321	.318	.320	.325	.330	.323	(.003)
	50th perc.	.416	.411	.414	.420	.428	.415	(.004)
	75th perc.	.538	.533	.537	.542	.553	.536	(.004)
	95th perc.	.753	.748	.752	.755	.765	.761	(.005)
New Jobs	Min	.188	.185	.187	.187	.189	.181	(.002)
	25th perc.	.274	.265	.273	.267	.289	.296	(.004)
	50th perc.	.365	.353	.364	.354	.386	.377	(.005)
	75th perc.	.486	.472	.486	.471	.510	.492	(.006)
	95th perc.	.703	.688	.703	.682	.724	.727	(.010)
B. Racial Gaps								
Transitions								
	JFP	.012	.014	-.004	.056	.114	.102	(.008)
	JLP	.000	-.002	.000	-.003	.007	.011	(.001)
Wages								
All Jobs	Min	.027	-.015	.010	.010	.016	.019	(.003)
	25th perc.	.025	.004	.009	.037	.041	.037	(.003)
	50th perc.	.022	.012	.009	.042	.058	.054	(.004)
	75th perc.	.016	.018	.008	.037	.078	.064	(.005)
	95th perc.	.003	.022	.007	.012	.102	.060	(.010)
New Jobs	Min	.027	-.015	.010	.010	.016	.012	(.004)
	25th perc.	.026	-.011	.009	.013	.061	.035	(.005)
	50th perc.	.023	-.008	.008	.010	.095	.049	(.007)
	75th perc.	.017	-.006	.007	-.002	.127	.064	(.014)
	95th perc.	.003	-.004	.007	-.035	.170	.051	(.025)
Criterion		567.6	597.9	628.0	464.5	147.6		

Notes.—Model fit of the five specifications: Negative Stereotyping (NS) in column (1), Screening Discrimination (SD) in column (2), differences in Unemployment Valuation (UV) in column (3), both types of statistical discrimination (NS-SD) in column (4) and the three mechanisms (NS-SD-UV) in column (5). Columns (6) and (7) contain the means and standard deviations of the bootstrapped moments with 500 replications. Transition rates are at monthly frequency, wages are hourly. Panel A. corresponds to the levels of wages of white workers. Panel B. reports the racial gaps. Wages in all jobs refer to the unconditional residual wage distribution, wages in new jobs to the entry wage distribution.

Table 3: Parameter estimates

Parameter	NS (1)	SD (2)	UV (3)	NS-SD (4)	NS-SD-UV (5)
$\overline{p_0}_W$	0.566	0.556	0.566	0.557	0.586
	(.004)	(.004)	(.004)	(.004)	(.003)
$\overline{p_0}_B$	0.556	0.556	0.566	0.538	0.516
	(.004)	(.004)	(.004)	(.005)	(.003)
$\sigma_{p_0 W}$	0.083	0.083	0.084	0.081	0.081
	(.001)	(.001)	(.001)	(.001)	(.001)
$\sigma_{p_0 B}$	0.089	0.083	0.084	0.095	0.070
	(.002)	(.001)	(.001)	(.002)	(.002)
σ_{XW}	234.48	217.00	238.22	193.63	239.54
	(6.45)	(5.65)	(6.68)	(4.53)	(8.18)
σ_{XB}	234.48	324.17	238.22	398.83	165.84
	(6.45)	(19.13)	(6.68)	(28.78)	(7.35)
b_W	-2.025	-1.910	-1.984	-1.951	-2.488
	(.041)	(.039)	(.040)	(.040)	(.051)
b_B	-2.025	-1.910	-2.055	-1.951	-0.594
	(.041)	(.039)	(.044)	(.040)	(.040)
λ	0.522	0.556	0.510	0.569	0.528
	(.016)	(.021)	(.016)	(.020)	(.012)
δ	0.020	0.020	0.020	0.019	0.022
	(.000)	(.000)	(.000)	(.000)	(.000)

Notes.—Estimation by SMM of the five specifications: Negative Stereotyping (NS) in column (1), Screening Siscrimination (SD) in column (2), differences in Unemployment Valuation (UV) in column (3), both types of statistical discrimination (NS-SD) in column (4) and the three mechanisms (NS-SD-UV) in column (5). Bootstrapped standard errors in parentheses. $\overline{p_0}_i$ and $\sigma_{p_0 i}$ correspond to the mean and standard deviation of the group-i specific prior distribution, respectively.

noisier when the worker is black. In column 3, only UV is accounted for. UV is actually larger for whites than for blacks, which explains column 3 in Table 2 displays results opposite to Proposition 4.

We now explain the trade-off between fitting the quantile differentials and fitting the job loss differential. In our data, the quantile differentials of the (residual) wage distributions are large and increasing in quantile, whereas the job loss rate differential is substantial. Proposition 2 shows that NS predicts decreasing quantile differentials. Moreover, though generating positive job loss differential is theoretically possible, the estimated differential is nil. Proposition 3 shows that SD can predict increasing quantile differentials. However, SD also implies that whites lose their jobs more rapidly.

bution does not stochastically dominate the black one at first order. Thus the pdf of the black distribution is slightly higher than the pdf of the white one in the neighborhood of $p = 0.9$.

Table 4: Endogenous variables

Variable	NS (1)	SD (2)	UV (3)	NS-SD (4)	NS-SD-UV (5)
$\underline{p}_W$	0.526	0.524	0.526	0.523	0.526
$\underline{p}_B$	0.520	0.531	0.524	0.528	0.518
$g'(\underline{p}_W)$	301.13	297.91	304.77	266.59	269.07
$g'(\underline{p}_B)$	297.85	454.18	299.72	535.20	349.37
$\sigma_{pW} \times 100$	0.532	0.575	0.523	0.644	0.520
$\sigma_{pB} \times 100$	0.532	0.384	0.524	0.312	0.753
rU_W	0.248	0.250	0.246	0.261	0.246
rU_B	0.221	0.243	0.236	0.213	0.256

Notes.—Endogenous variables of interest of the five specifications: Negative Stereotyping (NS) in column (1), Screening Discrimination (SD) in column (2), differences in Unemployment Valuation (UV) in column (3), both types of statistical discrimination (NS-SD) in column (4) and the three mechanisms (NS-SD-UV) in column (5).

In column 5, statistical discrimination is combined with UV heterogeneity. The estimated parameter configuration corresponds to NS, anti-SD and larger UV for blacks. Blacks still draw their initial belief on match quality from a worse distribution than whites. However, they benefit from a faster learning process: now the standard deviation of output signals is larger for whites. Lastly, blacks enjoy higher utility flows once in unemployment.

Anti-SD guarantees that blacks lose their jobs faster, whereas NS implies that the job-finding rate differential is positive and blacks are paid less than whites. According to Propositions 2 and 3, NS and anti-SD predict decreasing quantile differentials. Meanwhile Proposition 4 shows that larger UV for blacks implies increasing quantile differentials, which explains why $b_B > b_W$ in column 5.

Anti-SD means the true match quality is easier to observe during employment when the worker is black. This result is at odds with standard assumptions in the literature. In their paper quantifying racial differences in unemployment, Ritter and Taylor (2011) argue managers face more difficulties to assess the productivity of black workers both at interview and during employment. They refer to the theory of language discrimination put forward by Lang (1986). According to this theory, blacks can be seen as speaking a different language, generating transaction costs within firms. Ritter and Taylor add that managers who had to choose between reducing the white noise or the black one would prefer reducing the white noise because whites are more numerous. Cavounidis and Lang (2015) argue against this view, thereby providing a possible rationale to the result of anti-SD. They also consider managers' incentive to supervise the different groups of workers. When blacks occupy unproductive jobs more often than whites, managers spend more resources monitoring blacks.

UV heterogeneity is here beneficial to blacks. This can be interpreted in terms of heterogenous preference for leisure, domestic production, access to the informal sector, or unemployment stigma. A controversial implication of our estimation is that blacks and whites have the same expected utility when unemployed. Thus the model can rationalize large residual disparities in terms of wage and transition rates despite the typical black worker enjoys the same utility level as the typical white worker.

In all estimates, the utility flows b_B and b_W obtained in unemployment are negative. In job search models, what matters is the utility differential between the employment and unemployment states and not the utility level obtained in each state. We just lose the ability to measure the unemployment utility flow in percentage of the wage.

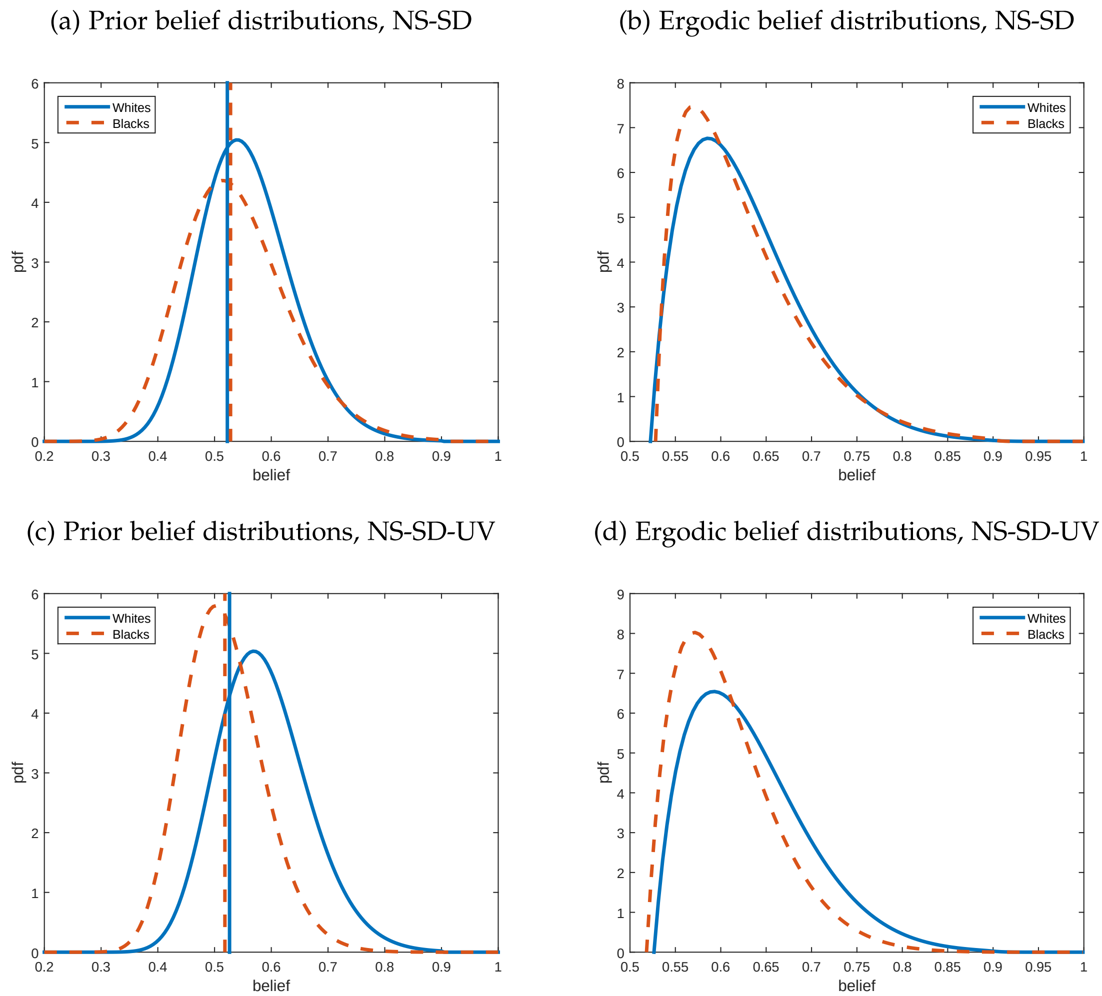
Belief and wage distributions.—Figure 2 depicts the initial and ergodic belief distributions in two cases: NS-SD vs NS-SD-UV. The ergodic distributions are more dispersed than the initial distributions and feature fat right tails. In both cases, the threshold beliefs are very close. Blacks are slightly more selected in the NS-SD case, but slightly less selected in the NS-SD-UV case. Under NS-SD, the prior distribution of blacks has a larger variance. This property extends to the ergodic distribution where the pdf is slightly higher for blacks at large beliefs. Under NS-SD-UV, both the mean and the variance of the prior distribution are larger for whites. Therefore the initial and ergodic white distributions stochastically dominate at first order the corresponding black distributions.

Figure 3 shows the corresponding entry and unconditional wage distributions. They plot the predicted and empirical wage distributions in the cases of NS-SD and NS-SD-UV. In both cases, the entry wage distribution is less well fitted than the unconditional wage distribution. The estimation procedure uses efficient weighting, which implies that the moments with the largest variance receive the lowest weight in the loss function. Quantiles of the entry wage distribution are less precisely estimated because fewer workers are concerned.

4.3 Returns to tenure

In this section we discuss the returns to tenure for blacks and whites. Fryer Jr et al. (2013) use a dataset from the Princeton University Survey Research Center

Figure 2: Prior and ergodic belief distributions



Notes.— Vertical lines correspond to threshold beliefs $\underline{p}_W$ and $\underline{p}_B$.

and estimate that the return to tenure is larger for blacks by 1.1 percentage points. They rationalize this result through a stylized three-period model of statistical discrimination. NS implies blacks are more selected than whites. Therefore the scope for wage improvement is larger for blacks than for whites and the wage gap decreases with tenure.

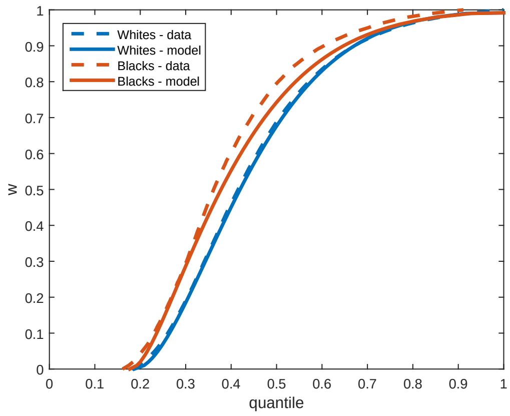
We simulate individual labor market histories for the different specifications of our model. We discretize time at monthly frequency and compute 500-month long histories for 25,000 individuals of each group. The details of the algorithm are provided in Appendix F. We then perform the following OLS regression:

$$\ln w_i = a_0 + a_1 Black_i + a_2 Tenure_i + a_3 Black_i \times Tenure_i + \varepsilon_i. \qquad (15)$$

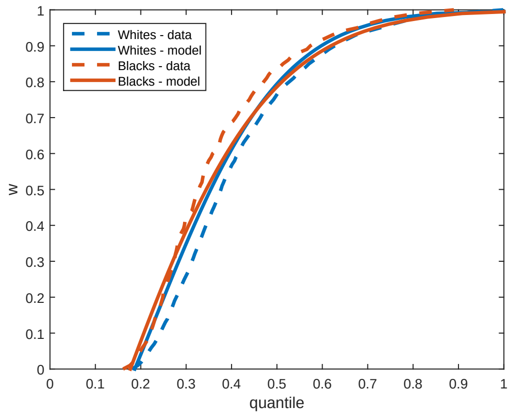
Table 5 reports the estimates. Each column corresponds to a particular specification of the model. Statistical discrimination alone does not predict the narrowing

Figure 3: Fit to wage moments

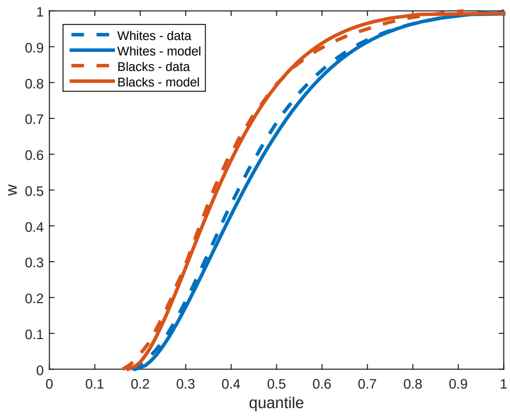
(a) All jobs, NS-SD



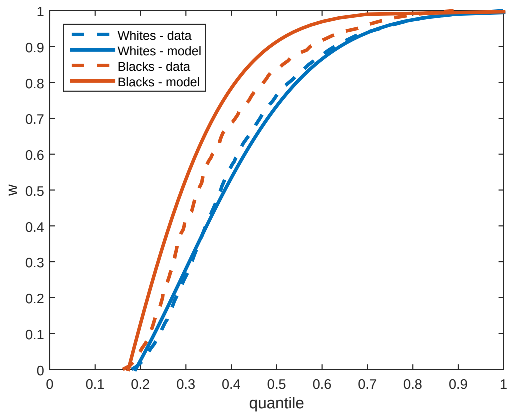
(b) New jobs, NS-SD



(c) All jobs, NS-SD-UV



(d) New jobs, NS-SD-UV



Notes.— Wages in all jobs refer to the unconditional residual wage distribution and wages in new jobs to the entry wage distribution.

of racial differences in wages with tenure, as reported by Fryer et al. Column 1 corresponds to the case advocated by Fryer et al. NS effectively implies that the return to tenure is larger for blacks. However, the estimated belief thresholds are very close to each other so that selection effects are quantitatively small. In column 4 NS is combined with SD. SD dominates selection effects induced by NS and the predicted return to tenure is larger for whites by 1 percentage point. By contrast, the combination of NS, SD and UV implies that the return to tenure is larger for blacks by 1.9 percentage points. The small R^2 in all regressions reflect that random draws on match quality account for a very large part of wages at all tenures. Moreover the wage predicted by our model is not log-linear in tenure; a log-linear regression therefore leads to systematic errors that reduce the R^2. Adding tenure squared and its interaction with the racial dummy increases the R^2 by 5 percentage

Table 5: Black and white estimated returns to tenure on simulated data

	NS (1)	SD (2)	UV (3)	NS-SD (4)	NS-SD-UV (5)
Black	-.063	.011	-.026	-.044	-.204
Tenure	.017	.019	.016	.021	.016
Tenure × Black	.001	-.007	.001	-.010	.019
Constant	-.954	-.974	-.956	-.966	-.918
R^2	.032	.026	.026	.035	.097

Notes.— Results of the OLS estimation of equation (15) using pooled data from the simulated individual histories as implied by the estimates of the five specifications: Negative Stereotyping (NS) in column (1), Screening Discrimination (SD) in column (2), differences in Unemployment Valuation (UV) in column (3), both types of statistical discrimination (NS-SD) in column (4) and the three mechanisms (NS-SD-UV) in column (5). Tenure is in years. Simulated data for each specification are obtained by generating 500-month long history for 25,000 individuals of each group. Total number of observations ≈ 23,000,000 depending on specification.

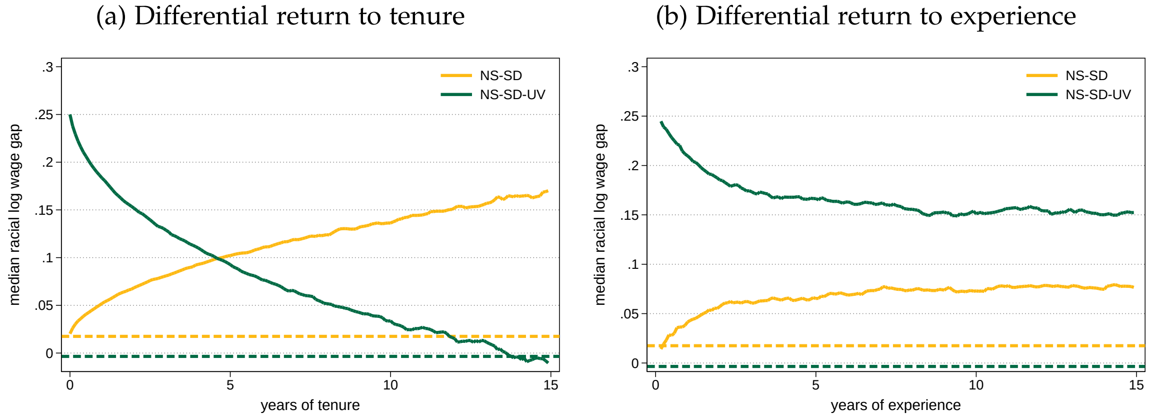
points on average.

The estimates performed from the NS-SD-UV sample have the same order of magnitude as Fryer et al. However, the economic mechanism strongly differs from theirs. They emphasize selection effects induced by NS. Here selection effects can be neglected and the differential return to tenure is entirely due to anti-SD.

Figure 4a depicts the predicted log median wage differential by tenure for the NS-SD and NS-SD-UV specifications. The horizontal line is the minimum log wage differential between the two groups reflecting the differential return to search. In the NS-SD case, the median wage differential strictly increases with tenure, starting relatively low around 2.5% and reaching over 15% after 15 years. In the NS-SD-UV specification, the median wage differential strictly decreases, starting relatively large around 25% and almost reaching 0 after 15 years. The slope of each curve decreases with tenure in absolute value. This confirms that the differential return to tenure decreases with tenure.

The differential return to tenure does not coincide with the differential return to experience because blacks and whites have different job durations. Figure 4b shows the log median wage differential by experience for the NS-SD and NS-SD-UV estimates. The NS-SD curve is below the NS-SD-UV curve partly reflecting the poor ability of the NS-SD case to fit the large quantile differentials of the unconditional wage distributions. In the NS-SD estimate, the job loss rate is slightly larger for whites, so that the differential return to experience is increasing but at a smaller pace than the differential return to tenure. In the NS-SD-UV specification, jobs last longer for whites, so that the differential return to experience is larger than the dif-

Figure 4: Differential return to tenure and experience according to model specification



Notes.— Median log wage gap is defined as the black-white difference in median log wage by tenure. It is calculated from the simulated data obtained by generating 500-month long history for 25,000 individuals of each group. The dashed lines measure the wage gap that results from differences in workers' outside options.

ferential return to tenure. Overall, the predicted differential return to experience is slightly decreasing at low experience and roughly constant over 15 years.

5 Conclusion

In the US, black workers spend more time in unemployment, lose their jobs more rapidly, and earn lower wages than white workers. This paper quantifies the contributions of statistical discrimination, as portrayed by negative stereotyping and screening discrimination, to such employment and wage disparities. We develop an equilibrium search model of statistical discrimination with learning based on Moscarini (2005) and estimate it by indirect inference. We show that statistical discrimination alone cannot simultaneously explain the observed differences in residual wages and monthly job loss probabilities between black and white workers. However, a model with negative stereotyping, larger unemployment valuation and faster learning about the quality of matches for black workers can account for these facts. One implication of our findings is that black workers have larger returns to tenure.

There are several avenues for research. First, the model and estimation methodologies can be applied to alternative datasets and groups of workers. Second, it would be worth providing microfoundations to exogenous key parameters like the ethnic-specific degree of observability of true match quality (see, e.g., Cavounidis and Lang (2015) who argue that managers have more incentive to monitor blacks because these workers historically had lower skills). A third set of extensions would enrich the current model. For instance, our model neglects human capital accumulation as a potential factor of wage growth. Estimating the effects of experience in structural search models is a non-trivial task (see Bagger et al. (2014)). The problem is a difficult one in our case where learning offers a competitive explanation for wage growth within firms. Therefore, we mention it as a limit of our work and call for additional research.

References

Aigner, D. J. and G. G. Cain, "Statistical theories of discrimination in labor markets," *Industrial and Labor Relations Review* 30 (1977), 175–187.

Arrow, K. J., "The Theory of Discrimination," in O. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets* (Princeton: Princeton University Press, 1973).

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

Becker, G. S., "The Economics of Discrimination," *University of Chicago Press Economics Books* (1971).

Black, D., "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics* 13 (1995), 309–334.

Borowczyk-Martins, D., J. Bradley and L. Tarasonis, "Racial discrimination in the US labor market: Employment and wage differentials by skill," *Labour Economics* 49 (2017), 106–127.

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

Cavounidis, C. and K. Lang, "Discrimination and Worker Evaluation," Technical Report, National Bureau of Economic Research, 2015.

Cornell, B. and I. Welch, "Culture, Information, and Screening Discrimination," *Journal of Political Economy* 104 (1996), 542–71.

DellaVigna, S. and M. D. Paserman, "Job Search and Impatience," *Journal of Labor Economics* 23 (2005), 527–588.

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

Elsby, M. W., B. Hobijn and A. Sahin, "The labor market in the Great Recession," Technical Report, National Bureau of Economic Research, 2010.

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

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

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

Gayle, G.-L. and L. Golan, "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997," *Review of Economic Studies* 79 (2012), 227–267.

Gourieroux, C., A. Monfort and E. Renault, "Indirect inference," *Journal of Applied Econometrics* 8 (1993), S85–S118.

Holden, S. and Å. Rosén, "Discrimination and employment protection," *Journal of the European Economic Association* 12 (2014), 1676–1699.

Jovanovic, B., "Matching, Turnover, and Unemployment," *Journal of Political Economy* 92 (1984), 108–22.

Lang, K., "A Language Theory of Discrimination," *Quarterly Journal of Economics* 101 (1986), 363–382.

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

Mailath, G. J., L. Samuelson and A. Shaked, "Endogenous inequality in integrated labor markets with two-sided search," *American Economic Review* 90 (2000), 46–72.

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

Mortensen, D. T. and C. A. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies* 61 (1994), 397–415.

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

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

Phelps, E. S., "The Statistical Theory of Racism and Sexism," *American Economic Review* 62 (1972), 659–61.

Ritter, J. A. and L. J. Taylor, "Racial disparity in unemployment," *Review of Economics and Statistics* 93 (2011), 30–42.

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

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

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

A Measuring the ins and outs of unemployment

To measure the monthly probability that an employed worker becomes unemployed, JLP_t, and the monthly probability that an unemployed worker finds a job, JFP_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 than a month; then JFP_t can be backed out from the data using:

$$u_{t+1} = (1 - JFP_t)u_t + u_{t+1}^s, \tag{16}$$

which implies

$$JFP_t = 1 - \frac{u_{t+1} - u_{t+1}^s}{u_t}. \tag{17}$$

To compute the job loss probabilities, we account for time aggregation bias and solve the following equation in jlr_t:

$$u_{t+1} = \frac{jfr_t}{jfr_t + jlr_t}(1 - e^{-jfr_t - jlr_t})l_t + e^{-jfr_t - jlr_t}u_t, \tag{18}$$

where l_t is the labor force and $jfr_t = -\ln(1 - JFP_t)$. Then, $jlr_t = -\ln(1 - JLP_t)$.

B Value functions

Let $w_{\alpha i}(p)$ be the wage and $W_{\alpha i}(p)$ be the value of holding a job when the belief on match quality is p. 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 workers' values solve the Hamilton-Jacobi-Bellman (HJB) equations:

$$rU_{\alpha i} = b_i\alpha + \lambda \int \max\{W_{\alpha i}(p) - U_{\alpha i}, 0\}\, dG_i^0(p), \tag{19}$$

$$rW_{\alpha i}(p) = w_{\alpha i}(p) + \frac{1}{2}\sigma_{pi}^2(p)W_{\alpha i}''(p) + \delta[U_{\alpha i} - W_{\alpha i}(p)], \tag{20}$$

The value of opening a vacancy is arbitrarily set to zero. In Appendix C, we close the model and introduce a standard constant returns to scale matching function and costly entry for firms. The value of a filled job with belief p solves the following HJB equation:

$$rJ_{\alpha i}(p) = \alpha\overline{\mu}(p) - w_{\alpha i}(p) + \frac{1}{2}\sigma_{pi}^2(p)J_{\alpha i}''(p) - \delta J_{\alpha i}(p). \tag{21}$$

Conditional on belief p, the equilibrium wage is pinned down by the generalized

Nash bargaining solution so that

$$W_{\alpha i}(p) - U_{\alpha i} = \beta S_{\alpha i}(p), \tag{22}$$

where $\beta \in [0,1]$ is the worker' bargaining power and $S_{\alpha i}(p) \equiv W_{\alpha i}(p) - U_{\alpha i} + J_{\alpha i}(p)$ is the total match surplus.

The match is stopped when the posterior belief of a good match falls to a belief, $\underline{p}_{\alpha i}$, where the firm-worker pair separate endogenously and restart searching on their own. Using (20) and (21), we can rewrite the total surplus as the following second-order differential equation:

$$S_{\alpha i}(p) = \frac{\alpha\overline{\mu}(p) + \frac{1}{2}\sigma_{pi}^2(p)S''_{\alpha i}(p) - rU_{\alpha i}}{r+\delta}, \tag{23}$$

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:

$$S_{\alpha i}(p) = c_{\alpha i} p^{\frac{1}{2}-\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}}(1-p)^{\frac{1}{2}+\sqrt{\frac{1}{4}+2\frac{r+\delta}{s_i^2}}} + \frac{\alpha\overline{\mu}(p) - rU_{\alpha i}}{r+\delta} \tag{24}$$

where $s_i \equiv (\mu_H - \mu_L)/\sigma_{Xi}$ is the signal-to-noise ratio. The coefficient $c_{\alpha i}$ and the optimal stopping belief $\underline{p}_{\alpha i}$ solve the system of value matching and smooth pasting equations. Existence and uniqueness of the solution is given in Appendix D. The resulting match surplus increases with belief $p \in [\underline{p}_{\alpha i}, 1]$.

Nash bargaining implies $\beta J''_{\alpha i}(p) = (1-\beta)W''_{\alpha i}(p)$. Using this fact yields a simple expression for the equilibrium wage:

$$w_{\alpha i}(p) = \beta\alpha\overline{\mu}(p) + (1-\beta)rU_{\alpha i}. \tag{25}$$

Lemma 1 (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., $\underline{p}_{\alpha i} = \underline{p}_i$.

C 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 $\lambda(x) = Ax^a$, $0 < a < 1$, where x 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

$$rV = -\kappa + \frac{\lambda(x)}{x} \sum_i m_i \int \int_{\underline{p}_i}^{1} [J_{i\alpha}(p) - V] dG_i^0(p) d\Psi_i(\alpha). \tag{26}$$

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

$$\kappa = \frac{\lambda(x)}{x} \sum_i m_i \int \int_{\underline{p}_i}^{1} \alpha J_i(p) dG_i^0(p) d\Psi_i(\alpha). \tag{27}$$

Thus for a given set of parameter estimates and a given θ, we compute the right-hand side of equation (27) with $\lambda(x) = \lambda$, set $A = \lambda x^{-a}$ and κ as the left-hand side of equation (27).

D The solution of the HJB equation

The solution is a slight amendment to Moscarini (2005). We first plug the match surplus solution (24) into the value of unemployment (19), use the Nash bargaining solution (22) and obtain the return to search:

$$rU_{\alpha i} = \frac{b\alpha + \beta\lambda \int_{p \geq \underline{p}_{\alpha i}} \left[c_{\alpha i} p^{\frac{1}{2} - \sqrt{\frac{1}{4} + 2\frac{r+\delta}{s_i^2}}} (1-p)^{\frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{r+\delta}{s_i^2}}} + \frac{\alpha \overline{\mu}(p)}{r+\delta} \right] dG_i^0(p)}{1 + \frac{\beta\lambda(1 - G_i^0(\underline{p}_{\alpha i}))}{r+\delta}}. \tag{28}$$

We then check that the proposed function solves equation (23) and show that the optimal stopping belief is well-defined and belongs to the interval $(0, 1)$. Hereafter we neglect type α and demographic group i. Let $n = \sqrt{1/4 + 2(r+\delta)/s^2}$, $s \equiv (\mu_H - \mu_L)/\sigma_X$. We obtain

$$\begin{aligned} S'(p) &= cp^{-1/2-n}(1-p)^{-1/2+n}(1/2 - n - p) + \frac{\mu_H - \mu_L}{r+\delta}, \\ S''(p) &= -c(1/4 - n^2)p^{-3/2-n}(1-p)^{-3/2+n}. \end{aligned}$$

Plugging $S'(p)$ and $S''(p)$ into (23) shows that (24) defines the solution. Moreover, c and $\underline{p}$ solve the following system of equations:

$$S'(\underline{p}) = c\underline{p}^{-1/2-n}(1-\underline{p})^{-1/2+n}(1/2 - n - \underline{p}) + \frac{\mu_H - \mu_L}{r+\delta} = 0, \tag{29}$$

$$S(\underline{p}) = c\underline{p}^{\frac{1}{2}-n}(1-\underline{p})^{\frac{1}{2}+n} + \frac{\overline{\mu}(\underline{p}) - rU}{r+\delta} = 0, \tag{30}$$

The value of unemployment and the optimal stopping belief solve $rU = rU_1(\underline{p}) = rU_2(\underline{p})$, where

$$
\begin{aligned}
rU_1(x) &= \mu_L + x(\mu_H - \mu_L) + (\mu_H - \mu_L)\frac{x(1-x)}{n+x-1/2}, && (31)\\
rU_2(x) &= \frac{b + \beta\lambda\int_{p\geq x}\left[c(x)p^{1/2-n}(1-p)^{1/2+n} + \frac{\bar{\mu}(p)}{r+\delta}\right]dG^0(p)}{1+\beta\lambda[1-G^0(x)]/(r+\delta)}, && (32)\\
c(x) &= \frac{\mu_H - \mu_L}{n+x-1/2}x^{1/2+n}(1-x)^{1/2-n}/(r+\delta). && (33)
\end{aligned}
$$

Let $\phi(x) = rU_1(x) - rU_2(x)$. We have

$$
\begin{aligned}
\phi(0) &= \frac{(\mu_L - b)(r+\delta) - \beta\lambda(\mu_H - \mu_L)\int_0^1 p dG^0(p)}{r+\delta+\beta\lambda} < 0,\\
\phi(1) &= \mu_H - b > 0,
\end{aligned}
$$

by assumption (7). Therefore there is $\underline{p} \in (0,1)$ such that $\phi(\underline{p}) = 0$. Moreover,

$$
\begin{aligned}
rU_1'(x) &= \frac{(\mu_H - \mu_L)(n^2 - 1/4)}{(n+x-1/2)^2},\\
\{1+\beta\lambda[1-G^0(x)]/(r+\delta)\}rU_2'(x) &= -\beta\lambda\left[c(x)x^{1/2-n}(1-x)^{1/2+n} + \frac{\bar{\mu}(x)}{r+\delta} - \frac{rU_2(x)}{r+\delta}\right]g^0(x)\\
&\quad +\beta\lambda\int_x c'(x)p^{1/2-n}(1-p)^{1/2+n}dG^0(p),\\
c'(x) &= \frac{(\mu_H - \mu_L)(n^2 - 1/4)}{(n+x-1/2)^2}x^{-1/2+n}(1-x)^{-1/2-n}/(r+\delta).
\end{aligned}
$$

In $x = \underline{p}$, we have

$$
\begin{aligned}
\phi'(\underline{p}) &= \frac{(\mu_H - \mu_L)(n^2 - 1/4)}{(n+\underline{p}-1/2)^2} - \frac{\beta\lambda}{1+\beta\lambda[1-G^0(\underline{p})]/(r+\delta)}\int_{\underline{p}} c'(\underline{p})p^{1/2-n}(1-p)^{1/2+n}dG^0(p)\\
&> \frac{(\mu_H - \mu_L)(n^2 - 1/4)}{(n+\underline{p}-1/2)^2}\left[1 - \int_{p\geq\underline{p}}(p/\underline{p})^{1/2-n}((1-p)/(1-\underline{p}))^{1/2+n}dG^0(p)\right] > 0
\end{aligned}
$$

because the term $p^{1/2-n}(1-p)^{1/2+n}$ decreases with p on the interval $[0,1]$. It follows that $\underline{p}$ is uniquely defined by the requirement $\phi(\underline{p}) = 0$.

E Proofs

Proof of Proposition 1. Ergodic belief distribution

We neglect the index i. Let $y(p) = p^2(1-p)^2 g(p)$. We have

$$
\begin{cases}
y''(p) + f(p) - (\nu^2 - 1/4)\frac{y(p)}{p^2(1-p)^2} = 0,\\
\int_{\underline{p}}^1 g(p)dp = m - u \in [0,m],\\
y(\underline{p}) = 0.
\end{cases}
\qquad (*)
$$

The general solution to problem $(*)$ has the form:

$$\begin{aligned} \nu y(p) &= [p(1-p)]^{1/2-\nu} \times \\ &\left\{ c_2 p^{2\nu} + c_1(1-p)^{2\nu} + p^{2\nu} \int_p^1 \frac{\lambda u g^0(x) x^{1/2-\nu}(1-x)^{1/2+\nu}}{2} dx \right. \\ &\left. -(1-p)^{2\nu} \int_p^1 \frac{\lambda u g^0(x) x^{1/2+\nu}(1-x)^{1/2-\nu}}{2} dx \right\}, \end{aligned} \tag{34}$$

where c_1 and c_2 are two constant terms. According to the first boundary condition, $\int_{\underline{p}}^1 g(p)dp$ converges, which implies that $c_2 = 0$. The second boundary condition gives $c_1 = \lambda uk$. The unnormalized density results by dividing $y(p)$ by $p^2(1-p)^2$. Three terms remain. The first term is always definite when $p < 1$. The second term is finite for all $p < 1$ when $g_i^0(1)$ is finite. As for the last term, the integral is finite under the assumption that there exist $A < \infty$ and $a > -1$ such that $\lim_{x\to 1} g_i^0(x)(1-x)^{1/2-\nu_i}/[A(1-x)^a] < 1$.

Proof of Proposition 2. Negative stereotyping

Part A. Suppose $U_W \leq U_B$. Using (19) and (22), the value of unemployment can be rewritten as

$$rU_W = b + \beta\lambda \int_0^1 \max\{S_W(p), 0\} g_W^0(p) dp. \tag{35}$$

The surplus equation (23) implies that $S_W''(p) = S_B''(p)$ and thus $S_W(p) \geq S_B(p)$, for all $p \in [0,1]$. Using this fact

$$rU_W \geq b + \beta\lambda \int_0^1 \max\{S_B(p), 0\} g_W^0(p) dp. \tag{36}$$

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

$$b + \beta\lambda \int_0^1 \max\{S_B(p), 0\} g_W^0(p) dp > b + \beta\lambda \int_0^1 \max\{S_B(p), 0\} g_B^0(p) dp = rU_B. \tag{37}$$

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

Part B. (i) Using equations (29) and (30), we obtain

$$rU_i = \mu_L + (\mu_H - \mu_L)(n + 1/2) \frac{\underline{p}_i}{n + \underline{p}_i - 1/2}. \tag{38}$$

The ratio is strictly increasing in $\underline{p}_i$. From part A, we have $U_W > U_B$, which implies that $\underline{p}_W > \underline{p}_B$.

(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.

(iii). At any time, the flow number of group-i workers who lose their job is $\delta(1 - u_i) + 0.5\sigma^2(\underline{p}_i)g_i'(\underline{p}_i)$. Thus the group-$i$ specific job-loss rate is $jlr_i = \delta + 0.5\sigma^2(\underline{p}_i)g_i'(\underline{p}_i)/(1 - u_i)$, $i = B, W$. The result follows.

Part C. (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.

(ii) We have $\omega_{iq} = \beta(\mu_H - \mu_L)\tilde{G}_i^{-1}(q) + (1 - \beta)rU_i + \beta\mu_L$. Therefore $z(0) = \beta(\mu_H - \mu_L)(\underline{p}_W - \underline{p}_B) + (1 - \beta)r(U_W - U_B) > z(1) = (1 - \beta)r(U_W - U_B)$ by parts A and B (i).

Proof of Proposition 3. Screening discrimination

Part A. Let $U \equiv U(\sigma_X)$ denote the value of unemployment when the standard deviation of output is σ_X. By construction, we have $U_B = U(\sigma_B)$ and $U_W = U(\sigma_W)$. Using the notations of Appendix D, we have $rU(\sigma_X) = rU_1(\underline{p}, \sigma_X) = rU_2(\underline{p}, \sigma_X)$, where the dependence vis-à-vis σ_X has been highlighted. Similarly we define $c(x, \sigma_X) \equiv c(x)$.

We now show that $rU'(\sigma_X) < 0$. Let $f(x, n) = \frac{1}{n+x-1/2}\left[\frac{x/(1-x)}{p/(1-p)}\right]^n$. We have

$$\begin{aligned}
\frac{\partial(rU_1)}{\partial\sigma_X} &= -(\mu_H - \mu_L)\frac{x(1-x)}{(n+x-1/2)^2}\frac{dn}{d\sigma_X} < 0, \\
\frac{\partial(rU_2)}{\partial\sigma_X} &= \frac{\beta\lambda\int_x \frac{\partial}{\partial n}(c(x,n)p^{1/2-n}(1-p)^{1/2+n})dG^0(p)}{1 + \frac{\beta\lambda}{r+\delta}[1 - G^0(x)]}\frac{dn}{d\sigma_X}.
\end{aligned}$$

But $\frac{\partial}{\partial n}(c(x,n)p^{1/2-n}(1-p)^{1/2+n})$ has the sign of $f_n(x, n) = \ln\left[\frac{x/(1-x)}{p/(1-p)}\right]f(x, n) - \frac{f(x,n)}{n+x-1/2} < 0$. Therefore $\frac{\partial(rU_1)}{\partial\sigma_X} < 0$ and $\frac{\partial(rU_2)}{\partial\sigma_X} < 0$ for all $x \in (0, 1)$. It follows that $rU'(\sigma_X) < 0$ and so $U(\sigma_{X_W}) > U(\sigma_{X_B})$.

Part B. (i) Let $\phi(x, \sigma_X) = rU_1(x, \sigma_X) - rU_2(x, \sigma_X)$. We have $d\underline{p}/d\sigma_X = -\frac{\phi_{\sigma_X}(\underline{p},\sigma_X)}{\phi_x(\underline{p},\sigma_X)}$, which has the sign of $-\phi_{\sigma_X}(\underline{p},\sigma_X)$. As shown in the proof of Part A, we have

$$\phi_{\sigma_X}(\underline{p},\sigma_X) = -\frac{(\mu_H-\mu_L)\underline{p}(1-\underline{p})}{(n+\underline{p}-1/2)^2}\frac{dn}{d\sigma_X} + \frac{\beta\lambda[1-G^0(\underline{p})]}{r+\delta+\beta\lambda[1-G^0(\underline{p})]}\frac{dn}{d\sigma_X}\times$$
$$\frac{(\mu_H-\mu_L)\underline{p}(1-\underline{p})}{n+\underline{p}-1/2}\int_{\underline{p}}^{1}\left[\frac{1}{n+\underline{p}-1/2}-\ln\frac{p/(1-p)}{\underline{p}/(1-\underline{p})}\right]p(1-p)\left(\frac{\underline{p}/(1-\underline{p})}{p/(1-p)}\right)^n\frac{dG^0(p)}{1-G^0(\underline{p})}$$
$$= -\frac{(\mu_H-\mu_L)\underline{p}(1-\underline{p})}{(n+\underline{p}-1/2)^2}\frac{dn}{d\sigma_X}\times\left\{1-\frac{\beta\lambda[1-G^0(\underline{p})]}{r+\delta+\beta\lambda[1-G^0(\underline{p})]}\times\right.$$
$$\left.\int_{\underline{p}}^{1}\left[1-(n+\underline{p}-1/2)\ln\frac{p/(1-p)}{\underline{p}/(1-\underline{p})}\right]\left(\frac{\underline{p}/(1-\underline{p})}{p/(1-p)}\right)^n\frac{p(1-p)dG^0(p)}{1-G^0(\underline{p})}\right\}$$
$$< -\frac{(\mu_H-\mu_L)\underline{p}(1-\underline{p})}{(n+\underline{p}-1/2)^2}\frac{dn}{d\sigma_X}\left\{1-\frac{1}{2}\int_{\underline{p}}^{1}\left[1-(n+\underline{p}-1/2)\ln\frac{p/(1-p)}{\underline{p}/(1-\underline{p})}\right]\left(\frac{\underline{p}/(1-\underline{p})}{p/(1-p)}\right)^n\frac{dG^0(p)}{1-G^0(\underline{p})}\right\}.$$

which has the sign of $I=\frac{1}{2}\int_{\underline{p}}\left[1-(n+\underline{p}-1/2)\ln\frac{p/(1-p)}{\underline{p}/(1-\underline{p})}\right]\left(\frac{\underline{p}/(1-\underline{p})}{p/(1-p)}\right)^n\frac{dG^0(p)}{1-G^0(\underline{p})}-1$. Let $f(y)=y^n\ln y$. The function f is such that $f(0)=f(1)=0$ and $f'(y)=y^{n-1}(1+n\ln y)$. Therefore $f(y)\geq-(ne)^{-1}$ for all $y\in[0,1]$. It follows that $I<\frac{1}{2}+\frac{n+\underline{p}-1/2}{ne}-1<1/e-1/2<0$. Thus $d\underline{p}/d\sigma_X>0$ and $\underline{p}_B>\underline{p}_W$.

(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).

(iii). See the proof of Part B (iii) of Proposition 2.

Part C. (i) The result follows from Part A.

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

Proof of Proposition 4. Unemployment valuation heterogeneity

Part A. Let $U\equiv U(b)$ denote the value of unemployment when the utility flow derived from unemployment is b. By construction, we have $U_B=U(b_B)$ and $U_W=U(b_W)$. Using the notations of Appendix D, we have $rU(b)=rU_1(\underline{p})=rU_2(\underline{p},b)$, where the dependence vis-Ã -vis b has been highlighted. We have $\partial(rU_2(x,b))/\partial b>0$ for all $x\in[0,1]$, which implies that $rU'(b)>0$. Therefore $U(b_B)>U(b_W)$.

Part B. (i) Let $\phi(x,b)=rU_1(x)-rU_2(x,b)$. We have $d\underline{p}/db=-\phi_b(\underline{p},b)/\phi_x(\underline{p},b)$, which has the sign of $-\phi_b(\underline{p},b)$. But the proof of Part A shows that $\phi_b(\underline{p},b)=-\partial(rU_2(\underline{p},b))/\partial b<0$. Therefore $d\underline{p}/db>0$ and $\underline{p}_B>\underline{p}_W$.

(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).

(iii). See the proof of part B (iii) of Proposition 2.

Part C. (i) The result follows from Part A.

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

F Algorithm to simulate individual histories

To simulate individual labor market histories, we discretize time at monthly frequency and use the following steps to compute 500-month long histories for 25,000 individuals of each group.

1. All individuals start unemployed in the first period.

2. At the beginning of each period, unemployed find a job with probability $1 - \exp(-\lambda[1 - G_i^0(\underline{p}_i)])$; the initial prior about match quality p_{i0} is drawn from the race-specific truncated distribution $G_i^0(p|p_{i0} > \underline{p}_i)$. The true match quality is determined by an additional draw where the probability of being in a good-quality match is $p_{i0} = \Pr(\mu = \mu_H)$.

3. At the beginning of each period, employed workers transit to unemployment with probability $1 - \exp(-\delta)$.

4. Employed workers who are not hit by an exogenous shock in Step 3 remain active and flow output is $\mu + \sigma_{Xi} Z_1$ where $Z_1 \sim \mathcal{N}(0,1)$.

5. The probability that the match is good p_{it} is updated using equation (4). Workers in matches where $p_{it} < \underline{p}_i$ become unemployed in the following period.