Racial Discrimination in the U.S. Labor Market: Employment and Wage Differentials by Skill

Job Market Paper

Daniel Borowczyk-Martins
University of Bristol

Jake Bradley
University of Bristol

Linas Tarasonis*
Aix-Marseille School of Economics

September 2013

Abstract

It is well-known that in the US labor market the average black worker is exposed to a lower employment rate and earns a lower wage compared to his white counterpart. Less attention has been given to the profile of these differences along workers’ skill distribution. Lang and Lehmann (2012) argue that black-white wage and employment gaps are smaller for high-skill workers. In this paper we show that a model of employer taste-based discrimination in a labor market characterized by search frictions and skill complementarities in production can replicate these regularities. We build on Shimer and Smith (2000) and assume that a positive share of employers are prejudiced against workers of a certain race. The model generates sorting along two dimensions: race/prejudice and skill. We estimate the model with US data using indirect inference estimation methods. Our quantitative results portray the degree of employer prejudice in the US labor market as being strong and widespread. We also find differences in the ability distributions of black and white workers, but they are quantitatively less important to explain observed racial differences in labor market outcomes.

Keywords: employment and wage differentials, discrimination, sorting.

JEL codes: J31; J64; J71.

*The authors are indebted to Grégory Jolivet, Fabien Postel-Vinay and Hélène Turon for their encouragement and very insightful comments and suggestions. We also thank Pierre Cahuc, François Fontaine, Pieter Gautier, Jean-Marc Robin and various seminar audiences. Daniel Borowczyk-Martins acknowledges financial support from Fundação para a Ciência e a Tecnologia under grant SFRH/BD/38968/2007, co-funded by the European Social Fund. Part of the research for this paper was carried out while Linas Tarasonis was at CREST (INSEE) and University of Bristol.
1 Introduction

In their recent survey of the economic literature on racial discrimination Lang and Lehmann (2012) highlight negative black vs. white employment and wage differentials as the two main empirical regularities a model of discrimination should replicate. Critically, they stress that in the United States (US) these differentials vary substantially by skill. In particular, wage gaps ‘are smaller or nonexistent for very high-skill workers’ and the employment gaps are ‘somewhat smaller among high-skill than among low-skill workers’ (p.12). They also emphasize that ‘no existing [discrimination] model can fully explain these regularities’ (idem).¹

In this paper we develop a model of discrimination that successfully replicates these regularities. The first contribution of the paper is to propose a mechanism that generates the observed skill profiles of wage and employment race differentials. Existing models of taste-based discrimination with search frictions explain black-white mean wage differentials (see Lang and Lehmann (2012)) and, to a less extent, can also explain black-white mean employment differentials. However, they cannot explain simultaneously the existence of employment and wage race differentials that differ for low- and high-skill workers.

The model developed in this paper shares several features with other taste-based search-discrimination models, like Black (1995), Bowlus and Eckstein (2002), Rosén (2003) and Flabbi (2010). The labor market is characterized by random search frictions and populated by two types of workers (who differ by race or gender) and two types of employers, where one type (prejudiced) incurs a utility cost from hiring a worker of a specific type. We enrich this setup by adding a second dimension of heterogeneity (skill) on both sides of the market.² Both workers and jobs are heterogeneous in skill (workers in ability and jobs in technology), which combines in a complementary way to produce match output. Two-sided skill heterogeneity and the assumption of production complementarities, coupled with the assumption of a utility cost of prejudice that is independent of workers’ and firms’ skill, gives us the means to generate a smooth decreasing profile of black-white employment and wage gaps in worker ability.

We estimate the model using various sources of publicly available data for the US manufacturing sector and indirect inference estimation methods.³ A critical feature the model

¹In this literature wage differentials are defined as one minus the ratio of mean black to white wages, whereas employment differentials refer to the percentage point difference between mean white and black employment rates.

²Bowlus and Eckstein (2002) allow the two types of worker to draw their productivity from two separate (degenerate) distributions, whereas Rosén (2003) and Flabbi (2010) model the productivity of the match (and not of workers and jobs) as being heterogeneous.

³Two-sided skill heterogeneity is a distinctive feature of the model which calls for an estimation strategy that makes use of matched employer-employee data. Unfortunately, for the US labor market no comprehen-
must satisfy to make its empirical implementation plausible is to allow for the possibility that black and white workers have different skill distributions. Indeed, there is substantive evidence of persistent black-white gaps in educational attainment and cognitive skill (see Neal (2006)), which suggests differences in skill endowments across races are likely to play an important role in shaping mean employment and wage differentials. Previous taste-based search-discrimination models have been estimated using publicly available data and structural methods (see Bowlus and Eckstein (2002) and Flabbi (2010)). Our paper is the first to take to the data a search-discrimination model based on Shimer’s and Smith’s (2000) partnership model. Structural estimations of this vintage of models are very recent in the applied search literature (see Jacquemet and Robin (2012) and Lise et al. (2013)).

The second contribution of the paper is to offer a precise quantitative account of the sources underlying the empirical regularities it seeks to replicate. Our structural decomposition allows us to quantify the contribution of employer prejudice and differences in ability across races to the observed black-white differences in wages and employment. This is an important undertaking as the literature identifies discrimination and skill differences as the two main competing explanations for the existence of differences in labor market outcomes across races.

A brief preview of the main empirical results in the paper is as follows. We find that, to describe the differences in labor market outcomes of black and white males in the US, both ability differences across races and employer prejudice are required, but discrimination generated by prejudiced employers is quantitatively more important. We estimate that about half the employers (49%) are prejudiced against black workers and that the utility cost of employing a black worker is about 8.8% of the average productivity of a match involving white workers. These results portray the degree of employer prejudice in the US labor market as being strong and widespread. We also find differences in the skill distributions of black and white individuals. The mean ability of blacks workers is estimated to be 4.7% lower than whites’.

The richness of our setting provides our model with a number of interesting theoretical results and implications, some of which are novel to the literature. First, the presence of prejudiced employers generates both wage discrimination and hiring discrimination against black workers. We follow the literature (see Cain (1986)) and define economic discrimination as the unequal treatment of equals on account of nonproductive factors, where equal workers are those with the same skill level. When a black worker is paid a wage lower than an equally able white worker we say there is wage discrimination. When an employer accepts to form a match with a white worker but refuses to match with an equally able black worker, we say
there is hiring discrimination. Second, our model incorporates sorting along two dimensions: race/prejudice and skill. Previous studies in the literature incorporated either of these two dimensions separately. The taste-based discrimination literature initiated in Becker (1971) embodies sorting on race/prejudice, while some contributions to the equilibrium job search literature focus on the empirical possibility of sorting on skill (see Abowd et al. (1999), Eeckhout and Kircher (2011), Hagedorn et al. (2012) and Lopes De Melo (2013)), or estimate the degree of skill complementarities in production using structural models (see Bagger and Lentz (2012) and Lise et al. (2013)).

The paper is structured in the following way. In section 2 we present the theoretical model, derive its equilibrium and establish some theoretical results pertaining to labor market outcomes of black and white workers. Section 3 describes the data. The estimation procedure is described in detail in section 4. Section 5 presents the fit of the model, while section 6 describes the structural parameter estimates. In the remaining sections we explore several applications using the estimated model. Section 7 quantifies the relative importance of prejudice and ability differences in explaining differences in employment and wages of blacks and whites. Section 8 analyzes the equilibrium sorting patterns of our model economy. Section 9 concludes.

2 The Model

The model we develop in this section builds on Shimer’s and Smith’s (2000) partnership model, extending it to a labor market where some employers are prejudiced vis-a-vis a specific type of workers and in which there is free entry of jobs. Because the model applies to any market where some employers are prejudiced against a certain type of worker, we will adopt a more general terminology in this section and return to the racial discrimination application we have in mind in the estimation section. Section 2.6 contains the main results of interest. It describes the properties of the equilibrium in the extended model — what we call a dual sorting equilibrium. The sections that precede section 2.6 set out the model in detail. Our exposition follows the structure in Shimer and Smith (2000) closely, but while their text emphasizes technical detail, ours discusses the pertinence of the model’s assumptions to study discrimination in the labor market. We refer the reader to Shimer’s and Smith’s original text and to the appendix at the end of the paper for technical details.

2.1 The Environment

We consider a labor market with $L$ workers and $G$ jobs. The number of jobs $G$ will be determined in equilibrium. There exist two types of jobs and two types of workers. A share
$m$ of workers are of type-1 and a share $1 - m$ of them are of type-2, with worker types being denoted by index $i = 1, 2$. Similarly, a share $\pi$ of jobs are operated by prejudiced employers ($P$) and a share $(1 - \pi)$ of those are not ($N$), with index $j = P, N$ denoting respectively prejudiced and nonprejudiced employers. In this model one firm is one job. Thus, throughout the text we use the terms jobs and firms interchangeably. Workers further differ in their ability $h$, which we assume is uniformly distributed over the unit interval, $h \in [0, 1]$, and firms in their level of technology (efficiency of labor inputs) $x$, which we also assume is uniformly distributed in the unit interval, $x \in [0, 1]$.¹ Let $\ell_i(h)$ and $g^j(x)$ denote respectively the population measures of type-$i$ workers of ability $h$ and type-$j$ firms of technology $x$.² The (endogenous) measures of type-$i$ unemployed workers of ability $h$ and type-$j$ vacant firms of technology $x$ are respectively denoted $u_i(h)$ and $v^j(x)$, with total measures of type-$i$ unemployed workers and type-$j$ vacant jobs given respectively by $u_i = \int u_i(h) \, dh$ and $v^j = \int v^j(x) \, dx$.

Time is continuous and both workers and firms are risk neutral, with discount rate $\rho$. Firms and workers maximize the present discounted value of future utility streams, measured in monetary terms. As in Becker (1971), prejudiced firms incur a psychic cost $d$ of employing a type-2 worker. When a worker and firm meet, their flow output depends on their levels of skill (worker ability and firm technology), denoted $f(h, x)$ and satisfying certain regularity conditions. We take complementarities in skill as a descriptive feature of modern labor markets and so assume a supermodular production function.³ This means the own marginal product of any worker and firm is increasing in his partner’s skill. Formal details are provided in appendix A.⁴

We assume that only unemployed workers and vacant firms search for a partner, ruling out on-the-job search. All unemployed workers search for jobs with equal search intensity. Job offers and unemployed job applicants arrive respectively to unemployed workers and vacant firms following a Poisson process. At each point in time the job and unemployed arrival rates are a function of the number of searchers on each side of the market via the aggregate matching function $M(u_1 + u_2, v^P + v^N)$. In the meeting process, type-1 and type-2 workers, and type-$N$ and type-$P$ firms, are perfect substitutes. Meeting is

---

¹These assumptions about the supports and densities of the skill distributions of workers and firms are normalizations. In the empirical application we allow all skill distributions to be defined over distinct supports and their densities to be non-uniform. Since we assume the production function is increasing in worker’s and firms’ skill, one can think of $h$ and $x$ as the skill ranks of the underlying skill distributions.

²So that $m \ell = \int \ell_i(h) \, dh$ and $\pi G = \int g^P(x) \, dx$.

³There is a growing consensus in the applied search literature on the descriptive relevance of positive assortative matching (see Lise et al. (2013), Lopes De Melo (2013) and Bagger and Lentz (2012)).

⁴In our empirical application we assume a specific degree of skill complementarities. However, the model, as well as the equilibrium characterization described in section 2.6, hold with any production function that satisfies the regularity conditions and supermodularity.
random and vacant jobs and unemployed workers of different types and skills are effectively exposed to the same arrival rates, respectively $\lambda^W = \frac{M(u_1 + u_2, v^F + v^N)}{u_1 + u_2}$ for workers, and $\lambda^F = \frac{M(u_1 + u_2, v^F + v^N)}{v^F + v^N}$ for jobs.

Once a firm and worker meet they decide whether or not to form a match. We define a match indicator function $\alpha^j_i(h, x)$, which is equal to 1 if a type-$i$ worker of ability $h$ and a type-$j$ firm of technology $x$ decide to match upon meeting. Matches are randomly destroyed by a Poisson process with arrival rate $\delta$, in which case both the worker and the firm reenter the pool of searchers. We should note that in the US labor market this assumption is at odds with the data — in particular, black workers experience, on average, lower employment spells. As argued in Lang and Lehmann (2012), differences in job destruction rates are likely to be quantitatively important to explain differences in unemployment rates across blacks and whites. Though appealing in practice, assuming exogenous differences in job destruction rates across races affects sorting patterns, via changes in the relative value of matches across types, and therefore job finding rates. Since we do not model this source of heterogeneity as a result of discrimination, we prefer to ignore it and quantify the effects of the sources of discrimination present in the model.

### 2.2 Value Functions

Workers can be in one of two different states: employed or unemployed. The flow value of employment for a type-$i$ worker of ability $h$ employed in a type-$j$ firm of technology $x$ is given by equation 1, where $w^j_i(h, x)$ is the wage she earns in this job and $U_i(h)$ is the value of being unemployed.

$$ \rho W^j_i(h, x) = w^j_i(h, x) + \delta \left[ U_i(h) - W^j_i(h, x) \right]. $$

(1)

While unemployed a worker receives a flow utility $b$. Therefore, the value of being unemployed for a worker of type-$i$ and ability $h$, $U_i(h)$, independent of the worker’s employment history, is given by the following equation.

$$ \rho U_i(h) = b + \lambda^W \sum_{j=P,N} \int \alpha^j_i(h, x) \left[ W^j_i(h, x) - U_i(h) \right] \frac{\nu^j(x)}{v^F + v^N} \, dx. $$

(2)

Note that, since $\nu^j(x)$ are density measures, $\frac{\nu^j(x)}{v^F + v^N}$ is the probability of sampling a vacant type-$j$ job of technology $x$ from the pool of unmatched firms.

Firms can be in three states: vacant, filled or idle. $J^j_i(h, x)$ is the value of a filled vacancy of a type-$j$ firm of technology $x$, filled with a type-$i$ worker of ability $h$:

$$ \rho J^j_i(h, x) = f(h, x) - d_{[i,j)=(2,P]} - w^j_i(h, x) + \delta \left[ V^j_i(x) - J^j_i(h, x) \right]. $$

(3)
where the indicator function, $1_{[(i,j)=(2,P)]}$ takes value 1 if the match involves a prejudiced firm and a type-2 worker.

In our model $d$ is a psychic cost valued in monetary terms that a prejudiced employer incurs upon matching with a type-2 worker. This psychic cost does not affect the production value of the match but it does affect the utility value of the match for a prejudiced employer. Note that in our specification $d$ enters additively in the firms’ value equation.\(^8\) This implies the degree of prejudice is independent of workers’ and firm’s skill levels. In other words, sorting on prejudice/race does not interact with sorting on skill. This assumption is crucial to replicate the skill profile of wage and employment differentials.

We assume that posting a vacancy has a nonnegative flow cost $\kappa \geq 0$. The value of posting a vacancy for a type-$j$ firm of technology $x$, $V^j(x)$, depends on the probability of the vacancy being filled by each of the two types of worker and it is given by the following equation.

$$\rho V^j(x) = -\kappa + \lambda \sum_{i=1,2} \int \frac{u_i(h)}{u_1 + u_2} dh.$$  \hspace{1cm} (4)

Again, since $u_i(h)$ is a density measure, $\frac{u_i(h)}{u_1 + u_2}$ is the probability of sampling an unemployed type-$i$ worker of ability $h$ from the pool of unmatched workers.

### 2.3 Entry

We assume that jobs remain active in the market if the present discounted value of keeping a job unfilled is nonnegative, i.e. if $V^j(x) \geq 0$. To determine the total mass of active jobs in equilibrium, $G$, we assume free entry. We show in Proposition 2 in appendix A that $V^j(x)$ is strictly increasing in $x$. We can therefore make a normalization and assume the least efficient job operated by a nonprejudiced employer makes zero profit. Free entry of jobs implies the following conditions respectively for jobs operated by nonprejudiced and prejudiced employers:

$$V^N(0) = 0,$$  \hspace{1cm} (5)

$$V^P(x^{P*}) = 0,$$  \hspace{1cm} (6)

where $x^{P*}$ is the technology level of the least efficient job operated by a prejudiced employer. We show in Corollary 4 that, if there is employer prejudice in this economy, i.e. if $\pi \in (0,1)$ and $d > 0$, then $x^{P*} > 0$. The least efficient job operated by a prejudiced employer in

\(^8\)This assumption is standard in search models of taste-based discrimination with bargaining and match-specific heterogeneity (see Rosén (2003) and Flabbi (2010)), as well as in search models with wage posting (see Bowls and Eckstein (2002)).
the market has a higher technology level compared to the least efficient job operated by a nonprejudiced employer.⁹

Prior to entry the characteristics of the employer who will operate the job and the technology level of the job are unknown. With probability \( \pi \) the employer is prejudiced. The technology level of new jobs is given by the probability distributions, \( g^j(x) \). As the mass of jobs increases, the expected value of keeping the job unfilled decreases for all active jobs. This process stops when the expected value of keeping a job unfilled is zero for the least efficient jobs operated by prejudiced and nonprejudiced employers. One possible interpretation of this model of entry in the context of labor market discrimination is the following. At any point in time an economy generates a certain number of jobs. These jobs differ in their efficiency of labor usage and on the racial attitude of the individual who is responsible for hiring a worker to operate it and who captures the surplus generated by the job. Because prejudice is costly, in equilibrium, the least efficient jobs operated by prejudiced employers have to be more efficient compared to the least efficient jobs operated by nonprejudiced employers. A prejudiced employer who draws a job with technology level lower than \( x^{P*} \) immediately leaves the market.¹⁰

This model of entry allows jobs operated by prejudiced employers to remain in the market, a question first addressed by Arrow (1973). Note that this model of entry rules out direct competition between prejudiced and nonprejudiced employers at every level of technology, as well as takeovers. In other words, we do not allow that, for every active job operated by a prejudiced employer, whatever her level of technology, there always exists an equally efficient nonprejudiced inactive employer who is willing to pay the prejudiced employer the latter’s expected value of being on the market, since it earns a positive profit from doing so. Allowing for this possibility would indeed imply that no jobs would be operated by prejudiced employers in the steady-state equilibrium. Whether the latter assumption provides a more accurate description is an open question.¹¹

2.4 Match Surplus

From the four value functions written above, we can determine the total surplus generated by any match. The surplus of a match between a type-\( i \) worker of ability \( h \) employed in

⁹The normalization we make here implicitly assumes \( d \geq 0 \). We only focus on equilibria satisfying this condition.

¹⁰See Larsen and Waisman (2012) for an alternative interpretation of this modeling approach in the context of immigrant discrimination.

¹¹This same remark is made, somewhat more concisely, in the closing section of Heckman (1998). Our model of entry resembles others descriptions in the literature. Black (1995) makes a similar argument to ours, but in his model prejudiced firms require a higher draw of entrepreneurial ability to enter the market, where the latter is independent of output. In the model developed by Rosén (2003), survival of prejudiced firms is achieved by a separation between owners and managers, where the owners are the residual claimants on output and managers bear the cost of prejudice.
a type-\textit{j} firm of technology \textit{x} is \( S_j^i(h, x) = W_j^i(h, x) - U_i(h) + J_j^i(h, x) - V^j(x) \) and it is split between the worker and firm in fixed shares. The worker takes a share \( \beta \) and the firm a share \( (1 - \beta) \), implying the following equalities:

\[
S_j^i(h, x) = \frac{J_j^i(h, x) - V^j(x)}{1 - \beta} = \frac{W_j^i(h, x) - U_i(h)}{\beta}
\]  

(7)

The wage equation that solves this bargaining problem is given by the following expression.

\[
w_j^i(h, x) = \beta \left[ f(h, x) - d^1_{[(i,j)=(2,P)]} - \rho V^j(x) \right] + (1 - \beta) \rho U_i(h)
\]  

(8)

Using equations 1 and 3, the expression for the total surplus can be rearranged and expressed by:

\[
S_j^i(h, x) = \frac{f(h, x) - d^1_{[(i,j)=(2,P)]} - \rho U_i(h) - \rho V^j(x)}{\rho + \delta}.
\]  

(9)

Whenever the surplus is positive a match is formed. Formally, we denote this by an indicator function:

\[
\alpha_j^i(h, x) = 1 \left[ f(h, x) - d^1_{[(i,j)=(2,P)]} - \rho U_i(h) - \rho V^j(x) > 0 \right].
\]  

(10)

A first remark about the bargaining process is that we assume firms and workers of different types have the same rent-sharing parameter. Allowing for different rent-sharing parameters across worker types would introduce another degree of heterogeneity across types. We choose not follow this route. Instead, we take the view of the strategic bargaining literature that rent-sharing parameters measure the relative impatience of bargaining participants and see no reason why it should differ across worker types and/or skill levels.\footnote{Some papers interpret differences in rent-sharing papers as the result of discrimination (see Eckstein and Wolpin (1999) and Bartolucci (Forthcoming)).}

Moreover, in the context of our model, absent taste-based discrimination (if \( d \) were zero) a lower \( \beta \) for type-2 workers by itself would not generate employment differences.

A second remark concerns the fact that \( d \) is transferable among match partners. This means the psychic cost of prejudiced firms is observable by both parties in the match and shared among them according to their rent-sharing parameters. We acknowledge that under a literal interpretation of the bargaining process, this feature can seem controversial. An alternative would be to model \( d \) as being nontransferable, which would be consistent with idea that the psychic cost of prejudice cannot be contracted among the parties. In practice, this would imply removing \( d \) from prejudiced firms’ value functions and altering the match feasibility condition between a type-2 worker and a type-P firm from equation 10 to \( \alpha_2^P(h, x) = 1 \left[ (1 - \beta)(f(h, x) - \rho U_2(h) - \rho V^P(x)) > d \right] \). In this specification a prejudiced firm only matches with a type-2 worker when its flow surplus is greater than the psychic cost \( d \). Though these two specifications tell a somewhat different story about how prejudice
translates into discrimination, they have similar properties. In particular, both generate hiring and wage discrimination in the presence of employer prejudice.

Having established the structure of agents’ payoffs, we can now define each agent’s strategy. For a type-\(i\) worker of ability \(h\) her strategy is given by two sets, \(M_i^P(h)\) and \(M_i^N(h)\). Similarly, a firm’s strategy is defined by two sets, \(M_j^i(x)\) and \(M_j^2(x)\). An agent’s matching sets contains all the acceptable partners with whom she is willing to match and who are willing to match with her. The symmetry of matching sets is due to the surplus-sharing rule being jointly privately efficient (i.e. the decision to match is mutually agreeable). Using the indicator function \(\alpha^j_i(h, x)\) we can express each worker’s matching set as:

\[
M_i^j(h) = \{x | \alpha^j_i(h, x) = 1\} \quad (11)
\]

and each firm’s matching set as:

\[
M_j^i(x) = \{h | \alpha^j_i(h, x) = 1\} \quad (12)
\]

2.5 Steady-state Equilibrium

A steady-state equilibrium of this model is characterized by four conditions: (i) workers and firms maximize their expected payoff, taking the strategies of all other agents as given; (ii) agents decide to match if it increases their payoff; (iii) all population measures of type-\(i\) workers of ability \(h\) and type-\(j\) firms of technology \(x\), \(\ell_i(h)\) and \(g^j(x)\), are in steady-state and (iv) the least productive active firms make zero profit. Conditions (i) and (ii) are given respectively by firms’ and workers’ value functions and their matching sets. Condition (iv) is given by the entry conditions. We now state the assumptions necessary to ensure condition (iii).

To fix all population measures, flow creation and flow destruction of matches for every type of agent must exactly balance. This is given by the following set of equations:

\[
\lambda^W \alpha^j_i(h, x) u_i(h) \frac{v^j(x)}{v^P + v^N} = \delta \gamma^j_i(h, x) , \forall x : V^j(x) \geq 0, \quad (13)
\]

where \(\gamma^j_i(h, x)\) is a joint measure of matched type-\(i\) workers of ability \(h\) and type-\(j\) firms of technology \(x\). These equations ensure that, for every possible match between a worker and firm of different types and skill levels, the number of matches being created at every point in time (the left-hand side of equation 13) is exactly the same as the number of matches being destroyed (the right-hand side of equation 13). Then, by definition, the steady-state stock of type-\(i\) employed workers of ability \(h\) is given by the following equation:

\[
\ell_i(h) - u_i(h) = \int \gamma^N_i(h, x) \, dx + \int \gamma^P_i(h, x) \, dx. \quad (14)
\]
That is, the total population of type-$i$ workers of ability $h$ must equal the sum of its unemployed and employed populations. Similarly, we can define the population of active type-$j$ firms of technology $x$ by:

$$g^j(x) - v^j(x) = \int \gamma^j_1(h, x) \, dh + \int \gamma^j_2(h, x) \, dh, \, \forall \, x : V^j(x) \geq 0,$$

so that, for each type-$j$ firm of technology $x$, the total number of firms $g^j(x)$ is equal to the total number of matched and vacant firms.

To obtain the equilibrium conditions of the model we first use the bargaining solution (equation 7) and firms’ and workers’ value functions to write each agent’s equilibrium reservation value (equations 16 and 17). We then use the flow-balance equations (equation 13) and the population accounting equations (equations 14 and 15) to express the equilibrium measures of unmatched agents (equations 18 and 19). Finally, the number of firms and the truncation point of prejudiced firms’ technology distribution is given the two entry conditions (equations 5 and 6). Formally, an equilibrium is defined in the following way.

**Definition 1 (Equilibrium):** Given exogenous parameters $L, m, d, \pi, \rho, \beta, b, \kappa, \delta$, the production function $f(h, x)$, a matching function $M(u_1 + u_2, v^P + v^N)$ and measures of firms and workers $\ell_i(h), g^j(x)$, an equilibrium is a fixed point $(\alpha^j_i(h, x), u_i(h), v^j(x), U_i(h), V^j(x), G, x^P)$

that solves the system of equations composed of the value functions of unmatched agents (equations 16 and 17), the entry conditions (equations 5 and 6), the measures of unmatched agents (equations 18 and 19) and the matching indicator functions of all the agents participating in the economy $\alpha^j_i(h, x)$ (equation 10):

$$\rho U_i(h) = \frac{b + \lambda^W \beta \sum_{j=\{N,P\}} \int \alpha^j_i(h, x) \left[ f(h, x) - d 1_{\{i,j\}=(2, P)} \right] - \rho V^j(x) \right] \frac{\nu^j(x)}{\nu^P + \nu^N} \, dx}{1 + \frac{\lambda^W \beta}{\rho + \delta} \sum_{j=\{N,P\}} \int \alpha^j_i(h, x) \frac{\nu^j(x)}{\nu^P + \nu^N} \, dx},$$

$$\rho V^j(x) = \frac{-\kappa + \lambda^P(1-\beta) \sum_{i=\{1,2\}} \int \alpha^j_i(h, x) \left[ f(h, x) - d 1_{\{i,j\}=(2, P)} \right] - \rho U^i(h) \right] \frac{u_i(h)}{u_1 + u_2} \, dh}{1 + \frac{\lambda^P(1-\beta)}{\rho + \delta} \sum_{i=\{1,2\}} \int \alpha^j_i(h, x) \frac{u_i(h)}{u_1 + u_2} \, dh},$$

$$u_i(h) = \frac{l_i(h)}{1 + \frac{\lambda^W}{\delta} \sum_{j=\{N,P\}} \int \alpha^j_i(h, x) \frac{\nu^j(x)}{\nu^P + \nu^N} \, dx},$$

and

$$v^j(x) = \frac{g^j(x)}{1 + \frac{\lambda^W}{\delta} \sum_{i=\{1,2\}} \int \alpha^j_i(h, x) \frac{u_i(h)}{u_1 + u_2} \, dh}, \forall \, x : V^j(x) \geq 0.$$
2.6 Dual Sorting Equilibrium

We now explore some implications of equilibrium for workers and jobs of different types and skill levels. Proofs of stated results can be found in appendix A. We characterize equilibria where, for all possible combinations of workers and firms of different types, some but not all matches are feasible. In other words, the equilibria we are interested in imply that, for any combination between a type-\(i\) worker and a type-\(j\) firm, some matches between workers of ability \(h\) and firms of technology \(x\) are formed, but not for all possible combinations of abilities and technology levels. In practice this implies that, on the one hand, the flow value of unemployment \(b\), the psychic cost borne by prejudiced employers \(d\) and firms’ vacancy cost \(\kappa\) are sufficiently small with respect to the value of production \(f(h, x)\), so that, for all combinations of workers and firms of different types, there exist combinations of skill levels \((h', x')\) that satisfy the match feasibility condition. On the other hand, it implies that, for certain combinations \((h, x)\), the value of production \(f(h, x)\) is small enough with respect to \(b\), \(\kappa\) and \(d\) to render the match between them not feasible, where this holds for any combination (type-\(i\), type-\(j\))\(^{13}\). In the various simulations of the model carried out in the execution of the paper we always found equilibria satisfying this description.

An equilibrium in our model economy is characterized by two forms of sorting across workers and jobs of different types an skill levels. To obtain positive assortative matching in skill we assume the production function is supermodular. We do not prove formally that, in our environment, assuming a supermodular production function implies positive assortative matching. Shimer and Smith (2000) show that when the production function is log-supermodular and satisfies the regularity conditions, there is positive assortative matching in skill. Proving this result in our environment, where not only are there complementarities in skill but also in race/prejudice, is not a trivial task. However, for all the simulations of the model we performed using different ranges of parameter values, we always observe positive assortative matching in skill. Regarding the second form of sorting – the patterns of negative assortative matching between black/white workers and prejudiced/nonprejudiced employers --, we are able to establish some results.

Before proceeding with the statements of the model’s implications we establish some necessary definitions. In the economics literature discrimination is said to exist when equally productive workers are treated differently based on nonproductivity related factors, such as race or gender (see Cain (1986))\(^{14}\). In our model, the psychic cost \(d\) reduces the utility value

\(^{13}\)These two conditions can be stated formally in the following way: \(\forall(i, j) \in [1, 2] \times [N, P], \exists(h', x')\) and \((h'', x'') \in [0, 1]^2\) such that \(\alpha_i(h', x') = 1\) and \(\alpha_j(h'', x'') = 0\).

\(^{14}\)As Cain (1986) emphasizes ‘although physical productivity excludes the psychic component [it] is intended to be broad and to include such characteristics of the workers as their regularity in attendance at work, dependability, cooperation, expected future productivity with the firm, and so on.’
of the match for prejudice employers, but, importantly, it does not affect the production value of the match, so the way we model discrimination is broadly consistent with the traditional definition studied in the literature. The first instance of economic discrimination we are interested in characterizing pertains to workers’ wages.

**Definition 2 (Wage Discrimination)**: A type-\(i\) worker of ability \(h\) experiences wage discrimination if she is paid a lower wage than an equally able type-\(k \neq i\) worker when both are matched with type-\(j\) firms with the same technology \(x\), that is

\[
\text{for some } (h, x), \ w_j^i(h, x) < w_{k \neq i}^j(h, x).
\]

The second instance of economic discrimination we characterize relates to agents decision of whom to match with. In an economy with no match surplus losses due to prejudice, type-1 and type-2 workers with the same ability match with firms within the same range of technology. In an economy in which there is employer prejudice, in general, this will no longer be the case and the matching sets of two equally able workers of different types will differ. One reason why these matching sets differ is due to hiring discrimination. Formally, we have that:

**Definition 3 (Hiring Discrimination)**: A type-\(i\) worker of ability \(h\) experiences hiring discrimination if, upon meeting a firm of technology \(x\) of type \(j = N, P\), he is not hired, but an equally able type-\(k \neq i\) worker is; that is,

\[
\text{for some } (h, x), \ \alpha_j^i(h, x) = 0 \text{ and } \alpha_{k \neq i}^j(h, x) = 1.
\]

To fix ideas, note that hiring discrimination describes discriminatory behavior by employers that is materialized in the decision to hire a worker — a decision that is different from that of how much to pay him (wage discrimination), but that stems from the same cause, viz. prejudice. The first implication of our model is that, for a positive value of \(d\), type-P firms (those who are prejudiced) and type-2 workers (those who are the object of prejudice) face worse perspectives in the labor market compared to type-N firms and type-1 workers, respectively. This result is stated in the following proposition.

**Proposition 1 (Outside Option Effects)**: For any equilibrium such that \(\pi \in (0, 1)\) and \(d > 0\):

(i) for a worker of ability \(h\), the value of unemployment of a type-1 worker is higher than that of a type-2 worker, that is, \(U_1(h) > U_2(h), \forall h\); and

(ii) for a firm of technology \(x\), the value of a vacancy to a type-N firm is higher than to a type-P firm, that is, \(V_N(x) > V_P(x), \forall x\).
A corollary of Proposition 1 is that, if there are any prejudiced employers in this model economy, there will be wage discrimination against all type-2 workers.

**Corollary 1 (Type-2 Wage Discrimination):** For any equilibrium such that \( \pi \in (0, 1) \) and \( d > 0 \), all type-2 workers experience wage discrimination in both types of firms and of any technological level.

So far we have established that, for prejudiced employers, the decision to match with a type-2 worker differs from that of matching with an equally able type-1 worker in two ways. First, the psychic cost \( d \) directly reduces the utility value of the match with a type-2 worker. Second, due to the outside option effect, type-2 workers have a lower outside option, which increases the value of the match. Since the decision to match is only governed by the match surplus condition (see equation 10), matching between type-2 workers and type-\( P \) firms depends on the relative magnitude of these two effects. We can show that, in our environment, the presence of employer prejudice implies that some type-2 workers will not be hired by certain prejudiced firms (the first effect dominates the second and it is high enough to reduce the match surplus to zero) and that those same firms will hire an equally able type-1 worker. As stated in Definition 3 this conjugation of circumstances entails hiring discrimination. The result is stated below.

**Corollary 2 (Type-2 Hiring Discrimination by Prejudiced Firms):** For any equilibrium such that \( \pi \in (0, 1) \) and \( d > 0 \), some type-2 workers experience hiring discrimination by some prejudiced firms.

For nonprejudiced firms, the difference between matching with equally able type-1 and type-2 workers is only affected by the outside option effect. When the match surplus condition between a nonprejudiced firm and a type-1 worker is not satisfied, the lower outside option of an equally able type-2 worker may render that match feasible. In those circumstances type-1 workers suffer hiring discrimination by those nonprejudiced firms — we will refer to this conjugation of circumstances as reverse hiring discrimination. This result is stated below.

**Corollary 3 (Type-1 Hiring Discrimination by Nonprejudiced Firms):** For any equilibrium such that \( \pi \in (0, 1) \) and \( d > 0 \), some type-1 workers experience hiring discrimination by some nonprejudiced firms.

We now turn our attention to the distributions of technology among prejudiced and nonprejudiced firms in equilibrium. If there are any type-2 workers in the economy, prejudiced
firms have to be more efficient than equally efficient nonprejudiced firms, since they have to make up for the cost of prejudice $d$ — which affects negatively their matching opportunities and the match value of feasible matches with type-2 workers. This effect is stated in the following corollary.

**Corollary 4 (Threshold Technology Differences):** For any equilibrium such that $\pi \in (0, 1)$ and $d > 0$, $\exists x^{P^*} > 0$ such that $V^P(x^{P^*}) = 0$.

Corollary 4 implies that, conditional on the technology distributions $g(x)^j$ across firm types being the same, in equilibrium, prejudiced firms will be on average unambiguously more efficient than nonprejudiced firms.

## 3 Data

To estimate the model we use three sources of data: worker level data, firm level data and market level data. The worker side data comes from the Current Population Survey (CPS). We merge the Monthly Outgoing Rotation Groups (MORG) with the Basic Monthly (BM) extracts, thus gathering information on individual wages and transition rates across employment and unemployment. Our sample runs from May 2004 to December 2005. We limit the sample to include only individuals who declare themselves to be either black or white. We only keep males in the sample in order to avoid complications of modeling labor supply decisions and to be as precise as possible about the type of prejudice we are estimating. We also restrict our sample to individuals between the ages of 18 and 65 who remain active in the labor market throughout their spell in the sample. We only consider individuals in two labor market states: unemployed or employed in a full-time job in the private sector.\(^{15}\) Finally, we restrict the sample to individuals who at some point during the sample were employed at a manufacturing firm.

Following these restrictions, we are left with a sample of 114,894 males (8,730 blacks vs. 106,270 whites), of which 4,459 are unemployed (735 blacks vs. 3,733 whites). When defining the sample we face a trade-off between sample homogeneity and sample size. Because unemployed blacks represent a very small share of the population, in order to have a representative sample of these individuals (one that provides accurate estimates of moments related to job mobility) we need a large sample. Since the steady-state assumptions limit the sample size in the longitudinal dimension, we have to sacrifice the homogeneity of the sample. This is the main reason why we include individuals of all working ages (18-65 years-old) and education levels. Note that this does not conflict with our interpretation of

\(^{15}\)In addition, we excluded individuals with reported working hours outside the 35 - 70 hours interval.
the model. Age and education are strong predictors of individual ability, but the way ability is defined in the model captures all dimensions of ability, be them observed or unobserved to the econometrician.

Worker’s ages are weekly and measured at the time of the interview. Wages are observed at most twice per individual and are top-coded.\textsuperscript{16} To deal with these issues we trim the top and bottom 2% of the wage distributions for black and white males, replacing observed wages as missing. Using information on the number of working hours we convert wages to be hourly, which should further reduce the problem of top-coding.

Our source of firm level data is the NBER-CES Manufacturing Industry Database.\textsuperscript{17} To our knowledge, this is the only publicly available data set for the US with information on value-added that can be readily matched with CPS data. We match firm level data to worker level data by NAICS four-digit industry code. Our matched sample contains 76 distinct manufacturing industries from the year 2005. For each four-digit industry we compute the average value-added per worker per hour and interpret it as the level of production generated by the match between that worker-firm pair (i.e. $f(h, x)$). This means in our empirical application technology differences across firms are only explained by the four-digit industry at which they operate. We are aware that this estimation strategy implies a particularly restrictive interpretation of the model.

Finally, our sources of market level data are the Job Openings and Labor Turnover Survey (JOLTS), the CPS and the Current Employment Statistics (CES), available from the Bureau of Labor Statistics webpage. To estimate labor market tightness in the same period in the manufacturing sector, we use the unadjusted series of job openings in the manufacturing sector from the JOLTS (JTU30000000JOL), the unadjusted unemployment rate series in the manufacturing sector from the CPS (LNU03032232) and the series of unadjusted employment in the manufacturing sector from the CES (CES3000000001). Our estimate of market tightness is 0.3808, with a standard deviation of 0.013.

4 Estimation

The relevant moments predicted by the model do not have a closed-form expression. Therefore, we make use of simulation methods to estimate the model’s parameters. Specifically, we estimate the model by indirect inference (see Gouriéroux et al. (1993)). This procedure

\textsuperscript{16}There is a literature about the misreporting of wages in the CPS. The main source of measurement error is believed to be over reporting at low levels and under reporting at high levels (see Bollinger (1998)).

\textsuperscript{17}This database is a jointly produced by the National Bureau of Economic Research (NBER) and U.S. Census Bureau’s Center for Economic Studies (CES). It contains annual industry-level data from 1958-2009 on output and employment and other variables, for the 473 six-digit 1997 NAICS industries (see http://www.nber.org/nberces).
involves a simulated method of moments estimator, in which some or all of the moments the procedure seeks to match are parameters from reduced-form models that capture important aspects of the ‘true’ data-generating process (i.e. the structural model). The parameters of these auxiliary models are a function of the structural parameters we seek to estimate.

The mechanics of an indirect inference procedure are the following. Let $\theta$ denote the vector of structural parameters (to be specified in the next subsection), $\hat{m}^S(\theta)$ denotes the model-generated vector of parameters of the auxiliary models and $\hat{m}^S$ its empirical counterpart. (The contents of these vectors are defined in section 4.2). The estimation procedure finds $\theta$ such that the distance between the model-generated moments and their empirical counterparts is as small as possible, according to the following criterion function:

$$L_N(\theta) = -\frac{1}{2} (\hat{m} - \hat{m}^S(\theta))^T \Omega^{-1} (\hat{m} - \hat{m}^S(\theta))$$  \hspace{1cm} (20)$$

To obtain the theoretical moments we need to simulate the model. A description of how we carry this out in practice is provided in appendix B. We set the weighting matrix $\Omega$ such that all moments have equal weight. Finally, to obtain standard errors we plan use the bootstrap. [S.E.s COMING SOON]

### 4.1 Econometric Specification

In order to make the model developed in section 2 empirically operational we make certain parametric assumptions and calibrate some parameters. The population shares of worker types, $m$ and $1-m$, are observed (92% and 8%, respectively). We specify the meeting function to be Cobb-Douglas with Constant Returns to Scale (CRS) and meeting elasticities equal to 0.5, i.e. $M(u_1 + u_2, v^P + v^N) = \lambda(u_1 + u_2)^{0.5}(v^P + v^N)^{0.5}$, where $\lambda$ is the constant matching efficiency parameter to be estimated. The monthly discount rate, $\rho$, is set at 0.0043 (equivalent to 5% per annum).

We assume a multiplicative production function in the skill levels of workers and firms. When describing the model in section 2 we assumed, for exposition purposes, that all skill distributions were the same, namely uniforms with support in the unit interval. We relax these assumptions in the empirical implementation of the model. We assume that all three distributions are log-normals, but allow the parameters of each distribution to differ. This allows us to summarize skill heterogeneity of this economy in six parameters, respectively the mean and standard deviation of white and black workers’ ability distribution, $\{\mu_1, \mu_2, \sigma_1, \sigma_2\}$, and the mean and standard deviation of firms’ technology, $\{\mu_x, \sigma_x\}$. To specify the log-normal distributions we transform the production function, replacing the rank of workers’ and firm’s skill by the inverse of the normal CDF, $\Phi^{-1}$, which we then transform to log-normal distributions with mean, $\mu$, and standard deviation, $\sigma$. 

16
\[ f_i(h, x) = \exp\{\mu_i + \sigma_i \Phi^{-1}(h)\} \exp\{\mu_x + \sigma_x \Phi^{-1}(x)\}, \quad i = 1, 2. \]

Together these assumptions leave us the following vector of parameters to estimate:

\[ \theta = \{\lambda, \delta, d, \pi, \mu_1, \sigma_1, \mu_2, \sigma_2, \mu_x, \sigma_x, \beta, b, \kappa\} \]

using data on individuals (indexed \( i \)) at different months \( t \), namely data on their race, \( \bar{r}_i \), employment status, \( \bar{s}_{i,t} \), log-wage, \( \bar{w}_{i,j,t} \), and log value-added per worker of the 4-digit manufacturing industry in which they are employed, \( \nu_{j(i),t} \), where firms are indexed \( j \); and an estimate of the average market tightness in the manufacturing sector during the same time period \( \bar{\theta} \).

### 4.2 Auxiliary Models

The choice of auxiliary models is a key step in an indirect inference procedure. Our choices are based upon what we deem to be the crucial aspects of our theoretical model (in the sense that they capture aspects of the model that are informative about the value of certain parameters) and for which we have reliable empirical counterparts. Due to the complexity of the model, we can only provide an heuristic argument to support parameter identification.

#### 4.2.1 Labor Market Transitions

Our model predicts three different rates of transition across labor market states.\(^{18}\) We summarize the information about transitions in the model in three moments: the average rate of transitions from job to unemployment for all employed workers and the average rates of transition from unemployment to job for white and black unemployed individuals. Their theoretical counterparts are given below.

\[ jtu = 1 - e^{-\delta \times 1} \] \hspace{1cm} (21)

\[ ut_{ji} = \int \left[ 1 - \exp \left( -\lambda^W \sum_{j=\{P,N\}} \int_{\alpha''(h,x) v''(x) dx \times 1} \frac{u_{ji}(h)}{u_i} dh \right) \right] \] \hspace{1cm} (22)

The average rate of transitions from job to unemployment is essentially governed by \( \delta \), the parameter of the continuous Poisson process that determines when a match is destroyed.

---

\(^{18}\)Under the steady state assumption these determine unemployment rates. Empirically, we can also observe different job destruction rates between blacks and whites as well as job-to-job transition rates. As we have argued before, introducing on-the-job search is a nontrivial extension we leave for future research. Although different job destruction rates for blacks and whites could be easily estimated, we choose not to make this extension. This imposes another source of discrimination which is not modeled and which also affects the complementarities between the race of a worker and a firm’s output.
Similarly, the average rates of transitions from unemployment to job for white and black workers are informative about the Poisson process governing the meeting of workers and jobs. This can be seen clearly in equation 22, where $$\lambda^W = \lambda(u_1 + u_2)^{0.5}(v^P + v^N)^{0.5}$$. This expression also shows that the difference in mean unemployment-to-job transition rates across the two races is informative about the degree of employer prejudice ($$\pi$$ and $$d$$). This is because the latter parameters affect differently the matching rates of workers of different races, via the matching sets, $$\alpha_j^i(h, x)$$ (see section 8 for an illustration of these effects).

### 4.2.2 Distribution of Wages and Value-added

The wages paid to workers of different skill and race are a key object of the model. As can be clearly seen from the wage equation (equation 8), the distribution of wages is informative about the distributions of workers’ ability, the psychic cost of prejudice $$d$$ and the technology distribution of firms. Wages also depend on reservation values and so they are also affected by the share of prejudiced firms $$\pi$$ in the economy and the flow value of unemployment, $$b$$. To disentangle the ability distribution of black and white individuals, as well as employer prejudice parameters, we use moments from the wage distributions of black and white workers. In practice, we use the mean, standard deviation, skewness, kurtosis and the minimum of each of the two distributions. In the spirit of the argument in Flinn and Heckman (1982), the latter moment should be informative about the level of $$b$$.

To disentangle the contribution of workers’ and firms’ skill to the level of wages we make use of firm level information. While not a direct counterpart to the distribution of firms’ technology, the distribution of value-added per worker is informative about the level and variance of firms’ technology.\(^{19}\) We use the first two moments of the distribution of value-added.

### 4.2.3 Rent Shares Regressions

The choice of our last auxiliary model draws on arguments made in the rent-sharing literature.\(^{20}\) Conditional on the skill of workers and firms, $$h$$ and $$x$$, the match production is divided up according to the worker and firm types and the bargaining parameter, $$\beta$$ (see equation 8). We estimate the following regression models separately for white and black employed individuals:

\(^{19}\)We interpret firms’ value-added per worker as the counterpart of the match production $$f(h, x)$$. We do so because, when there is sorting on skill, firm value-added can no longer be taken as a direct measure of firm technology.

\(^{20}\)Blanchflower et al. (1996) determine the share of profits accruing to workers in US manufacturing firms by matching CPS worker data with information on profits per sector (defined by 2-digit industry code), and regressing the former on the latter.
\[ w_{i,j,t} = \gamma_{j(i),t} + \epsilon_{i,j,t}, \]

(23)

where, \( \epsilon_{i,j,t} \) is an i.i.d. disturbance and \( w_{i,j,t}, \nu_{j(i),t} \) respectively the log-wage and log-value-added.

The parameter \( \bar{\gamma} \) describes the mean relationship between wages and firm value-added. These two parameters are related to the structural parameters \( \beta, \pi \) and \( d \) in the following way. The level of the parameter \( \bar{\gamma} \) is informative about rent-sharing in this economy, namely the value of \( \beta \). Any differences in \( \bar{\gamma} \) estimated separately in the population of black and white workers reveal information about the extent of discrimination and, therefore, of the magnitude of employer prejudice parameters (\( \pi \) and \( d \)). Indeed, if there were no prejudiced employers (\( d = 0 \)), then, on average, white and black workers would get an equal share of value-added and \( \bar{\gamma} \) estimated on the two populations of workers would be the same.

4.2.4 Labor Market Tightness

The ratio of vacant jobs to unemployed workers, \( \bar{\theta} = \frac{\bar{V}}{\bar{U}} \), summarizes the degree of tightness in a labor market. The model counterpart of this statistic can be computed by integration of the accounting equations 14 and 15 respectively over the support of jobs’ and workers’ skill, which gives \( \bar{V} \) and \( \bar{U} \). Given free entry of jobs and an aggregate meeting function, the number of firms in equilibrium, \( \bar{G} \), is fundamentally determined by the cost of posting a vacancy \( \bar{\kappa} \), via the two entry conditions (equations 5 and 6). These entry conditions also determine the threshold technology level of the least efficient job operated by a prejudiced employer, \( x_{P\ast} \). In sum, given values for the other parameters of the model, market tightness informs the value of \( \bar{\kappa} \) needed to satisfy the entry conditions.

5 Fit of the Model

In this section we evaluate the fit of the model.\(^\text{21}\) The first column of Table 1 reports the set of simulated moments corresponding to estimates of the model that allows for both ability differences and prejudice to govern workers’ labor market outcomes. Columns 2 and 3 respectively present the empirical moments used in the estimation and their respective standard deviations.

We first describe the empirical moments used to identify the parameters of the model. The top panel of Table 1 shows moments that govern worker mobility to and from unem-

\(^{21}\)In the latest version of the paper before the current one, along with the baseline specification of the model, we reported the fit, parameters estimates and standard errors of two alternative model specifications. The current estimation of the model is far more time-consuming, as to estimate \( \kappa, G \) and \( x_{P\ast} \) have to be determined endogenously. We are currently working on estimating these alternative specifications.
Table 1: Model Fit

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrestricted</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Transition Rates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment to Job</td>
<td>White</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>0.264</td>
</tr>
<tr>
<td>Job to Unemployment</td>
<td></td>
<td>0.009</td>
</tr>
<tr>
<td><strong>Wage Distribution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>White</td>
<td>2.872</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>2.631</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>White</td>
<td>0.471</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>0.445</td>
</tr>
<tr>
<td>Skewness</td>
<td>White</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>0.236</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>White</td>
<td>2.157</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>2.210</td>
</tr>
<tr>
<td>Minimum Wage</td>
<td>White</td>
<td>1.929</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>1.796</td>
</tr>
<tr>
<td><strong>Value Added Distribution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>4.234</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
<td>0.609</td>
</tr>
<tr>
<td><strong>Rent Share Regression</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>White</td>
<td>0.668</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>0.634</td>
</tr>
<tr>
<td><strong>Labor Market Tightness</strong></td>
<td></td>
<td>0.381</td>
</tr>
</tbody>
</table>

employment. The first two moments are the average transition rates from unemployment to job of respectively white and black males in the sample. The third moment is the average transition rate from job to unemployment of all individuals in the sample. The observed difference in unemployment-to-job transition rates of whites and blacks is of 5 percentage points. The size of this difference is consistent with an average unemployment duration of black unemployed workers that is 20 - 30% longer than that of whites, which is reported in the various studies cited in Lang and Lehmann (2012). The job-to-unemployment transition
rate common to all workers is 0.87%.\footnote{The difference in job-to-unemployment transition rates across races is 24\% (1.02\% and 0.82\% for black and white employed workers respectively). The magnitude of this difference is in line with studies reported in the literature (see Lang and Lehmann (2012)).}

The middle-top panel presents the first four moments of the distributions of log-wages of black and white workers. The differential in mean wages (in levels) of black and white workers is 31.4\%, consistent with numbers found in the literature (see Lang and Lehmann (2012)). The standard deviations of both wage distributions have similar magnitudes, with whites’ mean wage exhibiting slightly more variation. Both distributions are positively skewed, but the distribution of black wages is substantially more skewed to the right than that of whites’ (0.234 vs. 0.060, that is four times as larger). This difference can be seen more clearly in Figure 1, by comparing the two empirical distributions (given by dashed lines). The two distributions have very similar kurtoses and, as expected, the minimum wage of black workers is smaller than that of white workers (a 23.2\% differential of wages in levels).

The middle-bottom panel displays the mean and standard deviation of the distribution of log-value-added of manufacturing 4-digit sector in the sample. The last panel of Table 1 shows, for the samples of white and black workers, the rent-share regression coefficients ($\gamma$), which we also refer to as wage elasticity of firm value-added. With an elasticity of 0.671 against an elasticity of 0.596, white workers take, on average, a larger share of the match compared to black workers. A final comment concerns the precision of the moments. Due to the large size of our sample, all moments are very precisely estimated.

The model fits the empirical moments quite well. Comparing the values across columns in the first rows of the top panel of Table 1, it is clear the model matches very well the difference in mean unemployment-to-job rates of whites and blacks. Regarding the fit of moments of the wage distributions, reported in the middle-top panel of Table 1, the model also does a good job. The predicted mean of white log-wages is slightly above its empirical counterpart, which slightly overstates the observed difference in mean log-wages across races. A similar comment applies to the performance of the model at fitting the standard deviations of the two log-wage distributions. The fit of skewnesses is very good, which is important given the sizeable difference in the empirical moments. The models does not fit so well the fourth moment of the log-wage distributions. A visual display of the overall fit of the two log-wage distributions is given in Figure 1. The dashed lines denote the empirical distributions and the solid lines the distributions predicted by the model. The model successfully captures the difference in the two distributions, even though the fit of whites’ log-wage distribution is somewhat poor. To conclude the discussion of model fit, we turn our attention to the rent share regression coefficients, reported in the bottom panel of
Table 1. The model captures the level of the coefficients well, but the difference in wage elasticities of value-added is somewhat smaller than the one in the data (0.024 vs. 0.075).

6 Structural Parameter Estimates

Having argued that our model does a good job at matching the moments in the data, we now analyze its economic content. Table 2 reports the parameter estimates of our structural model.23

The estimated job arrival and job destruction of the continuous-time Poisson processes are respectively 0.823 and 0.009. The size of these parameters are chiefly determined by the average transition rates between employment and unemployment in our sample. These moments are smaller than the average values reported in macro studies for the US economy over the same time period (see Shimer (2012)). This is consistent with the evidence on the relatively lower turnover in the manufacturing sector in the US (see Davis et al. (2013)).

The model produces evidence of employer prejudice against black workers. The loss of utility incurred by a prejudiced firm when matching with a black worker is estimated at 7.16 dollars per hour. This corresponds to 8.8% of the average productivity of a match involving white workers. The share of prejudiced employers is estimated at 49%. In the context of our model, these values provide a picture of discrimination in US labor market as being driven by the presence of widespread and strong employer prejudice. This pattern

---

23 At this point we are still estimating standard errors using the bootstrap. See footnote 21.
Table 2: Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ability (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching efficiency</td>
<td>λ 0.823</td>
</tr>
<tr>
<td>Probability of job destruction</td>
<td>δ 0.009</td>
</tr>
<tr>
<td>Cost of discrimination</td>
<td>d 7.160</td>
</tr>
<tr>
<td>Fraction of prejudiced firms</td>
<td>π 0.490</td>
</tr>
<tr>
<td>Worker productivity parameters</td>
<td>μ₁ 1.911</td>
</tr>
<tr>
<td></td>
<td>σ₁ 0.178</td>
</tr>
<tr>
<td></td>
<td>μ₂ 1.874</td>
</tr>
<tr>
<td></td>
<td>σ₂ 0.156</td>
</tr>
<tr>
<td>Firm productivity parameters</td>
<td>μₓ 2.357</td>
</tr>
<tr>
<td></td>
<td>σₓ 0.564</td>
</tr>
<tr>
<td>Worker bargaining power</td>
<td>β 0.050</td>
</tr>
<tr>
<td>Unemployment flow utility</td>
<td>b 4.364</td>
</tr>
<tr>
<td>Vacancy flow cost</td>
<td>κ 540.381</td>
</tr>
<tr>
<td>Criterion</td>
<td>-0.013</td>
</tr>
</tbody>
</table>
is not entirely consistent with results from surveys on racial attitudes conducted in the US and documented in Lang and Lehmann (2012). Our model delivers a degree of employer prejudice that comes closer to the picture provided by these studies compared to other structural estimations based on the US labor market. We take this as an indication that further progress can be achieved by extending the model in other directions.

Turning to differences in ability across black and white individuals, we estimate a mean black-white ability differential of 4.7%. There are also differences, though marginal, in the standard deviations of the two distributions: whites ability is more dispersed compared to blacks (0.191 vs. 0.187). However, we cannot ascertain whether these differences are statistically significant until we obtain standard errors. There is obviously no available direct measure of workers ability to which we can compare our results to, however, our evidence seems consistent with the persistent skill-gap of black adults with respect to their white counterparts documented in great detail in Neal (2006) — if anything his conclusions suggest the gap might be even larger. The worker bargaining parameter is estimated at 0.05%, a value very close to zero. This estimate should be interpreted very cautiously, as the firm side data used in our empirical exercise is of limited quality.

Finally, the cost of posting a vacancy is estimated to be 540.381 $ per hour. The order of magnitude of this parameter is surprisingly high. The first thing to note is that the estimate of $\kappa$ is essentially driven by one moment (market tightness), which is quite low in the US manufacturing sector. When we estimate the model calibrating the cost of posting a vacancy based on estimates reported in the literature, the model can no longer fit the moments. This said, we concede that a low value of market tightness in the manufacturing sector is probably not related to very high vacancy posting costs. Our estimate of $\kappa$ might be picking up other aspects of the manufacturing sector, like high capital costs of entry, which are not included in the model.

7 Decomposition of Wage and Employment Differentials

So far we have established that both ability differences and employer prejudice are important to match the patterns in the data pertaining to differences in blacks’ and whites’ labor

---

24They conclude: ‘We take the evidence from the surveys and the IAT [Implicit Association Tests] as suggesting that credible models of discrimination based on prejudice may rely on the presence of strong prejudice among a relatively small portion of the population and/or weak prejudice among a significant fraction of the population, but not on widespread strong prejudice’, Lang and Lehmann (2012).

25Bowlus and Eckstein (2002) estimate that 56% of employers are prejudiced and that the disutility of hiring a black worker is 31% of white workers’ productivity.

26We discuss these directions in the conclusion.

27This ratio is calculated in the following way: $1 - \exp(\mu_2 - \mu_1 - (\sigma_2^2 - \sigma_1^2)/2)$.

28Using data on the unemployment rate from the CPS, employment from the CES and on vacancies from the JOLTS, the estimated market tightness is 0.54, for the nonfarm economy, and 0.38, for the manufacturing sector.
market outcomes. We now assess their relative importance. We focus on the first, second (median) and third quartiles of the log-wage distribution and the mean unemployment rate. Columns 1 and 2 of Table 3 show the values of these moments, respectively for white and black employed individuals, generated by the parameter estimates of the benchmark model (column 1 of Table 2). The next two columns of Table 3 show how much of the difference in moments can be explained by differences in ability and the existence of prejudiced employers, as a share of the same difference produced by both sources. Column 3 of Table 3 is obtained by simulating a model with the benchmark parameter estimates (column 1 of Table 2) except for the psychic cost, $d$, which is set to zero, to see how much of the gap can be attributed to ability differences alone. Analogously, column 4 of Table 3 is obtained by simulating a model with the benchmark parameter estimates and setting the ability distributions of black and white workers to the population ability distribution.

Table 3: Decomposition of wage and unemployment gaps

<table>
<thead>
<tr>
<th></th>
<th>Share of gap (%):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whites</td>
</tr>
<tr>
<td>Wage quartiles</td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>2.50</td>
</tr>
<tr>
<td>Median</td>
<td>2.87</td>
</tr>
<tr>
<td>Q3</td>
<td>3.24</td>
</tr>
<tr>
<td>Unemployment (%)</td>
<td>2.23</td>
</tr>
</tbody>
</table>

A number of remarks are in order. First, discrimination is quantitatively more important than ability differences for all the location measures of the wage distribution, with the exception of the third quartile. Second, prejudice is especially important to explain the differences in outcomes of black vs. white workers who earn lower wages (those who have lower levels of ability) and decreases monotonically as one moves towards higher wages. Again, this follows from the structure of the model, where the effects of prejudice on blacks’ outcomes diminish as their level of skill increases (we discuss this more thoroughly in the next section). Finally, employer prejudice explains a much larger share of the difference in average unemployment rates across races.

8 Sorting on Race/Prejudice and Skill

Having estimated the parameters of the model we can simulate it and describe its main features. To simulate the model we follow the methodology described in appendix B, where
the model is specified as detailed in section 4.1. Figure 2 displays the equilibrium strategies of all the agents in our economy (their matching sets) and the labor market outcomes of black and white workers of different levels of skill. Figures 2a and 2b plot the contour lines of the set of matching sets of respectively prejudiced and nonprejudiced employers with white and black workers. Workers’ ability levels are displayed on the horizontal axis and firms’ technology levels on the vertical axis. The solid lines indicate the bounds of the set of matching sets between white workers and prejudiced firms, while the dashed lines indicate the bounds of the set of matching sets between black workers and prejudiced firms. The region in the interior of each of the two pairs of lines contain all matches that produce a positive surplus.

A first observation is that the shapes of the two sets of matching sets embody sorting on skill: low(high)-technology firms are matching with low(high)-skill workers. Matches in the northwestern corner generate negative surpluses: high-technology firms have a high outside option that can only be compensated when they meet high-skill workers and production is high. Conversely, in the southeastern corner, the outside option of high-skill workers can only be outweighed when they match with high-technology firms.

A second observation concerns differences in whites’ and blacks’ sets of matching sets. Clearly, if there were no prejudice \((d = 0)\) the two sets of matching sets would coincide. But, as can be seen by inspecting Figure 2a, the set of matching sets of blacks with prejudiced firms is contained in the one of whites. It implies that there is a range of technology levels for which prejudiced firms are willing to match with white workers, but not with equally able black workers. This is an instance of hiring discrimination stated in Corollary 3. The main mechanism underlying the difference in the two sets of matching sets is the direct effect of prejudice. It decreases the value of matches between black workers and prejudiced firms, to the point that even if they are paid a lower wage compared to an equally able white worker (wage discrimination), some matches do not generate a positive surplus. This decrease in the set of matching sets is asymmetric with respect to blacks’ ability: it is stronger in the southwestern corner and milder in the northeastern corner. This follows from the way we modeled discrimination, where the psychic cost for a prejudiced employer upon matching with a black worker is independent of workers’ and firms’ skill levels. This implies low-skill blacks — who due to sorting match with low-technology firms — suffer relatively more from prejudice compared to their high-skill counterparts.

A different mechanism underlies reverse hiring discrimination (see Corollary 2). As can be seen in Figure 2b, the set of matching sets of black workers is larger than that of white workers in the northwestern and southeastern corners. There is an interval of technology levels for which nonprejudiced firms are willing to match with black workers
Figure 2: Simulation using estimated parameters
but not with equally able white workers. The presence of prejudice employers decreases blacks’ opportunities in the labor market, which results in them having a lower outside option compared to equally able whites (see Proposition 1). This lower outside option makes matches of nonprejudiced employers of certain technology levels with blacks feasible compared to unfeasible matches with equally able whites.

The differences in labor market outcomes across equally able black and white workers will depend on the relative magnitude of the various forms of discrimination present in the model, which in turn are determined by $\pi$ and $d$ and the share of whites and blacks in the population ($m$ and $1 - m$). In both Figures 2c and 2d the solid line indicates white workers’ outcomes and the dashed line the outcomes of black workers, whereas the $x$-axis denotes the ability percentiles of the population of all workers. Figure 2c shows the unemployment rates of both types of workers for different levels of skill. One can immediately see that blacks have higher unemployment rates compared to whites for any level of skill. This highlights what is already apparent in Figures 2a and 2b: that hiring discrimination dominates reverse hiring discrimination. Second, the black-white unemployment rate differential is far higher for low levels of skill and vanishes as the level of skill increases. This is another illustration of the disproportionate effect of the psychic cost $d$ on low-skill black workers. Third, for both types of workers the distributions of unemployment rates across levels of skill describe a U-shaped curve, with workers in the middle-range of the skill distribution experiencing lower unemployment rates. This results from the shape of matching sets and is a general feature of other search and matching models with sorting on skill (see Eeckhout and Kircher (2011)).

Finally, Figure 2d displays the differences in expected wages for blacks and whites with different levels of skill. The main patterns are similar to those in Figure 2c. Whatever the level of skill, the expected wage of whites is always higher than that of equally able blacks. The white-black expected wage differential is high for low levels of skill and disappears as the level of skill increases. Similar to the effects of hiring discrimination, wage discrimination affects low-skill blacks proportionately more. This is already clear in the wage equation of the model (see equation 8).29

9 Conclusion

This paper develops a search model of the labor market with two-dimensional heterogeneous firms and workers. Our taste-based model of discrimination successfully replicates stylized

29A last clarifying remark is perhaps in order. Due to the estimated differences in ability across races, there are no black individuals with extremely high levels of skill. This is why the dashed lines in all the four plots end before the solid lines.
facts regarding mean differences in wages and employment between white and black workers, as well as across their respective skill levels. The model is estimated and the relative contribution of discrimination (resulting from employer prejudice) and ability differences across races to explain observed differences in labor market outcomes is assessed.

Despite the success of our approach, we believe there is ample room for further progress. First, our estimation results portray employer prejudice in the US manufacturing sector as strong and widespread. Lang and Lehmann (2012) consider this possibility implausible as it ‘does not seem likely that a large proportion of employers (…) are willing to forego significant profits in order to avoid hiring blacks’. We agree that a lower degree of employer prejudice would be a more credible result. We conjecture that this result can be explained in part by the exclusion of channels whereby prejudice translates into worse labor market outcomes for blacks. In particular, in our model all matches are exogenously destroyed at the same rate. If the decision to fire a worker were also modeled it would imply black workers were fired more often, which would subsequently affect the decision to match with them in a negative way. It would be trivial to impose different rates of exogenous job destruction across worker types, but we find this a rather ad hoc modeling strategy. We do not explore this conjecture further in this paper as doing so would require introducing idiosyncratic productivity shocks to worker-firm matches, which would add significant complexity to the model and to the estimation protocol. This is work for future research.

The lack of publicly available matched employee-employer data sets for the US economy tied our hands in terms of the parameters that could be estimated — in particular, the degree of complementarities in the production function could not be estimated. The estimation of production complementarities would have allowed us to make milder functional form assumptions regarding the production function, which would provide greater insight into the exact effect of differential sorting patterns governed by employer prejudice. Finally, it would be interesting to extend the study of discrimination using our approach to other industries in the US economy beyond manufacturing, other countries and other forms of discrimination (e.g. gender).

A final and a perhaps more fundamental critique to our approach is the way in which we model discrimination. We focus on taste-based discrimination and shut down any channels of statistical discrimination.\textsuperscript{30} Introducing statistical discrimination would require modeling uncertainty regarding worker’s skill. An interesting endeavour would be to introduce uncer-
tainty into the model and disentangle the relative effects of the two types of discrimination. This could be particularly valuable to guide empirical work. Indeed, despite the breadth of empirical studies in the racial discrimination literature, there is no systematic evidence pointing to one approach as being more plausible than the other (see Charles and Guryan (2011) and Lang and Lehmann (2012)).
References

ABOWD, J., F. KRAMARZ AND D. MARGOLIS, “High wage workers and high wage firms,” 
_Econometrica_ 67 (1999), 251–333.

ARROW, K., _The Theory of Discrimination_, volume 1 of _Discrimination in the Labor Mar-

mimeo (2012).

Matched Employer-Employee Data,” _Journal of Human Resources_ (Forthcoming).


BLACK, D. A., “Discrimination in an Equilibrium Search Model,” _Journal of Labor Eco-

BLANCHFLOWER, D., A. OSWALD AND P. SANFEY, “Wages, Profits and Rent-Sharing,” 

BOLLINGER, C., “Measurement Error in the CPS: A Nonparametric Look,” _Journal of 

BOWLUS, A. J. AND Z. ECKSTEIN, “Discrimination and Skill Differences in an Equilibrium 

CAIN, G. G., _Chapter 13 The economic analysis of labor market discrimination: A survey_, 


EECKHOUT, J. AND P. KIRCHER, “Identifying Sorting — In Theory,” _The Review of Eco-
nomic Studies_ 78 (2011), 872.


A Omitted proofs

In all the proofs presented in this section we assume the production function \( f(h, x) \) satisfies the regularity conditions stated in assumption A0 in Shimer and Smith (2000) and strict supermodularity. Formally, we have that:

**Assumption 1** (Production Function): The production function \( f(h, x) \) is nonnegative, symmetric, continuous, and twice differentiable, with uniformly bounded first partial derivatives on \([0,1]^2\). The production function \( f(h, x) \) is strictly supermodular. That is, if \( h > h' \) and \( x > x' \), then \( f(h, x) + f(h', x') > f(h, x') + f(h', x) \).

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

Part 1 of Lemma 1 in Shimer and Smith (2000) states that, for any worker of any ability level, her unemployment value can only be smaller if evaluated at some alternative (non-optimal) matching set. This result applies in our environment as well.\(^{31}\) In particular, it implies that, for any type-1 worker with ability \( h \):

\[
\rho U_1(h) \geq b + \lambda^W \frac{\beta}{\rho} \sum_{j=P,N} \int \alpha^j_2(h, x) \left[ W^j_1(h, x) - U_1(h) \right] \frac{v^j(x)}{vP^P + vN} dx. \tag{24}
\]

Subtracting \( \rho U_2(h) \) to both sides of this inequality, substituting in the bargaining solution (equation 9) and rearranging one obtains the following inequality:

\[
U_1(h) - U_2(h) \geq \frac{d}{\rho} \times \frac{\lambda^W \frac{\beta}{\rho} \int \alpha^P_2(h, x) v^P(x) dx}{1 + \frac{\lambda^W \frac{\beta}{\rho} \sum_{j=P,N} \int \alpha^j_2(h, x) v^j(x) dx}{vP^P + vN} dx} \tag{25}
\]

If \( \pi \in (0,1) \) then all workers face a positive probability of meeting a prejudiced firm due to random matching frictions, i.e. \( \frac{v^P(x)}{vP^P + vN} > 0, \forall x \). Since we are characterizing equilibria in which at least some matches of every type are feasible, the integral in the numerator is always positive and so, when \( d > 0 \), \( U_1(h) > U_2(h) \), \( \forall h \).

*Mutatis mutandis*, one can prove that \( V^N(x) > V^P(x) \), \( \forall x \).

**Proof of Corollary 1. Wage Discrimination**

Take an arbitrary \( h \) and \( x \). If \( \pi \in (0,1) \) and \( d > 0 \), then \( U_1(h) > U_2(h) \), \( \forall h \) and :

\[
w^P_2(h, x) = \beta \left[ f(h, x) - d - \rho V^P(x) \right] + (1 - \beta) \rho U_2(h) < \beta \left[ f(h, x) - \rho V^P(x) \right] + (1 - \beta) \rho U_2(h) < \beta \left[ f(h, x) - \rho V^P(x) \right] + (1 - \beta) \rho U_1(h) = w^P_1(h, x). \tag{26}
\]

\(^{31}\)The proof is available from the authors upon request.
We have proven that \( w^T_1(h, x) > w^T_2(h, x), \forall(h, x) \).

_Mutatis mutandis_, one can prove that \( w^N_1(h, x) > w^N_2(h, x), \forall(h, x) \).

**Proof of Corollary 2. Type-2 Hiring Discrimination by Prejudiced Firms**

Using equation 9 one can write \( S^P_1(h, x) \) as a function of \( S^P_2(h, x) \):

\[
S^P_1(h, x) = S^P_2(h, x) + \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta}
\]

First, proving \( d + \rho(U_2(h) - U_1(h)) > 0 \) will imply \( S^P_1(h, x) > S^P_2(h, x), \forall(h, x) \). The proof is similar to the proof of Proposition 1 except that we are interested in the upper limit of the difference \( U_1(h) - U_2(h) \). Specifically,

\[
U_1(h) - U_2(h) \leq \frac{d}{\rho} \times \frac{\lambda^W \beta}{\rho + \delta} \sum_{j = 1}^{N} \int \alpha'_1(h, x) \nu^P(x) \, dx
\]

Using the same argument as in the proof of Proposition 1, we have \( U_1(h) - U_2(h) < \frac{d}{\rho} \).

Take an arbitrary \((h, x)\) and assume that a type-2 worker will not be hired, i.e. \( S^P_2(h, x) \leq 0 \). She will suffer from hiring discrimination by prejudiced firms for all combinations of \((h, x)\) such that \( 0 < S^P_1(h, x) < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta} \).

Recall that we are characterizing equilibria where some but not all matches of every kind are feasible, which implies \( \exists(h, x) : S^P_1(h, x) > 0 \) and \( \exists(h', x') : S^P_1(h', x') \leq 0 \). Part 2 of Lemma 1 in Shimer and Smith (2000) states that the values of unmatched agents are Lipschitz and thus continuous. This result applies in our environment as well.\(^{32}\) In particular, it implies that \( S^P_1(h, x) \) is continuous with respect to both \( h \) and \( x \). Hence, by the intermediate value theorem, \( \exists(h', x') : 0 < S^P_1(h', x') = \epsilon < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta} \).

We have shown that there exists at least one combination of \((h', x')\) between a type-1 worker and a prejudiced firm such that \( 0 < S^P_1(h', x') < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta} \). Since the support of workers’ ability is the same, \( S^P_2(h', x') \) is well-defined and we know that \( S^P_2(h', x') \leq 0 \), i.e. a type-2 worker of the same ability suffers hiring discrimination by the very same firm.

**Proof of Corollary 3. Type-1 Hiring Discrimination by Nonprejudiced Firms**

\( S^N_2(h, x) \) can be expressed as a function of \( S^N_1(h, x) \):

\[
S^N_2(h, x) = S^N_1(h, x) + \frac{\rho(U_1(h) - U_2(h))}{\rho + \delta}
\]

Proposition 1 implies \( S^N_2(h, x) > S^N_1(h, x), \forall(h, x) \). The rest of the proof is analogous to the proof of Corollary 2.

\(^{32}\)The proof is available from the authors upon request.
Proof of Corollary 4. Threshold Productivity Differences

We first state a result that will be instrumental in showing both results of interest. It states that, for each type of firm, the value of posting a vacancy is a monotonically increasing function in firm’s technology.

**Proposition 2 (Monotonically Increasing Outside Options):**

(i) $V^j(x)$ is monotonically increasing in $x$, for $j = N, P$; and

(ii) $U_i(h)$ is monotonically increasing in $h$, for $i = 1, 2$.

The proof is, again, based on Part 1 of Lemma 1 in Shimer and Smith (2000). The value inequality lemma states that for an arbitrary $x'$:

$$\rho V^j(x) \geq -\kappa + \lambda F \sum_{i=1,2} \int \alpha^j_i(h, x') \left[ f_i^j(h, x) - V^j(x) \right] \frac{u_i(h)}{u_1 + u_2} \, dh.$$  

Hence, for all $x_1 < x_2$

$$V^j(x_2) - V^j(x_1) \geq \frac{\lambda F (1-\gamma)}{\rho + \delta} \sum_{i=1,2} \int \alpha^j_i(h, x_1) \left( f(h, x_2) - f(h, x_1) \right) \frac{u_i(h)}{u_1 + u_2} \, dh.$$  

Since we assume production complementarities, specified by a supermodular production function, $f(h, x_2) - f(h, x_1) > 0$ for all $x_1 < x_2$, we also have that $V^j(x_2) - V^j(x_1) > 0$.

*Mutatis mutandis*, the proof is the same for $U_i(h)$.

We now turn to the proof of Corollary 4. In the type of equilibria we study we always assume, without loss of generality, that the least productive nonprejudiced firm makes zero profit, i.e. $V^N(0) = 0$. Proposition 2 states $V^N(x)$ is monotonically increasing in $x$, which implies all nonprejudiced firms are active. Whenever $\pi \in (0, 1)$ and $d > 0$, by Proposition 1 $V^P(0) < 0$. Then, by Proposition 2 and continuity of $V^P(x)$, there exists a threshold productivity, $x^P_\star > 0$, such that $V^P(x^P_\star) = 0$.

## B Simulation Algorithm

The application of our chosen estimation method requires the equilibrium of the model to be solved numerically. This section describes how this is carried out in practice.

1. We set initial values for all the equilibrium objects (values, densities, matching sets, the total number of firms in the economy and the threshold technology level of prejudiced firms) and parameters of the model $\theta = \{\lambda, \delta, d, \pi, \mu_1, \sigma_1, \mu_2, \sigma_2, \mu_x, \sigma_x, \beta, b, \kappa\}$. 

36
We discretize the supports of the ability and technology measures of firms and workers in the economy.

2. Given all initial values we find new values for $G$ and $x^{P*}$ such that $V^P(x^{P*}) = 0$ and $V^N(0) = 0$

3. Using the initial values and updated $G$ and $x^{P*}$ we iterate over equations (16) and (17) to determine $U_i(h)$ and $V^j(x)$, at each stage updating the region of feasible matches determined by equation (10).

4. New values for $u_i(h)$ and $g^j(x)$ are obtained from equations (14) and (15).

5. Given the values determined in the previous step, new values of $u_i$ and $v^j$ as well as $\lambda^W$ are determined.

6. Steps 3 through 5 are updated until the endogenous distributions $u_i(h)$ and $g^j(x)$ converge.