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

Daniel Borowczyk-Martins[†]
Copenhagen Business School and IZA

Jake Bradley[‡]
University of Cambridge and IZA

Linas Tarasonis[§]
Bank of Lithuania and Vilnius University

August 2017

Abstract

In the U.S. the average black worker has a lower employment rate and earns a lower wage compared to his white counterpart. Lang and Lehmann (2012) argue that black-white wage and employment gaps are smaller for high-skill workers. We show that a model combining employer taste-based discrimination, search frictions and skill complementarities can replicate these regularities, and estimate it using data from the U.S. manufacturing sector. We find that discrimination is quantitatively important to understand differences in wages and job finding rates across workers with low education levels, whereas skill differences are the main driver of those differences among workers with high education levels.

Keywords: employment and wage gaps, discrimination, job search, sorting.

JEL codes: J31; J64; J71.

[†]danielbm@gmail.com.
[‡]jake.c.bradley@gmail.com.
[§]ltarasonis@gmail.com.

1 Introduction

In their survey of the economic literature on racial discrimination Lang and Lehmann (2012) document persistent differences in employment and wages across black and white workers in the United States (U.S.). They argue that negative black-white employment and wage gaps are the two main empirical regularities a model of discrimination should seek to replicate.[1] Critically, their review of the evidence suggests these gaps vary considerably by skill. In particular, wage gaps '*are smaller or nonexistent for very high-skill workers*' and employment gaps are '*somewhat smaller among high-skill than among low-skill workers*' (p.12). The authors also assess the ability of existing discrimination models to replicate these facts. They conclude that '*no existing* [discrimination] *model can fully explain these regularities*' (idem).

In this paper we develop a model of discrimination that successfully replicates these regularities. We show that a model of taste-based employer discrimination can deliver simultaneously mean black-white wage and employment gaps, as well as a decreasing profile of these gaps as the skill of workers increases.[2] Recent evidence in the economics literature suggests labor market discrimination remains a plausible hypothesis to rationalize observed wage and employment gaps across races (see Fryer et al. (2013)). Correspondence and audit studies find pervasive evidence of unequal treatment of black workers vis-a-vis seemingly equally skilled white workers (see Bertrand and Mullainathan (2004) and Charles and Guryan (2011)). The latest evidence produced by regression-based studies using the methodology proposed by Neal and Johnson (1996) points to the conclusion that, although differences in premarket factors are likely to play a major role in explaining observed mean black-white wage gaps, a substantial wage gap remains after controlling for premarket factors (Lang and Lehmann, 2012). Ritter and Taylor (2011) use Neal and Johnson's methodology to measure the importance of premarket factors for observed mean black-white employment gaps and also find that a substantial gap remains after controlling for premarket factors and a number of other variables.[3]

---

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

[2]We are not claiming that the model developed in this paper is the only one that can successfully replicate these empirical regularities. In fact, recent models of customer taste-based discrimination and statistical discrimination can also replicate those patterns (see Cavounidis and Lang (2016) and Decreuse and Tarasonis (2016)).

[3]Neal's and Johnson's investigation of the role of premarket factors in explaining mean black-white wage gaps had a major impact in the literature. In that study the authors argue that controlling for differences in premarket factors (measured by the Armed Forces Qualification Test at an individual's young age) can fully explain the mean wage gap across blacks and whites in the U.S. As mentioned in the main text, recent evidence (see Carneiro et al. (2005) and Lang and Manove (2011)) suggests a more nuanced view of the importance of premarket factors.

Two competing approaches dominate the economic literature on discrimination: prejudice or taste-based models, pioneered by Becker (1971), and models of statistical discrimination, starting with Phelps (1972) and Arrow et al. (1973).[4] Despite the voluminous empirical literature on racial discrimination, there is no systematic evidence pointing to one approach as being more plausible than the other (Charles and Guryan (2011) and Lang and Lehmann (2012)). We find both approaches theoretically compelling and existing evidence suggests both are important to rationalize the data. However, in this paper we set statistical discrimination aside and focus only on the consequences of taste-based discrimination for differences in labor market outcomes of blacks and whites. In particular, we build on a modeling approach that combines employer taste-based discrimination and random search frictions to describe differences in labor market outcomes of workers who differ in terms of a nonproductive attribute (e.g. race or gender). The assumption of search frictions is a natural modeling choice in this context, since there are sizable differences in mean unemployment durations of black and white individuals.

Previous models in the discrimination literature have shown how the combination of taste-based discrimination and random search frictions generates mean employment and wage differentials across races.[5] The central intuition is the following. Consider an economy populated by two types of workers (who differ by race) and two types of employers, where one type (prejudiced) incurs a utility cost from hiring a black worker. Since prejudice reduces the match value between prejudiced employers and black workers, the matching opportunities of black workers are smaller compared to those of white workers. Under random search, black workers cannot direct their search away from prejudiced employers (their probability of meeting a prejudiced employer is the same as that of whites), so in this setup black workers have lower employment prospects compared to white workers. This delivers mean employment differentials across races. Lower employment prospects in turn imply black workers have lower reservations values. Since all employers know it takes longer for black workers to find a job, they will take advantage of that and offer black workers lower wages. This delivers mean racial wage differentials.

To generate wage and employment gaps that are smaller for high-skill workers compared to low-skill workers we extend this modeling approach in two directions. We start by assuming that workers and firms differ respectively in their levels of skill and technology and that the production value of the match is a complementary function of firms' and workers' skill levels. Under production complementarities and capacity constraints,

---

[4]The literature that followed Phelps (1972) focuses on the possibility that blacks' productivity is more difficult to observe than that of whites, while the literature that builds on Arrow et al. (1973) stresses the effects of differences in employers' beliefs about blacks' and whites' productivity.

[5]This result was first shown in Black (1995).

2

in equilibrium, high(low)-skill workers will be matched more frequently with high(low)-technology firms. In other words, there will be positive assortative matching on skill. Because production is also an increasing function of workers' and firms' skill levels, matches involving high-skill workers will involve higher levels of production. If we further assume the utility cost for prejudiced employers of employing a black worker is constant, then the cost of prejudice represents a smaller share of the value of matches as the skill of workers increases. This will translate into lower employment and wage differentials across races as the level of workers' skill increases.

The model developed in this paper shares several features with other models of taste-based discrimination in a random search environment, like Black (1995), Bowlus and Eckstein (2002), Rosén (2003) and Flabbi (2010).[6] In all these models prejudiced employers incur a psychic cost of employing a worker who belongs to the minority group.[7] Our model setup differs from these papers by assuming two-sided skill heterogeneity, production complementarities and endogenous vacancy posting. These assumptions have strong empirical support. The analysis of matched employer-employee data sets over the past two decades established the importance of both worker and firm unobserved heterogeneity to explain observed differences in wages across workers (see Abowd et al. (1999) and Lopes De Melo (forthcoming)). Production complementarities are increasingly seen as a plausible description of the production technology in modern labor markets.[8] Although this question is not fully settled, there is a growing consensus that positive sorting on skill is an important feature of labor markets. Lastly, the assumption of endogenous vacancy posting is supported by evidence on the empirical relevance of an aggregate matching function in the U.S. labor market and on firms' response to changes in vacancy filling rates (see Petrongolo and Pissarides (2001), Borowczyk-Martins et al. (2013) and Davis et al. (2013)).

We estimate the model using various sources of publicly available data for the U.S. manufacturing sector.[9] A critical feature the model must satisfy to make its empirical

---

[6]We only refer to papers based on a similar modeling approach. See Lang and Lehmann (2012) for a comprehensive review of all discrimination models.

[7]Black (1995) assumes prejudiced employers never match with black workers. Bowlus and Eckstein (2002) develop a wage posting model and allow both types of worker to draw their productivity from separate (degenerate) distributions. Rosén (2003) and Flabbi (2010) model the productivity of the match (and not of workers and jobs) as being heterogeneous.

[8]Shimer and Smith (2000) established the equivalence between a specific form of production complementarities and positive assortative matching in a random search environment, generalizing the famous result in a competitive setup put forth by Becker (1973). Several papers in the applied search literature have tried to measure the sign and strength of assortative matching in skill nonparametrically (see Abowd et al. (1999), Eeckhout and Kircher (2011), Lopes De Melo (forthcoming) and Hagedorn et al. (2017)), or estimate the degree of skill complementarities in production using a structural approach (see Bagger and Lentz (2016) and Lise et al. (2016)).

[9]Two-sided skill heterogeneity is a distinctive feature of the model that calls for an estimation strategy

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) and Fryer Jr (2011)), which suggests differences in skill across races are likely to play an important role in shaping mean employment and wage differentials. Other models of taste-based employer discrimination have been estimated using structural methods (see Bowlus and Eckstein (2002) and Flabbi (2010)). However, our paper is the first to take to the data a search-discrimination model based on Shimer and Smith's (2000) two-sided skill heterogeneity partnership model. An important distinction of our estimation strategy with respect to other papers in the literature is that prejudice and skill differences are identified by matching jointly wage and unemployment-to-job transition moments.

To empirically distinguish between education and skill, we separately estimate the model among low- and high-educated individuals (those without any college education and those with some college education or more, respectively). We find stark differences in the relative importance of prejudice and skill differences across races to explain differences in wages and job finding rates. For those without a college degree, we estimate that 57% of potential employers are prejudiced and that the utility cost of employing a black worker is about 7 dollars per hour. In terms of skill, blacks and whites are estimated to be quite homogeneous, with the average white worker being 1.4% more skilled than his black counterpart. In contrast, among those with some college education, prejudice is much more concentrated (only 31% of potential employers are prejudiced) and milder (the disutility of employing a black worker is $1.4), whereas these workers are considerably more heterogeneous in terms of their productive abilities (both within and between races). The average black worker is close to 14% less productive than the equivalent white worker.

We use the estimated model to gauge the effectiveness of alternative policy approaches to improve labor market outcomes of black vis-a-vis white workers. In our counterfactual experiments, we consider two types of policies: enforcement of the equal treatment principle and programs aimed at reducing the premarket skill racial gap. Since our model features free entry of jobs, the effects of different policies are analyzed taking labor demand adjustment into account. Our results suggest policies can improve labor market outcomes of black workers without large decreases in social welfare.

The rest of the paper is structured in the following way. In Section 2 we present the theoretical model and derive its main properties. Section 3 describes the data. The estimation procedure is presented in detail in Section 4. Section 5 describes the fit of

---

based on matched employer-employee data. Unfortunately, at the time we started this project there was no matched employer-employee database for the U.S. labor market accessible for researchers based outside the U.S.

the model and reports parameter estimates. In Section 6 we simulate the estimated model to provide an interpretation of observed cross-racial differences in wages and job finding rates. Section 7 undertakes counterfactual analysis of different policies. Section 8 concludes.

2   Model

The model we develop in this section builds on Shimer 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 in the estimation section. Section 2.7 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 2.7 set out the model in detail and discuss the relevance of the model's assumptions to study discrimination in the labor market.

2.1   Environment

We consider a labor market with L workers and G jobs. The number of jobs G is determined in equilibrium while L is given. There exist two types of employers 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 π of jobs are operated by prejudiced employers (P) and a share $(1-\pi)$ by nonprejudiced ones (N), with index $j = P, N$. In this model one firm is one job. Thus, throughout the text we use the terms *jobs*, *employers* and *firms* interchangeably. In addition, workers differ in terms of skill, h, and firms in terms of technology (efficiency of labor inputs), x. We assume these distributions (often simply referred to as 'skill distributions') are uniform, i.e. $h \sim U[0,1]$ and $x \sim U[0,1]$). We [10] Let $\ell_i(h)$ and $g^j(x)$ denote respectively the population measures of type-i workers of skill h and type-j firms of technology x.[11] The endogenous measures of type-i unemployed workers of skill 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 $v^j = \int v^j(x) \, dx$.

---

[10]These assumptions about the supports and densities of the distributions are normalizations. In the empirical application we allow 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/firms' skill/technology, one can think of h and x as the ranks of the underlying skill/technology distributions.

[11]So that $mL = \int \ell_1(h) \, dh$ and $\pi G = \int g^P(x) \, dx$.

Time is continuous and both workers and employers are risk neutral, with discount rate ρ. Employers and workers maximize the present discounted value of future utility streams, measured in monetary terms. As in Becker (1971), prejudiced employers incur a psychic cost d of employing a type-2 worker. While this assumption (a constant disutility of prejudice) is standard in the literature, in a setting with two-sided skill heterogeneity, a more general specification would allow the degree of prejudice to vary with employers' technology.[12] In terms of the main motivating empirical pattern of the paper (decreasing wage and employment gaps in workers' skill), our model is able to replicate it under the assumption that d is positively (though not perfectly) related to employers' skill. However, under that assumption, and given our data, we can no longer confidently separately identify prejudice from average skill differences, which is a central goal of the paper.[13]

The match flow output depends on the worker's type and skill and on the firm's technology, denoted $f_i(h, x)$ and satisfying certain regularity conditions (see A).[14] We take complementarities in skill as a descriptive feature of modern labor markets and so assume a supermodular production function.[15] This means the own marginal product of any worker and job is increasing in his partner's skill. Formal details are provided in A.[16]

We assume that only unemployed workers and vacant firms search for a partner. By assuming that workers cannot search on the job we are potentially eliminating an important source of observed wage differentials between blacks and whites. The main reason why we do not include this mechanism in our model is because we lack the necessary data to inform it in a meaningful way. A limitation of the data set we use in this paper (the Current Population Survey, CPS) is that we do not observe wages for the same individual before and immediately following a job change. Therefore, any inference we were to make regarding wage changes after a job-to-job transition would be entirely driven

---

[12]A constant disutility of prejudice 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 Bowlus and Eckstein (2002)).

[13]In the context of our empirical application, to our best knowledge, there is no direct evidence that points to a specific relationship between employer's technology and the intensity of racial prejudice. Since d is a preference parameter, it is fundamentally unobservable. The type of evidence that is usually used to directly measure the extent of prejudice (or racial attitudes) is based on survey data and questions that involve a yes-or-no answer (whether respondents would vote for a black president or approve interracial marriage). While we find this evidence useful to describe the evolution of aggregate long-run trends and spatial differences across countries or states (as for example Lang and Lehmann (2012) and Charles and Guryan (2008) do), it is not straightforward to map the answers to those questions onto an exact specification of the relationship between prejudice and firms' technology.

[14]The worker's type influences output only if there are differences in the skill distributions across types.

[15]There is a growing consensus in the applied search literature on the descriptive relevance of positive assortative matching (see Lise et al. (2016), Lopes De Melo (forthcoming) and Bagger and Lentz (2016)).

[16]In our empirical application we assume a specific degree of skill complementarities. However, the model, as well as the equilibrium characterization described in Section 2.7, hold with any production function that satisfies the regularity conditions and supermodularity.

by the chosen wage determination structure. For wages of entrants from unemployment, Nash bargaining with unemployment as a threat point is a common assumption in the literature. However, how wages are determined following a job-to-job transition employed really needs to be disciplined by data. For example, papers by Cahuc et al. (2006), Hagedorn et al. (2017) and Bradley and Kuegler (2017) all use the same wage determination mechanism as ourselves for those hired from unemployment, but assume three different mechanisms for wage changes associated with job switches. Therefore, while recognizing that on-the-job search is a potentially very important channel to understand how prejudice translates onto differential racial outcomes along workers' careers, we deem the innovations in our model worthy of a separate analysis and so leave this important extension for future research.

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 meeting 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 random and vacant jobs and unemployed workers of different types and skills are effectively exposed to the same arrival rates, respectively $\lambda^W = \frac{M(u_1+u_2,v^P+v^N)}{u_1+u_2}$ for workers, and $\lambda^F = \frac{M(u_1+u_2,v^P+v^N)}{v^P+v^N}$ for jobs.

Once a firm and a worker meet they decide whether or not to form a match. We define a match indicator function $\alpha_i^j(h,x)$, which is equal to 1 if a type-i worker of skill 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 δ, in which case both the worker and the firm join the pool of searchers. This job destruction rate is assumed constant. We assume away differences in δ across types because it affects sorting patterns, via changes in the relative value of matches across types, and therefore differences in job finding rates. If matches with a specific worker type have shorter expected durations, employers will take this into account when deciding whether or not to form a match. This effectively imposes another source of discrimination at the hiring stage, which is not modeled, and that also affects the complementarities between the worker's type and a firm's output.[17]

2.2 Value Functions

Having defined the setup, we can now write each agents' payoffs. A worker of type-i and skill level h can be in one of two different states: employed or unemployed. The flow value of employment when he is employed in a type-j firm of technology x is given by equation

---

[17]We discuss this point in more detail in section 4.2.1.

(1):

$$\rho W_i^j(h, x) = w_i^j(h, x) + \delta \left[U_i(h) - W_i^j(h, x) \right]. \tag{1}$$

, where $w_i^j(h, x)$ is the wage she earns in this job. The wage $w_i^j(h, x)$ is endogenously determined as the solution to a Nash bargaining game. $U_i(h)$ denotes the worker's value of being unemployed.

While unemployed a worker receives a flow utility b, independent of his race, skill and employment history. The value of unemployment reads as follows:

$$\rho U_i(h) = b + \lambda^W \sum_{j=P,N} \int \alpha_i^j(h, x) \left[W_i^j(h, x) - U_i(h) \right] \frac{v^j(x)}{v^P + v^N} \, \mathrm{d}x, \tag{2}$$

where $\frac{v^j(x)}{v^P + v^N}$ is the sampling probability of a vacant type-j job of technology x from the pool of unmatched firms.

Firms can be in two states: vacant or filled. $J_i^j(h, x)$ is the value of a job for a type-j firm of technology x, filled with a type-i worker of skill h:

$$\rho J_i^j(h, x) = f_i(h, x) - d\mathbf{1}_{[(i,j)=(2,P)]} - w_i^j(h, x) + \delta \left[V^j(x) - J_i^j(h, x) \right], \tag{3}$$

where the indicator function $\mathbf{1}_{[(i,j)=(2,P)]}$ takes value 1 if the match involves a prejudiced employer and a type-2 worker. The psychic cost d that a prejudiced employer incurs when matching with a type-2 worker does not affect the production value of the match, but it does affect the utility value of the match for a prejudiced employer. Because d enters additively in the firm's value function, the degree of prejudice is independent of workers' and firms' skill levels.

We assume that keeping a vacancy open is costless, but requires paying an upfront sunk cost. Therefore, the value of posting a vacancy for a type-j firm of technology x, $V^j(x)$, is given by the following equation:

$$\rho V^j(x) = \lambda^F \sum_{i=1,2} \int \alpha_i^j(h, x) \left[J_i^j(h, x) - V^j(x) \right] \frac{u_i(h)}{u_1 + u_2} \, \mathrm{d}h, \tag{4}$$

where $\frac{u_i(h)}{u_1+u_2}$ is the probability of sampling an unemployed type-i worker of skill 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 $V^j(x) \geq 0$. Since there is no flow cost associated with keeping a vacancy open, by construction, this is always true. To determine the total mass

of active jobs in equilibrium, G, we assume a free entry condition, whereby new entrants make the decision to enter behind a veil of ignorance, i.e. without knowledge of either their productivity type x or their taste of prejudice (see Hopenhayn (1992)). In order to enter firms must incur a fixed sunk cost, denoted κ. One can think of it as the required capital investment to create a new job. After entry, vacancies stay open until they meet an acceptable worker. The total mass of active jobs, G, is determined by equalizing the expected value of opening a vacancy to the fixed cost, κ, as formalized in the equation below:

$$\pi \int V^P(x)\frac{g^P(x)}{\pi G}dx + (1-\pi) \int V^N(x)\frac{g^N(x)}{(1-\pi)G}dx = \kappa. \tag{5}$$

As equation (5) highlights, G directly affects the value of retaining an open vacancy. Since the matching function has diminishing returns in the number of vacancies, as their number of increases, the probability that an individual vacancy is filled decreases. This, in turn, depresses the value of keeping a vacancy open.

Our model of entry departs from other search-discrimination models in the literature by introducing free entry of jobs and endogenous meeting rates via an aggregate meeting function. Like previous search-discrimination models, prejudiced and nonprejudiced firms coexist on the market due to the presence of search frictions, and product market competition is not explicitly modeled. This is an important limitation of the model, insofar as product market competition affects the equilibrium distribution of prejudiced and nonprejudiced firms which, in turn, affects the existence and persistence of racial wage and employment gaps. Becker (1971) famously conjectured that competitive forces (namely competition in the product and capital markets) would drive prejudiced firms out of the market in the long run. Addressing the long-run implications of product market competition would require us to formalize the industry dynamics, including the processes of ownership transfer and firm entry and exit, which would greatly complicate the model and the estimation protocol. Therefore, while recognizing this limitation, we do not attempt to address this question in the paper.[18]

---

[18]In a recent paper, Ederington and Sandford (2016) assess Becker (1971)'s long-run prediction on the survival of prejudiced firms by formalizing a model of industry dynamics with prejudiced and nonprejudiced firms. Their findings offer a more nuanced view of Gary Becker's prediction. In their concluding remarks they write: '*Indeed, we show that some degree of market power (driven by product differentiation across firms in our model) is necessary for discriminatory firms to exist even in the short run. However, we also prove that the cost structure of an industry is of equal importance in explaining the survival of discriminating firms. Specifically, we show that, even in industries characterized by low degrees of market power (i.e., low levels of product differentiation and thus low price-cost margins), the presence of high entry costs, low per-period entry costs and sequential entry can generate the long-run survival of discriminatory firms).*' (p. 30).

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 skill h employed in a type-j firm of technology x writes: $S_i^j(h,x) = W_i^j(h,x) - U_i(h) + J_i^j(h,x) - V^j(x)$, and it is split between the worker and firm according to Nash bargaining. The worker takes a share β and the firm a share $(1-\beta)$ of the match surplus, implying the following equalities:

$$S_i^j(h,x) = \frac{J_i^j(h,x) - V^j(x)}{1-\beta} = \frac{W_i^j(h,x) - U_i(h)}{\beta}. \tag{6}$$

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

$$w_i^j(h,x) = \beta\left[f_i(h,x) - d\mathbf{1}_{[(i,j)=(2,P)]} - \rho V^j(x)\right] + (1-\beta)\rho U_i(h). \tag{7}$$

Using equations (1) and (3), the expression for the total surplus can be rearranged and expressed in the following way:

$$S_i^j(h,x) = \frac{f_i(h,x) - d\mathbf{1}_{[(i,j)=(2,P)]} - \rho U_i(h) - \rho V^j(x)}{\rho + \delta}. \tag{8}$$

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

$$\alpha_i^j(h,x) = \mathbf{1}\left[f_i(h,x) - d\mathbf{1}_{[(i,j)=(2,P)]} - \rho U_i(h) - \rho V^j(x) > 0\right]. \tag{9}$$

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 to 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 for it to differ across worker types and/or skill levels.[19] Moreover, in the context of our model, absent taste-based discrimination (i.e. if d were zero) a lower β for type-2 workers by itself would not generate employment differences across worker types.

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 alternative specifications (for instance, one in which d is nontransferable) may be equally

---

[19]Some papers interpret differences in rent-sharing parameters as the result of discrimination (see Eckstein and Wolpin (1999) and Bartolucci (2013)).

plausible. Although assuming a nontransferable d entails a somewhat different description of how prejudice translates into discrimination, it implies similar properties to our chosen specification. In particular, it also generates hiring and wage discrimination, as well as higher employment and wage gaps for low-skill workers compared to high-skill workers.

2.5 Matching Sets

Having established the structure of agents' payoffs, we can now define each agent's strategy. For a type-i worker of skill h her strategy is given by two sets, $\mathcal{M}_i^P(h)$ and $\mathcal{M}_i^N(h)$. Similarly, a firm's strategy is defined by two sets, $\mathcal{M}_1^j(x)$ and $\mathcal{M}_2^j(x)$. 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 $\alpha_i^j(h, x)$ we can express each worker's matching set as:

$$\mathcal{M}_i^j(h) = \{x \,|\, \alpha_i^j(h, x) = 1\}, \tag{10}$$

and each firm's matching set as:

$$\mathcal{M}_i^j(x) = \{h \,|\, \alpha_i^j(h, x) = 1\}. \tag{11}$$

2.6 Steady-state Equilibrium

A steady-state equilibrium in 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 measures of unmatched workers of type-i and skill h and firms of type-j and technology x, respectively $u_i(h)$ and $v^j(x)$, are in steady-state and (iv) in expectation the value of opening a vacancy is zero. Conditions (i) and (ii) are given respectively by firms' and workers' value functions and their matching sets. Condition (iv) is given by the free entry condition. We now state the assumptions necessary to ensure condition (iii).

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

$$\lambda^W \alpha_i^j(h, x) \, u_i(h) \, \frac{v^j(x)}{v^P + v^N} = \delta \gamma_i^j(h, x), \tag{12}$$

where $\gamma_i^j(h, x)$ is a joint measure of matched type-i workers of skill 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

11

point in time (the left-hand side of equation (12)) is exactly the same as the number of matches being destroyed (the right-hand side of equation (12)). Then, by definition, the steady-state stock of type-i employed workers of skill h is given by the following equation:

$$\ell_i(h) - u_i(h) = \int \gamma_i^N(h,x)\, \mathrm{d}x + \int \gamma_i^P(h,x)\, \mathrm{d}x. \tag{13}$$

That is, the total population of type-i workers of skill 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:

$$g^j(x) - v^j(x) = \int \gamma_1^j(h,x)\, \mathrm{d}h + \int \gamma_2^j(h,x)\, \mathrm{d}h \tag{14}$$

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 (6)) and firms' and workers' value functions to write each agent's equilibrium reservation value (equations (26) and (27)), where matching decisions satisfy equation (9). We then use the flow-balance equations (equation (12)) and the population accounting equations (equations (13) and 14) to express the equilibrium measures of unmatched agents (equations (28) and (29)). Finally, the number of firms is determined by the free-entry condition (equation (5)). The formal definition of equilibrium is given in B.

2.7   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 A. The results we present in this section are characteristic of a class of equilibria where, for all possible combinations of workers and firms of different types, some but not all matches are feasible. In other words, equilibria in which, for any combination between a type-i worker and a type-j firm, some matches between workers of skill h and firms of technology x are formed, but not for all possible combinations of skill and technology levels. In practice this implies that, on the one hand, the flow value of unemployment b and the psychic cost borne by prejudiced employers d are sufficiently small with respect to the value of production $f_i(h,x)$, 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 $f_i(h,x)$ is small enough with respect to b and d to render the match between them not feasible, where this holds for

12

any combination (type-i, type-j).[20] 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 and 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. 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 analytic results.

We start by establishing some basic 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)).[21] In our model, the psychic cost d reduces the utility value of the match for prejudiced 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 characterize pertains to workers' wages.

Definition 1 (WAGE DISCRIMINATION): *A type-i worker of skill 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*

$$\text{for some } (h,x), \ w_i^j(h,x) < w_{k \neq i}^j(h,x).$$

The second instance of economic discrimination 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 skill match with firms within the same range of technology. In an economy in which there is employer prejudice, in general, this is no longer the case and the matching sets of two equally skilled workers of different types will differ. One reason why these matching sets differ is due to hiring discrimination. Formally, we have that:

---

[20]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 $\alpha_i^j(h',x') = 1$ and $\alpha_i^j(h'',x'') = 0$.

[21]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.*' (p. 695).

Definition 2 (HIRING DISCRIMINATION): *A type-i worker of skill h experiences hiring discrimination if, upon meeting a firm of technology x of type-j, he is not hired, but an equally skilled type-k $\neq i$ worker is; that is,*

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

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 prospects 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): *If $\pi \in (0, 1)$ and $d > 0$, in any equilibrium:*

(i) *for a worker of skill 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): *If $\pi \in (0, 1)$ and $d > 0$, in any equilibrium all type-2 workers experience wage discrimination in both types of firms 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 skilled 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, 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 (9)), 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 2 this combination of circumstances entails hiring discrimination. The result is stated below.

Corollary 2 (Hiring Discrimination in Prejudiced Firms): *If $\pi \in (0,1)$ and $d > 0$, some type-2 workers experience hiring discrimination by some prejudiced firms.*

For nonprejudiced firms, the difference between matching with equally skilled 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 skilled type-2 worker may render that match feasible. We will refer to this combination of circumstances as *reverse hiring discrimination*. This result is stated below.

Corollary 3 (Hiring Discrimination in Nonprejudiced Firms): *If $\pi \in (0,1)$ and $d > 0$, some type-1 workers experience hiring discrimination by some nonprejudiced firms.*

3   Data

To estimate the model we use three sources of data: worker-level data, firm-level data and market-level data. In the following paragraphs we describe these data sources and the procedures we implemented to obtain the sample used in our estimation. The exact moments used in the estimation are presented and discussed in Section 5.2.

The worker-level data comes from the CPS produced by the Bureau of Labor Statistics. We pool the micro-data Basic Monthly extracts from January 2003 to December 2006 to retrieve information on wages and transitions across labor market states.

Our sample selection procedure is as follows. First, we only include individuals who declare themselves to be either black or white. Second, 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 further restrict our sample to individuals between the ages of 18 and 30 years old who remain active in the labor market throughout their spell in the sample. Since we do not model wage growth (human capital accumulation and search on the job), we believe our model is more adequate to interpret differences in labor market outcomes among individuals who are at the beginning of their career. Fourth, we only consider individuals in two labor market states: unemployed or employed in a full-time job in the non-farm business sector. Last, to provide a correspondence with our firm-level data, we restrict the sample to individuals who at some point in time were employed in a manufacturing firm.

Having defined the estimation sample, we split it in two education groups: those with less than a college education (at most 12 years of schooling) and those with at least some college education (more than 12 years of schooling). We do so based on the premise that labor markets are segmented by broad education groups. This assumption is common in the literature on structural estimation of job search models (see e.g. Lise et al. (2016)). Table 1 presents summary statistics on the two sub-samples. The 'No college' (NC) sample is composed of 9,540 males (8,464 whites and 1,076 blacks), whereas the 'Some college' (SC) sample is composed of 5,160 males (4,720 whites and 440 blacks). The share of blacks is larger in the NC sample compared to the SC one. Both samples are quite homogenous in terms of age and years of schooling, although SC whites are slightly older and more educated than their black counterparts. This partly explains the larger black-white wage and unemployment gaps. We will analyze these differences in more detail when describing the model fit (see section 5.2).

We explore the longitudinal dimension of the CPS to measure individual transition probabilities across employment and unemployment. We calculate average transitions rates for the two sub-samples using the methodology proposed by Shimer (2012). Workers' wages are measured at the household's fourth and eighth months in the survey, when information on usual weekly hours/earnings is recorded. 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 trim the top and bottom 2% of the hourly wage distributions of black and white males.[22]

Our source of firm-level data is the Manufacturing Industry Database jointly produced by the National Bureau of Economic Research (NBER) and the U.S. Census Bureau's Center for Economic Studies (CES). It contains annual industry-level data from 1958 to 2009 on output, employment and other variables for the 473 six-digit 1997 North American Industry Classification System (NAICS) industries.[23] To our knowledge, this is the only publicly available data set for the U.S. with information on value-added. For each six-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. $f_i(h, x)$). This means that, in our empirical application, technology differences across firms are only explained by the six-digit industry at which they operate.

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. We combine these data to obtain an estimate

---

[22]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)).

[23]See http://www.nber.org/nberces.

Table 1: Summary Statistics of Estimation Samples

	A. No college		B. Some college	
	Whites	Blacks	Whites	Blacks
	(1)	(2)	(4)	(5)
Age	24.47	24.40	26.26	26.09
	(.14)	(.42)	(.15)	(.46)
Years of schooling	11.02	11.37	15.19	14.66
	(.07)	(.13)	(.08)	(.24)
Unemployment rate	0.11	0.15	0.05	0.14
	(.01)	(.05)	(.01)	(.06)
Hourly wage	13.34	12.44	20.52	15.31
	(.26)	(.86)	(.60)	(1.30)
Number of observations	29875	3319	17270	1457
Number of males	8464	1076	4720	440

Notes. — Authors' calculations based on Basic Monthly extracts of the Current Population Survey, 2003-2006. Standard errors reported in parentheses. Age, years of schooling and unemployment rates refer to January 2004. Hourly wages refer to 2004 Q1.

of labor market tightness in the same period in the manufacturing sector. Specifically, 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).

4 Estimation Method

The relevant moments predicted by the model do not have a closed-form expression. Therefore, we estimate the model's parameters using a simulated method of moments (SMM) estimator. Specifically, let $\boldsymbol{\theta}$ denote the vector of structural parameters we seek to estimate, $\hat{\boldsymbol{m}}^S(\boldsymbol{\theta})$ the vector of model-generated moments and $\hat{\boldsymbol{m}}_N$ its empirical counterpart. The estimation procedure finds $\boldsymbol{\theta}$ such that the distance between the model-generated moments and their empirical counterparts is as small as possible, according to the following criterion function:

$$L_N(\boldsymbol{\theta}) = \frac{1}{2} \left(\hat{\boldsymbol{m}}_N - \boldsymbol{m}^S(\boldsymbol{\theta}) \right)' \boldsymbol{\Omega}^{-1} \left(\hat{\boldsymbol{m}}_N - \boldsymbol{m}^S(\boldsymbol{\theta}) \right). \tag{15}$$

To obtain the theoretical moments we need to simulate the model. A description of the standard simulation algorithm of the model is provided in C. The weighting matrix $\boldsymbol{\Omega}$ gives equal weight to all moments.[24] Asymptotic standard errors are approximated using

---

[24]We also experimented with an *optimal* weighting matrix based on the variance-covariance matrix of the bootstrapped moments. The estimation results are not very different but, because transition rates

a first-order Taylor series, the details of which are provided in D.

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 monthly discount rate, ρ, is set at 0.0043 (equivalent to 5% per year). The population shares of worker types, m and $(1 - m)$, are observed: they are simply the share of white and black workers in our two samples (11.3 and vs 7.8% respectively for the NC and SC samples).

We specify the aggregate 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 λ is the constant matching efficiency parameter to be estimated. Since our model does not generate aggregate dynamics, for a given set of parameter values, the aggregate number of vacancies and unemployed workers is constant. Therefore, we cannot use time-series variation in vacancies and the job finding rate to identify the meeting function elasticity. The selected value of 0.5 is in line with estimates of the aggregate matching function in the U.S. reported in Petrongolo and Pissarides (2001).

The bargaining parameter β is set equal to 0.25. It is well known that without worker-firm matched data identification of this parameter is extremely difficult. Initially we calibrated β to 0.5. We experimented with alternative values and found that a lower value, 0.25, substantially improves the fit of the value-added moments. This calibration is consistent with Flinn (2006), which has the same wage-setting mechanism for unemployed workers and finds estimates of workers' bargaining power between 0.34 and 0.44. Since we estimate the model on a sample of young workers, a slightly lower value of the bargaining parameter seems plausible.

The match production function is a key object in our model, as it regulates the degree of production complementarities and, by extension, the degree of assortative matching in the model. Since our model does not describe individual wage dynamics, to achieve identification of production complementarities (parametrically or not) we need reliable firm-level data that can be matched to individual-level data. Lacking the necessary data, we calibrate the degree of production complementarities. Specifically, we assume the

---

in our data set are less precisely estimated compared to other moments, an *optimal* weighting matrix gives much more weight to wage moments. As a result, the fit of other moments (namely transitions from unemployment to job), which we consider equally important from an economic point of view, is quite poor using the optimal weighting matrix. In addition, when moment conditions are based on relatively few observations (as is the case in our estimation) it has been shown that equally-weighted distance matrices can perform better than *optimally*-weighted ones (see Altonji and Segal (1996)).

following match production function:

$$f_i(h, x) = \exp\{\mu + s\mathbf{1}_{[i=2]} + \sigma_h \Phi^{-1}(h) + \sigma_x \Phi^{-1}(x)\} \tag{16}$$

$$\forall\, h > 0,\ x > 0 \text{ and } f_i(0, x) = f_i(h, 0) = 0,\ \forall\, h,\ x.$$

According to equation (16), production depends on the worker's skill and race and on the firm's technology. The transformations of the worker's and the firm's ranks (respectively $\exp\{\sigma_h \Phi^{-1}(h)\}$ and $\exp\{\sigma_x \Phi^{-1}(x)\}$) are equivalent to assuming that both skill and technology levels are log-normally distributed with parameters zero and σ_h and σ_x, respectively. The level of output is regulated by the parameter μ. The match production function depends on the worker's race through s, which is the proportional difference in output between equivalently ranked white and black workers.

Having calibrated and specified the model, the vector of parameters to be estimated is the following:

$$\boldsymbol{\theta} = \{\lambda, \delta, \pi, d, s, \mu, \sigma_h, \sigma_x, b, \kappa\}\,.$$

4.2   Moments

To estimate $\boldsymbol{\theta}$ we select moments 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. We summarize the information about transitions in the model in three moments: the average transition rate from job to unemployment (JTU hereafter) for all employed workers and the average transition rates from unemployment to job (UTJs hereafter) for white and black individuals. The theoretical moments are given by the following equations:

$$JTU = 1 - e^{-\delta}, \tag{17}$$

$$UTJ_i = \int \left[1 - \exp\left(-\lambda^W \sum_{j=\{P,N\}} \int \alpha_i^j(h, x) \frac{v^j(x)}{v^P + v^N}\, dx \right) \right] \frac{u_i(h)}{u_i}\, dh. \tag{18}$$

In our model, JTU is exclusively governed by δ, the parameter of the continuous-time Poisson process that determines when a match is destroyed. This implies a common JTU across the two races, which we know is at odds with the data. In the U.S. labor

market black workers experience, on average, shorter employment spells compared to white workers. In the presentation of our model, we assume away racial differences in δ because it affects sorting patterns. Since our model does not generate endogenous differences in JTUs as a result of employer discrimination and/or skill differences, in our baseline specification we ignore this source of heterogeneity and quantify the sources of heterogeneity captured by the model. In section 4 we report and discuss results obtained by fitting race-specific JTUs observed in our data set.

The UTJs for white and black workers are informative about the Poisson process governing the meeting of workers and jobs. This can be clearly seen in equation (18), where $\lambda^W = \lambda(u_1 + u_2)^{-0.5}(v^P + v^N)^{0.5}$. This expression also shows that the difference in UTJs across the two races is informative about employer prejudice (π and d). Recall that both parameters affect differently the matching rates of workers of different races, via the matching sets, $\alpha_i^j(h, x)$.[25] The matching sets also depend on workers' skill (h), so race-specific UTJs depend on the skill distributions of black and white workers. While we can be sure that skill differences across races and prejudice impact UTJs, we are less certain about their impact being sufficiently different to ensure separate parameter identification. We will return to this point in the next subsection.

4.2.2 Distribution of Wages and Value-added

The wages paid to workers of different skill and race are a key equilibrium object of the model. As can be clearly seen in the wage equation (equation (7)), the distribution of wages is informative about the distributions of workers' skill, 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 (π) and the flow value of unemployment (b). To disentangle the skill distribution of black and white individuals, employer prejudice parameters and the flow value of unemployment, we use the first four moments of the wage distributions of black and white workers.

In a model similar to ours, Flabbi (2010) illustrates how the shape of the wage distribution of the minority group (females in his application) is differently affected by skill differences and prejudice parameters. Specifically, he shows that, when all employers are prejudiced, the mass near the reservation value is much lower compared to the wage distribution that results when both prejudice parameters are set to zero. This is because matches with prejudiced employers are not viable for very low levels of productivity. When π is a number between zero and one, the wage distribution of the minority group is a mixture of the wage distributions in prejudiced and nonprejudiced employers. He also

---

[25]See Section 6 for an illustration of these effects.

demonstrates that π and d affect the shape of the minority's wage distribution differently. While π mainly affects the central tendency of the distribution, d flattens the distribution to the left of the mode.

Due to the extra channels present in our model (sorting on skill and reverse discrimination), it is not straightforward to verify the validity of Flabbi (2010)'s identification argument in our model. However, when we simulate the model varying one parameter at a time we do verify that, locally, different parameters distort the wage distributions of blacks and whites differently. This is why, in addition to the mean and standard deviation, we target the third and fourth moments of both wage distributions.

To disentangle the contribution of workers' and firms' skill to the level of wages we make use of firm-level information. Specifically, we use the first two moments of the distribution of value-added. 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.[26]

An important feature of our identification strategy is to target simultaneously race-specific wage distribution moments and UTJs. We find that a model of skill differences alone is not able to generate the magnitude of cross-racial differences in UTJs observed in the data. While the differences in the UTJs help us separately identify prejudice from differences in skill, it is less clear how we separately identify the two prejudice parameters (π and d). To reassure ourselves, we conduct a numerical analysis and find that the minimization criterion is well behaved and convex in π and d. We report the main results of this numerical analysis in E.

4.2.3 Labor Market Tightness

The ratio of vacant jobs to unemployed workers, $\Theta = \frac{V}{U}$, summarizes the degree of tightness in the labor market. We use this moment to pin down the aggregate number of firms G and, by extension, the sunk cost of posting vacancy κ.[27] Integration of equation (14) over the support of jobs' and workers' skill yields the following aggregate stock-accounting equation: $G - V = 1 - U$, where V denotes the aggregate number of vacancies and U aggregate unemployment. With simple algebraic manipulations we arrive at an expression that relates market tightness to the other aggregate stocks, namely $\Theta = \frac{G-1+U}{U}$. Hence, given a value of U, market tightness pins down G. Conditional on the estimated parame-

---

[26]We interpret firms' value-added per worker as the counterpart of the match production $f_i(h, x)$. We do so because, when there is sorting on skill, firm value-added can no longer be taken as a direct measure of firms' technology.

[27]In our estimation protocol, we treat G as a parameter to be estimated. See section C for further details on our estimation protocol.

ters, we can use the free-entry condition (equation (5)) to compute the fixed sunk cost of posting a vacancy κ.

5 Estimation Results

In this section we describe the empirical moments that we target in our estimation, evaluate the fit of alternative specifications of the model and comment on different sets of parameter estimates.

5.1 Empirical moments

Table 2 reports empirical moments in Columns 4 and 9, respectively for the NC and SC samples (panels A. and B.) Their bootstrapped standard deviations, reported in parentheses, are obtained by resampling the data 200 times.

The first three rows of each column show the empirical moments that govern worker mobility to and from unemployment. The first two moments are the UTJs respectively white and black males in the sample. The third moment is the JTU of all individuals in the sample. The observed difference in UTJs of whites and blacks is 15.9 percentage points (pp) for NC and 20.2 pp for SC. The size of these differences is consistent with an average duration of unemployment for blacks that is 20 - 30% longer than that of whites, and which is reported in various studies cited in Lang and Lehmann (2012). The JTUs are 4.2 and 1.9% for the NC and SC samples, respectively. Interestingly, the higher unemployment rates for those with no college education are driven exclusively by shorter employment durations.

The middle rows of each panel present the first four moments of the distributions of log-wages of black and white workers in each of the two samples. The difference in mean wages of black and white workers is 7 and 18 log points (lp) for the NC and SC samples, respectively. Given the selected nature of our sample, our numbers are not directly comparable to those reported in other studies. Still, they are in the same ballpark as those cited by Lang and Lehmann (2012).[28] The much lower wage gap among those with no college education is somewhat surprising.[29] The remaining moments show similar cross-racial patterns in both the NC and SC samples. The standard deviation of

---

[28]Lang and Lehmann (2012) argue that 10% is a plausible number for the mean wage gap among men, after controlling for other factors.

[29]The fact that mean wage and UTJ gaps are higher for more educated workers does not contradict the motivating facts of the paper: that wage and employment gaps are lower for high-skill workers. Lang and Lehmann (2012) review facts based on what they take to be a proxy measure of cognitive ability: the test scores of the Armed Forces Qualification Test (AFQT). Since AFQT scores are not available in the CPS, we cannot verify these facts in our data set. However, we do find that quantile wage gaps are lower at the top of the wage distribution in each of the two sub-samples.

wages is higher among whites, whereas skewnesses (both positive) and kurtoses are higher among blacks.

The bottom rows display the mean and standard deviation of the distribution of log-value-added of the 6-digit manufacturing industries in the sample.[30] The last moment used in the estimation is the average labor market tightness (the ratio of aggregate vacancies to aggregate unemployment) in the manufacturing sector during the sample period. Finally, to compare the overall fit across specifications, the last row of Table 2 reports the minimum criterion for each model (i.e. $L_N(\boldsymbol{\theta}^*)$).

5.2 Fit of the Model

Table 2 also reports simulated moments for the NC and SC samples. In each panel the first three columns report the set of simulated moments corresponding to three alternative model specifications. The first specification (Columns 1 and 6) assumes black and white workers draw their skill levels from a common distribution, so that differences in labor market outcomes are the result of employer prejudice alone ('Discrimination'). In the next specification (Columns 2 and 7) we assume there are no prejudiced employers in the economy and therefore shut down the effect of hiring and wage discrimination on workers' labor market outcomes ('Skills'). The last specification allows for both skill differences and discrimination to govern workers' labor market outcomes. This is our preferred specification and is displayed in Columns 3 and 8. We will also refer to this specification as the benchmark model ('Benchmark').

We focus on the NC sample results first. Comparing the values across columns in the two first rows of Table 2 indicates all models fit the UTJs reasonably well. The benchmark model matches the UTJs levels of blacks and whites more closely, but overshoots the observed UTJ racial gap by 1.9 pp (17.8 vs 15.9 pp). The model with skill differences only produces a considerably lower gap (11.9 pp) and the model with discrimination only generates a considerably greater gap (18.9 pp).

Overall, the fit of wage distribution moments is not as good. 'Discrimination' and 'Benchmark' match the mean wages of blacks very well, but overshoot whites' mean wages. All specifications produce a substantially greater wage gap compared to the observed one (7 lp). 'Skills' and 'Discrimination' generate a gap of 10 lp, while the benchmark model delivers a mean wage gap of 14 lp. However, 'Benchmark' fits the second and

---

[30]We use the same value-added moments for both NC and SC samples. Alternatively, we could match workers and firms based on their 4-digit industry code and obtain a different set of moments for both sub-samples. In theory, this could be informative about different sorting patterns across the two samples. In practice, we find that using 4-digit industry codes (as opposed to 6-digit) substantially reduces firm-level dispersion.

Table 2: Fit of the moments

		A. No college					B. Some college				
		Model specification			Data		Model specification			Data	
		Discrim.	Skills	Bench.	Mean	Sd.	Discrim.	Skills	Bench.	Mean	Sd.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Transitions											
UTJ	Whites	.422	.380	.417	.402	(.011)	.408	.367	.356	.379	(.041)
	Blacks	.233	.261	.239	.243	(.036)	.151	.158	.161	.177	(.143)
JTU		.042	.041	.042	.042	(.001)	.020	.021	.021	.019	(.001)
Wages											
Mean	Whites	2.60	2.64	2.61	2.54	(.005)	2.64	2.97	2.95	2.89	(.008)
	Blacks	2.49	2.54	2.47	2.47	(.011)	2.49	2.51	2.48	2.71	(.024)
Sd.	Whites	.289	.305	.293	.299	(.003)	.435	.430	.421	.407	(.004)
	Blacks	.273	.282	.270	.273	(.009)	.421	.355	.355	.395	(.018)
Skewness	Whites	.259	.254	.254	.256	(.026)	.114	.111	.109	.106	(.024)
	Blacks	.402	.450	.414	.409	(.082)	.402	.507	.513	.651	(.110)
Kurtosis	Whites	2.23	2.20	2.22	2.42	(.041)	2.28	2.24	2.23	2.16	(.030)
	Blacks	2.43	2.33	2.44	3.00	(.169)	2.42	2.59	2.60	2.72	(.242)
Value-added											
Mean		3.83	3.63	3.85	4.08	(.041)	3.88	4.31	4.31	4.08	(.041)
Sd.		.570	.491	.574	.548	(.038)	.491	.549	.549	.548	(.038)
Tightness		.492	.493	.491	.494	(.001)	.497	.504	.507	.494	(.002)
Criterion		**.027**	**.052**	**.025**			**.118**	**.049**	**.046**		

Notes.—Model fit of the three specifications. Standard errors of the moments in the data reported in parentheses.

24

third moments of both wage distributions better. 'Discrimination' generates a standard deviation for whites wages that is too low, whereas 'Skills' substantially overpredicts the skewness of blacks' wage distribution.

In sum, 'Discrimination' and 'Benchmark' perform very similarly. Allowing for differences in skills improves the fit of the UTJs and wage skewnesses of both races, as well as the standard deviation of wages of blacks, but at the cost of a poorer fit of the racial wage gap.

Turning to the SC sample, it is noticeable that the qualitative performance of the various specifications changes somewhat. As we argue in more detail below, this is related to the differences in mean wage levels and UTJs across the two samples. 'Discrimination' still overpredicts the racial UTJ gap, but targets the mean wage gap more closely. The predicted levels of mean wages are, however, considerably below their empirical targets, the standard deviations of wages too high and the skewness of blacks' wages too low. On the other hand, 'Skills' and 'Benchmark' hit the levels of both UTJs quite well, but at the cost of overpredicting the wage gap compared to the data. This very large gap is mostly due to the poor fit of blacks' mean wages.

In conclusion, although all specifications match quite well important subsets of moments, our preferred specification comes closer to capturing simultaneously all the key aspects of the data: the difference in UTJs of whites and blacks and the moments of blacks' and whites' log-wage distributions. This information is summarized in the value of the minimization criterion at the bottom of Table 2, which is lower for the benchmark model (0.025 and 0.046) compared to 'Discrimination' (0.027 and 0.118) and 'Skills' (0.052 and 0.049), respectively for the NC and SC samples.

5.3   Structural Parameter Estimates

Having argued that our model does a good job at matching the empirical moments, we now analyze its economic content. Table 3 reports the parameter estimates of all specifications for both the NC and SC samples, with standard errors reported in parentheses.

The estimated monthly probabilities of job arrival are 69% and 66%, respectively for the NC and SC samples.[31] A comparison with the UTJs in Table 2 suggests that, for the NC group, approximately a quarter of all job offers are rejected, while a third of offers are rejected by those with some college education. The corresponding job finding rates (UTJs) are not too dissimilar, indicating that the large employment differentials between the two strata are governed by very different separation rates (δ).

In terms of the relative importance of discrimination and skill differences in explaining

---

[31]Calculated as $1 - e^{-\lambda^W}$.

Table 3: Parameter Estimates

Parameter	A. No college			B. Some college		
	Discrim (1)	Skills (2)	Bench. (3)	Discrim (4)	Skills (5)	Bench. (6)
π	.619		.565	.879		.311
	(.074)		(.112)	(.410)		(.039)
d	6.75		7.37	4.90		1.41
	(1.438)		(1.29)	(.52)		(.262)
s		-.048	-.014		-.143	-.136
		(.008)	(.001)		(.021)	(.040)
λ	1.19	1.24	1.17	1.91	1.16	1.07
	(.428)	(.076)	(.116)	(.194)	(.057)	(.030)
δ	.043	.042	.043	.020	.021	.021
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
μ	3.83	3.63	3.85	3.88	4.32	4.32
	(.104)	(.039)	(.056)	(.209)	(.184)	(.110)
σ_h	.101	.141	.101	.151	.129	.124
	(.008)	(.006)	(.007)	(.032)	(.019)	(.019)
σ_x	.568	.445	.572	.418	.500	.506
	(.039)	(.027)	(.038)	(.061)	(.031)	(.051)
b	7.02	8.07	6.99	5.71	6.89	6.68
	(1.309)	(.417)	(.722)	(1.09)	(.98)	(1.50)
G	.959	.959	.959	.979	.976	.976
	(.040)	(.004)	(.024)	(.045)	(.023)	(.028)

Notes.—Estimation by SMM of the three specifications. Asymptotic standard errors reported in parentheses.

differences in labor market outcomes, we obtain very distinct patterns in the two samples. In the NC sample, there is evidence of strong and widespread employer prejudice against black workers. The loss of utility incurred by a prejudiced employer when matching with a black worker is estimated at \$7.4 per hour. The share of prejudiced employers is estimated at 56.5%. The estimated average differences in skill across black and white individuals are very small (a mean black-white skill differential of 1.4%). On the other hand, in the SC sample we find evidence of very a substantial skill gap across races (13.6%), as well as mild employer prejudice concentrated among a small share of employers (the utility cost of employing a black worker is \$1.4 and the share of prejudiced employers is 31.1%).

The large differences across strata in terms of the racial skill gap may reflect the differences in skill heterogeneity across the two samples. For a given stratum, the distribution of output depends on the worker's race and skill, as well as firms' productivity. Conditioning on the firm being of median type, we can express the mean and variance of output as a function of μ, σ_h and s.[32] Inserting our estimated values of μ, σ_h and s in that expression, for an employed worker in the median firm, the variance of output is 4 (3) times larger for a white (black) with some college education compared to a white (black) without any college education.

5.4 Heterogeneous job-destruction rates

Up until this point we have estimated the model under the assumption that both types of workers are exposed to the same exogenous job destruction rate. In this section we relax that assumption and estimate a version of the model where job destruction arrival rates are allowed to vary across races by fitting the observed racial-specific JTUs.[33] For sake of space, we report the main tables of results in F. In terms of differences across the two samples, the cross-racial difference in JTUs is higher among individuals with some college vs no college. In the NC sample blacks' JTU is 0.072 vs. 0.039 for whites, whereas those figures are 0.044 and 0.017 in the SC sample. Overall, the estimation results are consistent with those obtained using a common job destruction rate across races: employer prejudice fits the main patterns in the NC sample, whereas skill differences match the SC sample moments.

The main effect of introducing different job destruction rates is to introduce an exogenous wage gap and a negative UTJ gap in the absence of either employer prejudice

---

[32]$f_i(h, 0.5) = \exp\{\mu + s\mathbf{1}_{[i=2]} + \sigma_h \Phi^{-1}(h)\}$. Since h is uniformly distributed, taking the variance operator through this yields $\mathrm{var}\{f_i(h, 0.5)\} = \exp\{2\mu + 2s\mathbf{1}_{[i=2]}\}.\mathrm{var}\{\sigma_h \Phi^{-1}(h)\} = \exp\{2\mu + 2s\mathbf{1}_{[i=2]} + \sigma^2\}[\exp\{\sigma^2\} - 1]$.

[33]A more complete route to address this issue would be to extend the model in order to generate endogenous racial differences in JTUs as a result of prejudice and skill differences. An extension of the model in that direction is beyond the scope of this paper.

or skill differences. Since matches with black workers have shorter duration, the present-discounted value of a match with a black worker is lower compared to an equally skilled white. This translates into lower wages for feasible matches and a lower outside option for black workers. Since the match feasibility condition (see equation (9)) only depends on job destruction rates through the outside option (and negatively), the matching sets of black workers expand relative to whites. As a result, blacks' UTJ is higher compared to whites. To summarize, the inclusion of separation rates that vary by race does not significantly change our parameter estimates and model fit.

5.5 Taking Stock

The fact that, based on the same model, we obtain very different interpretations of facts that are qualitatively similar (compared to whites, blacks have lower UTJ, average and standard deviations of wages but higher wage skewness) suggests the performance of the two polar specifications is affected by the levels of UTJs and mean wages. To generate empirically reasonable UTJ gaps, the model based only on skill differences needs to generate a substantial wage gap. On the other hand, the model based only on prejudice matches the wage and UTJ gaps well, but struggles to generate high mean wages for blacks and whites.

In what concerns the empirical estimates, given that we estimated the model in a particular sample (young workers employed in the manufacturing sector) we cannot rule out that they are specific to this segment of the labor market. The patterns of employer prejudice that we find in the NC sample are quite elevated. Audit studies (see e.g. Bertrand and Mullainathan (2004)) usually find that, in quasi-experiments, observationally-equivalent black workers receive between 15 to 30% less calls relative to whites. This could suggest that the share of prejudiced employers is lower. However, that interpretation rests on the assumption that prejudiced employers never call black workers and that all nonprejudiced employers do. This is only true for special configurations of parameter values. As our model highlights, in a more realistic environment with two-sided skill heterogeneity, some prejudiced employers hire black workers and some nonprejudiced employers do not.

A broader perspective to interpret our empirical results should take account of the fact that, in addition to evidence on taste-based discrimination, there is ample evidence suggesting that statistical discrimination is an important driver of cross-racial differences in labor market outcomes in the U.S. (see e.g. Charles and Guryan (2011)). If statistical discrimination generates similar predictions (in terms of size and direction of employment and wage gaps), our empirical strategy will overestimate the extent of employer prejudice. On the contrary, if statistical discrimination predicts opposite effects, it will underestimate

the extent of prejudice. Exploring this issue is an important avenue for future research.

6   Sorting on Race/Prejudice and Skill

Having estimated the parameters of the model in both the NC and SC samples, we can simulate two versions of the model featuring both skill differences and employer prejudice to provide an interpretation of the observed empirical patterns.[34] By doing so, a clear relationship emerges between matching decisions (captured by matching sets) and differences in labor market outcomes for workers of different races. To simulate the model we follow the methodology described in C, where the model is specified as detailed in Section 4.1.
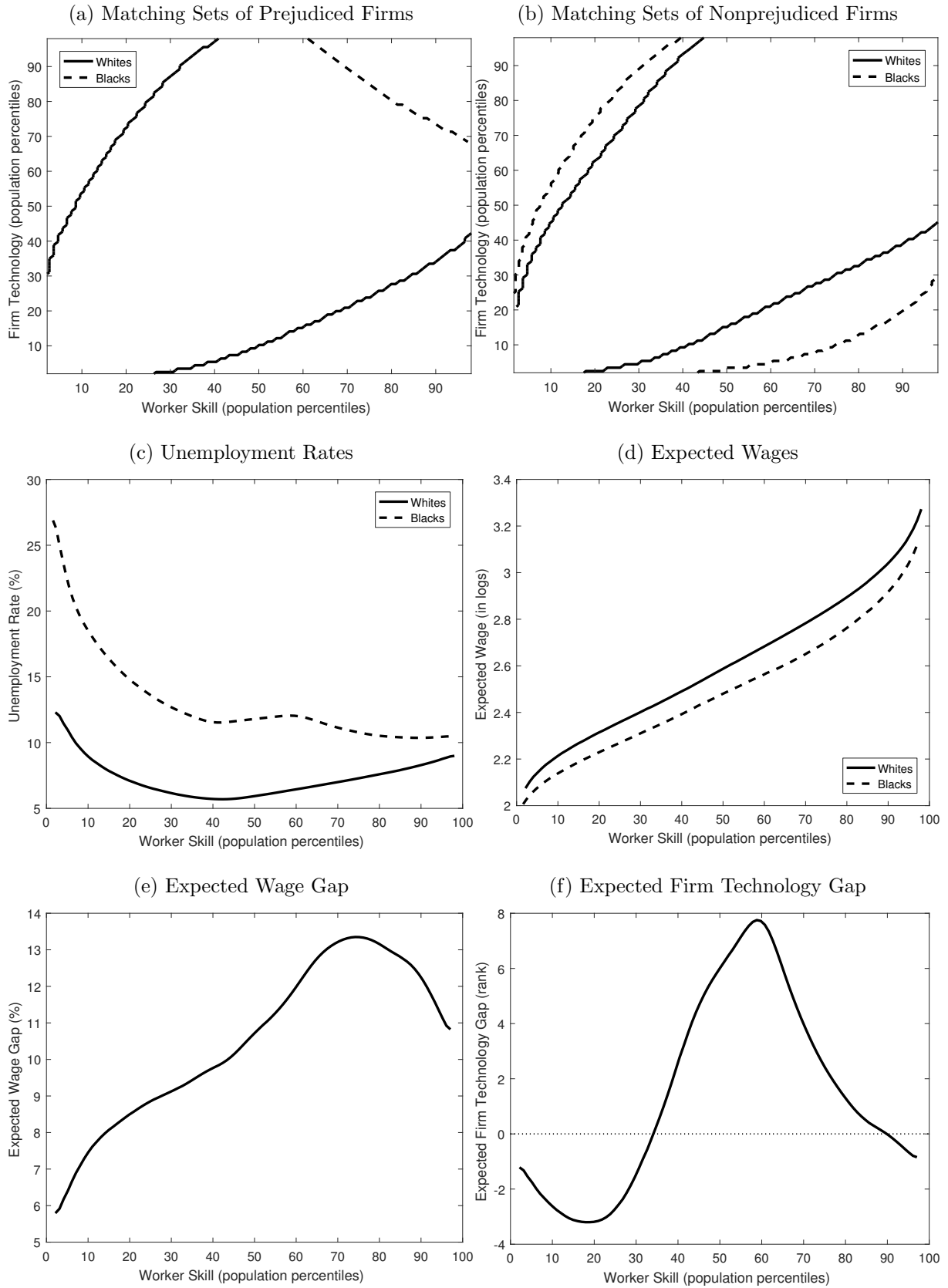
6.1   No College sample

We start by describing the patterns of the model based on the NC sample. The two plots in the top panel of Figure 1 display the equilibrium strategies of all the agents in the economy (their matching sets). Figures 1a and 1b plot the contour lines of the set of matching sets of respectively prejudiced and nonprejudiced employers with white and black workers. Workers' skill (expressed in percentiles of the population distribution) is displayed on the horizontal axis and firms' technology (expressed in percentiles of the population distribution) on the vertical axis. The solid/dashed lines indicate the bounds of the set of matching sets with white/black workers. 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 match more often 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 1a, 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 skilled black workers. This is an instance of hiring discrimination

---

[34]To simplify the description of main effects, we focus on a specification of the model with a common δ. All arguments that follow can be made, albeit with an additional moving part, with a model with separation rates dependent on race

Figure 1: Simulation of Baseline Model — No College sample

(a) Matching Sets of Prejudiced Firms

(b) Matching Sets of Nonprejudiced Firms

(c) Unemployment Rates

(d) Expected Wages

(e) Expected Wage Gap

(f) Expected Firm Technology Gap

stated in Corollary 2. 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 when 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' skill: it is much 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 3). As can be seen in Figure 1b, 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 but not with equally skilled white workers. The presence of prejudiced employers decreases blacks' opportunities in the labor market, which results in them having a lower outside option compared to equally skilled 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 skilled whites.

The differences in labor market outcomes between black and white workers of similar skill depend on the relative magnitude of the various forms of discrimination present in the model, which in turn are determined by π and d and the share of whites and blacks in the population, m and $(1 - m)$. In both Figures 1c and 1d the solid lines indicate white workers' outcomes and the dashed lines the outcomes of black workers, whereas the x-axis denotes the skill percentiles of the population of all workers. Figure 1c shows the unemployment rates of both types of workers for different levels of skill. More skilled workers experience lower unemployment rates. For both types of workers the distributions of unemployment rates across levels of skill describe a U-shaped curve: unemployment rates decrease up until percentile 40 and increase from there on for whites and remain flat for blacks. This effect follows from the shape of matching sets (wider in the middle of the skill distribution) and is a general feature of other search and matching models with sorting on skill (see Eeckhout and Kircher (2011)). A more important feature of this plot is that, whatever the level of skill of black workers, they experience higher unemployment rates compared to equally skilled white workers. This highlights what is already apparent in Figures 1a and 1b: that hiring discrimination dominates reverse hiring discrimination.

31

This pictures also highlights that the black-white unemployment rate differential is far higher for low levels of skill and becomes smaller as the level of skill increases, consistent with a disproportionate effect of the psychic cost d on low-skill black workers' matching opportunities.

Figure 1d displays the differences in expected wages for blacks and whites with different levels of skill. Expected wages are increasing in workers' skill. As can be seen more clearly in Figure 1e, the black-white expected wage differential displays a non-linear relationship in workers' skill: the gap increases up until percentile 70 and decreases thereafter. This pattern can be explained by changing relative importance of two distinct channels. The first channel is sorting by skill. Until percentile 60 (see Figure 1a) black workers are only employed by nonprejudiced employers, so the profile of racial differences in expected wages along the skill distribution in this region is not affected by direct discrimination (since d does not enter blacks' wages in this region). Hence, it must be explained by changing patterns of skill sorting among blacks and whites as workers' skill increases.

Figure 1b displays the racial gap in the average technology level of firms with whom workers are matched for different skill levels. From percentile 0 to 20, the average technology of firms with whom blacks are matched is higher (hence the negative gap) and rises faster compared to whites' as skill increases (this can be seen in the steeper slope of the dashed line vs the solid line in Figure 1b). Therefore, the only reason why the wages of whites are increasing faster than those of blacks in this region is because whites' outside options are also increasing. From percentile 20 to 60 the expected firm technology with whom whites are matched improves relative to blacks' and that drives the increase in the wage gap (see bottom lines of Figure 1b). From percentile 60, blacks start matching with high-technology prejudiced firms (see Figure 1a). As a result, the effect of direct discrimination also sets in, so that, as the skill of workers increases, the wage gap also decreases. This is because $-\beta d$ represents an ever decreasing proportion of wages (see equation (7)). In this region blacks reduce their average firm technology gap relative to whites as their skill increases, but that effect is quantitatively smaller.

6.2   Some College sample

Figure 2 displays various labor market outcomes for the model estimated on the SC sample. Since many of the points we made regarding the NC results are also present in the SC results, we focus on the main differences between them. This segment of the labor market is characterized by mild prejudice concentrated among a small share of employers, and large skill differences across workers of different races. The lower level of prejudice

translates into small differences in matching sets across races.[35] The degree of hiring discrimination is much smaller (see 2a), and there is almost no reverse hiring discrimination (see Figure 2b). Consequently, the differences in expected wages and unemployment across workers with the same level of skill are much lower (see Figures 2c and 2d). In addition, the milder differences in matching sets generate smoother patterns of wage and unemployment gaps along workers' skill. Both the unemployment rate gap and the expected wage gap decrease monotonically as the worker's skill increases (see Figures 2e and 2f).

7   Counterfactual Policy Analysis

In this section we use the estimated model to perform counterfactual experiments aimed at evaluating the effect of two policy approaches in reducing differences in labor market outcomes of blacks relative to whites. Beyond labor market outcomes, we are interested in quantifying th effects of policies on individual and social welfare. We measure individuals' welfare by their expected present discounted lifetime utility value as described by the equation below:

$$o_i(h) = u_i(h)U_i(h) + \sum_{j=\{N,P\}} \int \gamma_i^j(h,x)W_i^j(h,x)\mathrm{d}x. \qquad (19)$$

Because we focus on the effects of policies in equilibria satisfying the free-entry condition (by which employers make zero profits in expectation), we only compare the effects of policies on workers' welfare.[36] However, for descriptive purposes it is useful to define the welfare of firms by the expected present discounted lifetime utility value, as in the equation below:

$$o^j(x) = v^j(x)V^j(x) + \sum_{i=\{1,2\}} \int \gamma_i^j(h,x)J_i^j(h,x)\mathrm{d}h. \qquad (20)$$
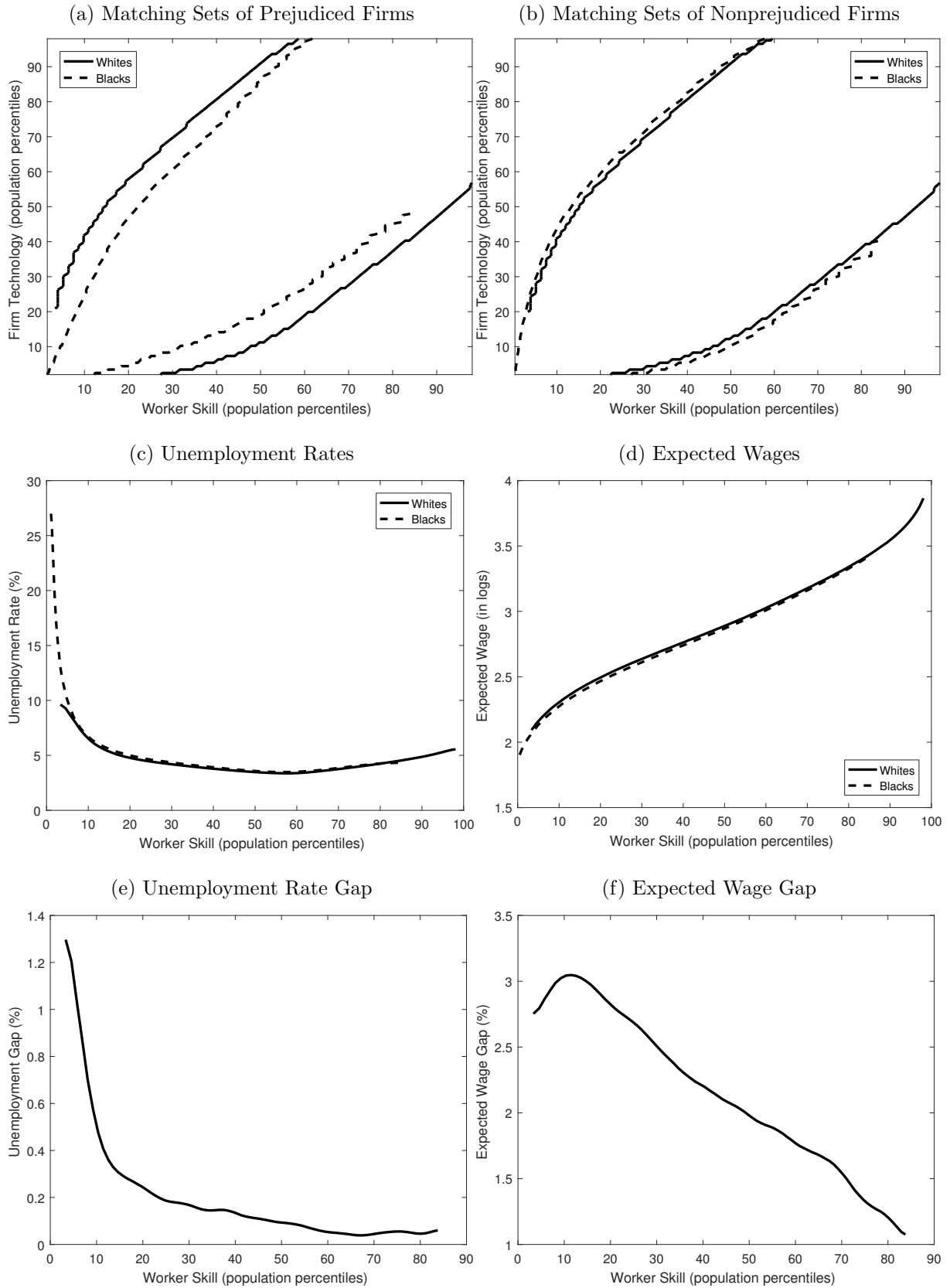
To evaluate the social desirability of different policies, we use a utilitarian social welfare function given by the sum of workers' welfare.[37] An important caveat to the conclusions

---

[35]Due to the large estimated differences in skill distributions across races, there are no blacks with very high levels of skill, nor whites with very low levels of skill. This is why the supports of the dashed and solid lines do not overlap at their extremes.

[36]In describing the effects of the policies, it is often convenient to refer to the short-run equilibrium, i.e. without satisfying the free-entry condition. For sake of space we do not report those results in the paper, but they are available from the authors on request.

[37]Our chosen social welfare function includes the losses due to prejudice. In our view, excluding them would make sense if the social cost of prejudice were fully internalized by prejudiced employers. In our baseline model that is not the case – the utility cost of prejudice is partly transferred from prejudiced employers to black workers–, so it is important to include those losses.

Figure 2: Simulation of Baseline Model — Some College sample

(a) Matching Sets of Prejudiced Firms

(b) Matching Sets of Nonprejudiced Firms

(c) Unemployment Rates

(d) Expected Wages

(e) Unemployment Rate Gap

(f) Expected Wage Gap

on the effect of the policies on social welfare is that we do not consider the costs of implementing the policies.

We consider two policies in our counterfactual analysis. The first policy is an illustration of an increase in the enforceability of the equal-pay principle (a pillar of anti-discrimination legislation). The second policy entails an increase in the premarket skills of black workers.[38]

To describe the effects of the two policies, we simulate the modified version of the model for the two sets of parameter values estimated respectively on the NC and SC samples. Table 4 reports statistics on workers' and firms' labor market outcomes, as well as on group and social welfare, for the baseline and the two counterfactual policy equilibria, for each of the two samples. The top panel displays average log-wages and unemployment rates for both worker types. The middle panel reports the mean vacancy rate of prejudiced and nonprejudiced firms, as well as the total number of active firms (G). The bottom panel provides information on social welfare, the welfare of the two worker types and the welfare accounted for by production and the utility losses due to prejudice.

7.1 Equal pay

In the context of our model, the principle of equal pay implies that equally skilled workers working in equally productive firms should be paid the same wage. In our baseline specification, we assume that workers and employers may agree wages that explicitly violate this principle. We believe this is a plausible description, insofar as worker skill and job requirements are imperfectly observed by parties external to a match, thereby making the enforcement of equal-pay legislation very difficult in practice. Against this backdrop, we consider a policy that improves the application of the equal-pay principle across workers of similar skill and different races. Specifically, we consider a setup in which the utility cost of prejudice is not transferable from prejudiced employers to black workers. Recall that, in the baseline model, when a black worker and a prejudiced firm negotiate the wage, the utility cost of prejudice enters the negotiation explicitly. Both parties recognize the presence of this cost and its negative effect on the utility value of the match, and they agree to share that loss according to their rent-sharing parameters. In this counterfactual experiment, the utility cost of prejudice is fully internalized by the employer. To specify it, we change the match feasibility condition between black workers (type-2) and

---

[38]It would be interesting to structurally evaluate the effects of minority quotas, i.e. policies that incentivize firms to hire a certain proportion of minority workers (see Miller (forthcoming) for a recent reduced-form analysis). In our model a firm is one job, so the notion of quota (a share of firms' employment) is not well defined.

prejudiced firms (type-P) from equation (9) to equation (21),

$$\alpha_i^j(h, x) = \mathbf{1}\left\{(1-\beta)[f_i(h, x) - \rho U_i(h) - \rho V^j(x)] > d\mathbf{1}_{[(i,j)=(2,P)]}\right\}. \tag{21}$$

This condition states that a prejudiced employer will only match with a black worker if its flow surplus is greater than the psychic cost d. In addition to the match feasibility condition, we also change the payoffs of black workers and prejudiced employers when they are involved in matches with each other. Now prejudiced firms fully internalize the psychic cost of prejudice and so none of it is transferred to black workers' wages.[39] It is important to note that, due to the structure of our model – one in which wages depend explicitly on outside options and search frictions are random–, full equal pay would only be enforced if all prejudiced firms were driven out of the market. Therefore, the aim of the policy is to *limit* wage discrimination against black workers. In the baseline model wage discrimination occurs via two channels: the direct effect of d on wages and its effect through the outside option. In this counterfactual the first form of wage discrimination cannot occur, but the latter remains operative. In this sense, we see this counterfactual as an illustration of the impact of policy instruments that limit the ability of prejudiced employers to pay lower wages to equally skilled black workers.

To describe the effects of enforcing equal pay we compare the values in columns (1) and (2) with those in columns (4) and (5) of Table 4, respectively for the NC and SC samples. The main direct effect of equal pay is to reduce the market returns of prejudiced employers via two channels: the number of feasible matches with black workers diminishes and the wages paid in feasible matches increase. This leads to an increase in the unemployment rate of black workers, along with an increase in their average wages. In the short run (i.e. keeping number of active firms constant), the lower outside option of prejudiced firms benefits white workers: their average unemployment rate is lower and their average wage higher. In the long run, however, the lower expected returns of firm entry lead to net firm exit, which overturns the benefits to white workers. Overall, the average unemployment rates of whites increase slightly relative to the baseline economy, while average wages decrease very slightly. The combined effect of the policy on cross-race differences in labor market outcomes is a reduction in the wage gap, but which comes at the cost of a higher unemployment gap. Put together, the latter lead to a reduction in the welfare gap across races (of less than one percentage point, from 15.12 to 14.34% in the NC sample, and from 39.76 to 39.27 in the SC sample). The policy leads to a very minor reduction in social welfare. The gains to black workers come at the cost of lower welfare to white workers, which is the result of a reduction in the number of active firms. The lower number of

---

[39]See G for a full mathematical description of this specification.

active firms leads to less matches being realized and, consequently, to lower output. This effect is partially mitigated by more efficient labor market sorting: the contraction in the set of matching sets between prejudiced firms and black workers improves the productivity of realized matches and diminishes the welfare losses due to prejudice.

While the the aggregate effects of equal pay are similar in the two samples, the distributional effects are somewhat different. In the NC sample's baseline economy, only very high-skill black workers match with nonprejudiced firms, so the direct policy effects will be concentrated on them. For these workers, the contraction in matching sets (see Figure 3a) leads to higher unemployment rates, but also much higher wages in feasible matches. This increase in expected wages leads to an increase in the outside option of high-skill black workers, implying that some matches with low-technology nonprejudiced firms are no longer feasible (see Figure 3b). As illustrated in Figures 3c and 3d, the changes in the unemployment rate and wage gaps are concentrated on the right tail of the skill distribution. Figure 3e highlights that the welfare gains to black workers are similarly concentrated on the right tail of the skill distribution. Figure 3f makes the same point by highlighting that the fall in firms lifetime values falls mostly on high-technology prejudiced firms. On the contrary, whites' welfare losses are evenly distributed along the skill distribution.

The distributional effects on the SC sample are illustrated in Figure 4. In the SC sample's baseline economy, there are feasible matches between black workers and prejudiced firms in the whole support of the skill distribution. The effect of equal pay is to contract the set of matching sets towards the 45° line. This contraction is particularly strong in the left tail of the skill distribution. This effect is so large that the wage gap increases in this part of the distribution, whereas it decreases everywhere else. Whereas almost all white workers are made worse off by the policy, blacks on the 40 percentile of the skill distribution or above benefit from the policy (those below are worse off).

7.2   Equal skill

To illustrate the impacts of policies aiming to equalize the pre-market skills of minority and majority workers, we conduct a counterfactual experiment whereby the skill distribution of black workers is set equal to that of whites. We interpret the measure of skill in our model as general one, i.e. it can relate both to cognitive or noncognitive skills. This counterfactual can speak to a wide variety of real-world policy interventions aimed at closing the racial skill gap (see Fryer Jr (2011) for a comprehensive account of such interventions).

In Table 4, Columns (3) and (6) report the outcomes of the estimated model with s set

Table 4: Counterfactual Policies Simulation Results

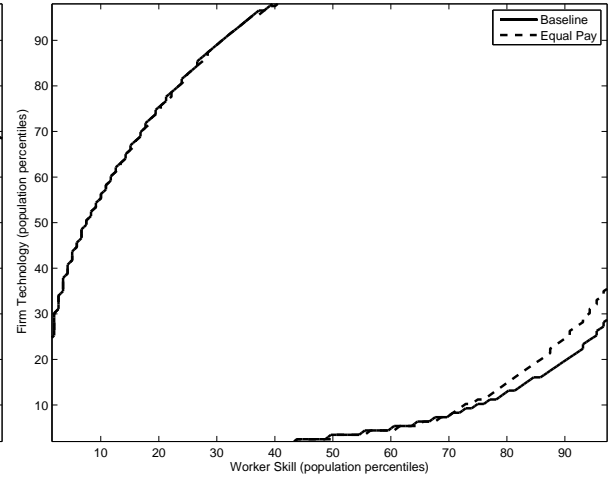| | | A. No college | | | B. Some college | | |
		Baseline (1)	Equal pay (2)	Equal skill (3)	Baseline (4)	Equal pay (5)	Equal skill (6)
Workers							
Mean wage	Whites	2.610	2.610	2.611	2.951	2.950	2.961
	Blacks	2.475	2.482	2.514	2.484	2.496	2.942
Unemployment rate (%)	Whites	7.265	7.322	7.200	4.449	4.482	4.130
	Blacks	13.37	14.16	12.80	9.81	10.92	4.291
Employers							
Vacancy rate (%)	Nonprejudiced	3.779	3.759	3.805	2.532	2.523	2.559
	Prejudiced	4.291	4.316	4.296	2.587	2.671	2.579
Measure of active firms		.9594	.9576	.9608	.9757	.9744	.9836
Welfare							
Social welfare (%)		1.000	1.000	1.005	1.000	0.999	1.041
Whites		12.14	12.13	12.15	18.55	18.52	18.68
Blacks		1.310	1.321	1.364	1.042	1.048	1.713
Black/white welfare gap (%)		15.12	14.34	11.74	39.76	39.27	1.66
Production		50.66	50.61	50.79	81.97	81.87	83.33
Utility loss due to prejudice		.0569	.0282	.0670	.0234	.0135	.0297

Notes.— 'Mean wage' is hourly and expressed in log points. 'Black/White welfare gap' is computed as one minus the ratio of the average lifetime value of blacks and whites. 'Social welfare' is expressed as a proportion of the level in the baseline economy. Unless indicated otherwise the metric of welfare is flow utility. 'Production' is the sum of flow production in all matches. 'Utility loss due to prejudice' is the obtained by multiplying d by the number of matches.
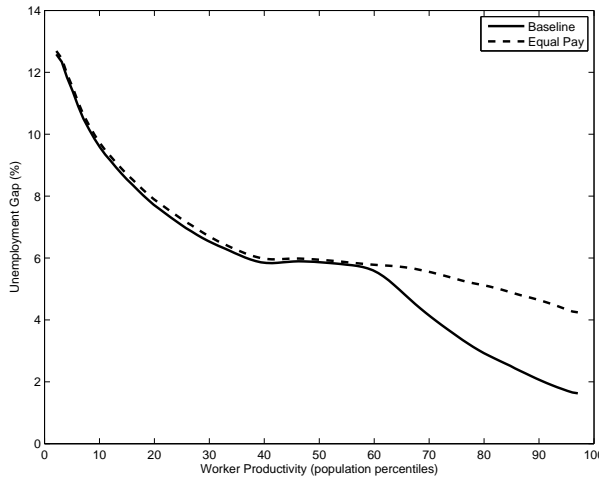
Figure 3: Equal Pay – No College sample
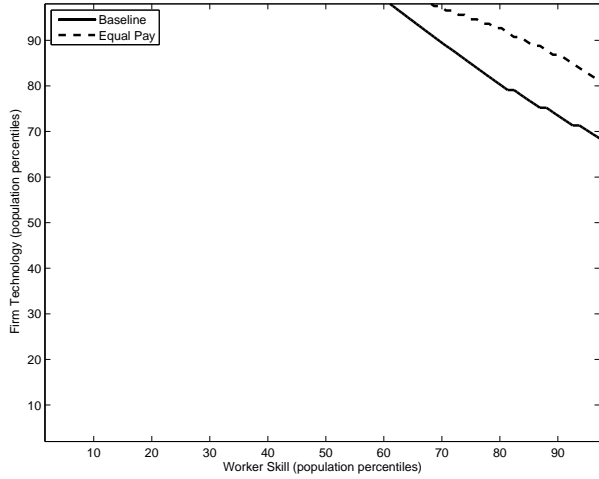
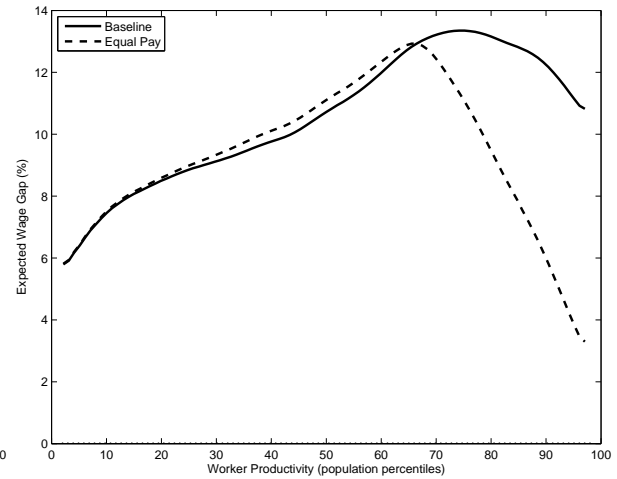(a) Matches of Prejudiced Firms and Blacks

(b) Matches of Nonprejudiced and Blacks

(c) Unemployment Rate Gap

(d) Expected Wage Gap

(e) Change in Workers Lifetime Values

(f) Change in Firms Lifetime Values

39

Figure 4: Equal Pay – Some College sample

(a) Matches of Prejudiced Firms and Blacks

(b) Matches of Nonprejudiced and Blacks
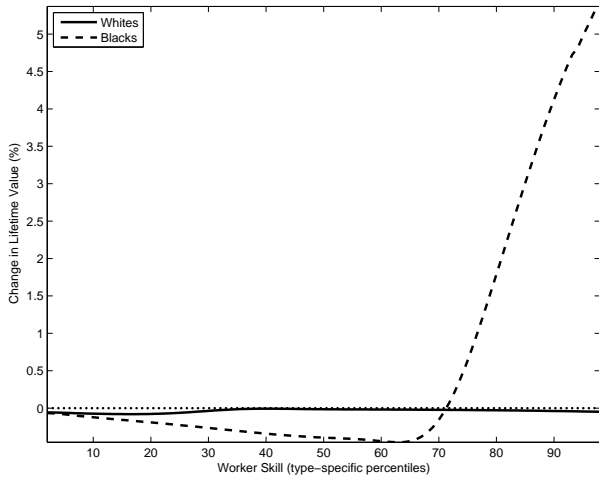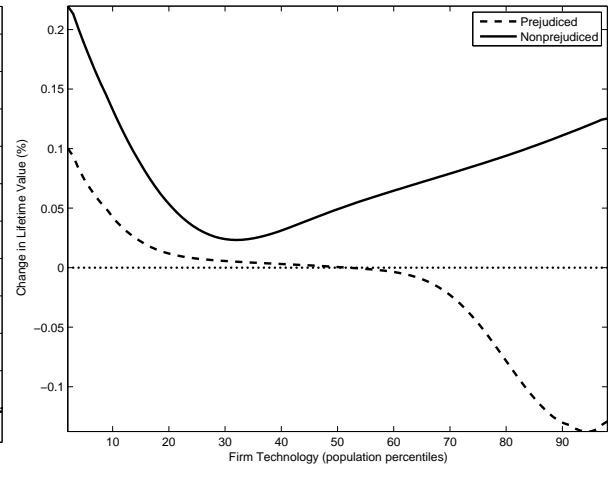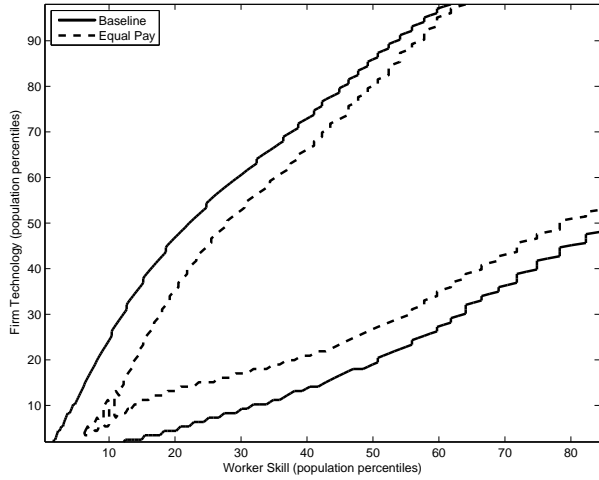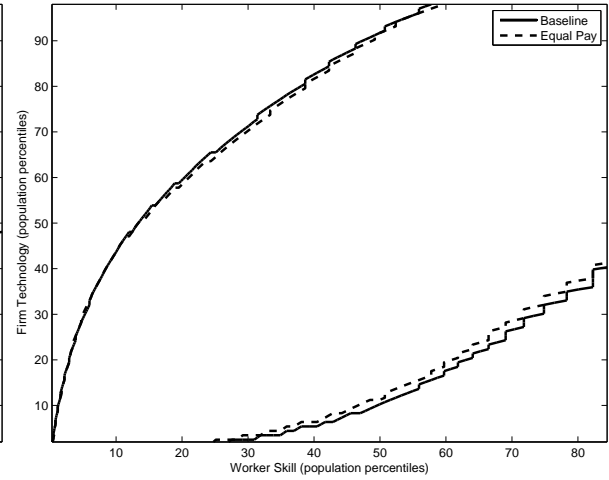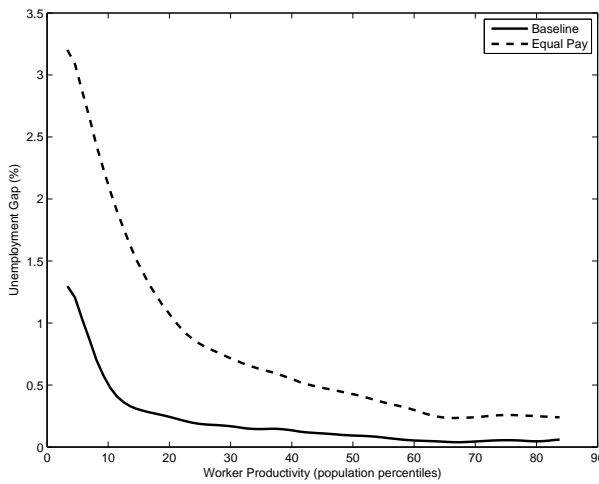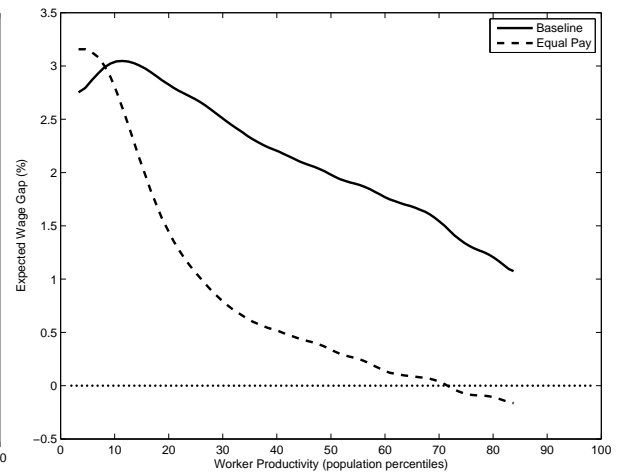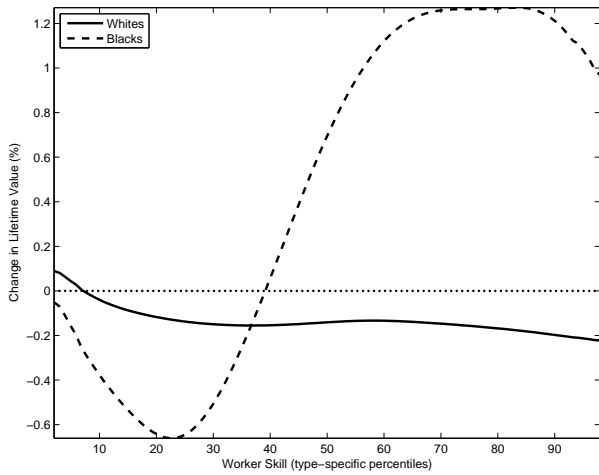

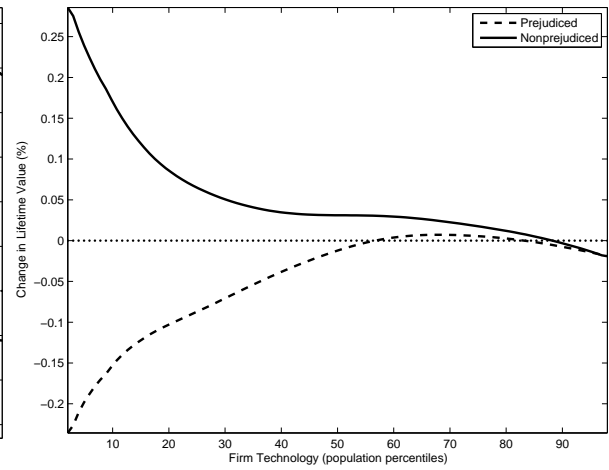(c) Unemployment Rate Gap

(d) Expected Wage Gap

(e) Change in Workers Lifetime Values

(f) Change in Firms Lifetime Values

to zero; which we compare to the values in Columns (1) and (2) pertaining to the baseline equilibrium. Black workers are now more skilled, so they match, on average, with firms with higher technology. This produces a positive effect on their wages via an outside option effect. Moreover, wages further increase as the productivity of black workers is higher. In the short run, the increase in blacks' skill entails an increase in the competition for jobs across workers, which depresses whites' matching opportunities and wages (via an outside option effect). However, since a more skilled labor force pushes firms' expected returns upwards, in the long run the mass of active firms increases, thereby overturning the negative effects on whites' wages. Social welfare is obviously greater under equal skill. This increase in driven both by a higher number of matched agents and a greater average production per match. The social losses due to prejudice are greater, but this is due to the higher number of matches between blacks and prejudiced firms.

The distributional effects of the policy are qualitatively similar, but the policy achieves a much greater reduction in the welfare gap across races in the SC sample. This is entirely justified by the much larger gap in average skills across races in the baseline SC sample vs the NC sample. Unsurprisingly, leveling the skill distributions of the two races in the SC sample entails a much larger shock to the economy. In both samples whites benefit from a more skilled labor force via increased firm entry.

8 Conclusion

In this paper we develop a search and matching model of the labor market with two-dimensional heterogeneous firms and workers to replicate key stylized facts pertaining to racial discrimination in the U.S. labor market. The model replicates facts regarding mean differences in wages and unemployment rates between white and black workers, as well as across their respective skill levels. We estimate the model using publicly available data for the U.S. manufacturing sector, and show that it captures well empirical moments retrieved from different data sources and pertaining to different aspects of the labor market.

We used the estimated model to characterize the equilibrium allocation of workers to jobs in a frictional environment in which workers differ in terms of a nonproductive attribute as well as skill, and where jobs have different technology levels and are operated by employers with different racial attitudes. The equilibrium allocation is characterized by positive assortative matching and wage and hiring discrimination, where these two channels operate with more intensity among low-skill workers compared to high-skill workers.

Our estimation results portray employer prejudice in the U.S. manufacturing sector as strong and widespread among workers without college education, and concentrated

and mild among workers with some college education or more. Given the incomplete and partial nature of our model, we are obviously cautious in taking our estimates as a definitive description of employer prejudice in the U.S. There is ample room to enrich our modeling approach along important dimensions, including trying to match other relevant empirical moments, and thereby obtain more credible estimates. In addition to including on-the-job search, it seems important to allow for endogenous job destruction, so that employer prejudice and skill differences are allowed to explain differences in separation rates across races. Another important limitation of our work is that we take the skill distribution of workers as given. In reality, skill acquisition is likely to be affected by the degree of labor market discrimination. Any of those extensions adds significant complexity to the model and the estimation protocol, and so we leave that effort for future research.

References

References

ABOWD, J., F. KRAMARZ AND D. MARGOLIS, "High wage workers and high wage firms," *Econometrica* 67 (1999), 251–333.

ALTONJI, J. G. AND L. M. SEGAL, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics* 14 (July 1996), 353–66.

ARROW, K. ET AL., "The theory of discrimination," *Discrimination in labor markets* 3 (1973), 3–33.

BAGGER, J. AND R. LENTZ, "An empirical model of wage dispersion with sorting," *mimeo* (2016).

BARTOLUCCI, C., "Gender Wage Gaps Reconsidered A Structural Approach Using Matched Employer-Employee Data," *Journal of Human Resources* 48 (2013), 998–1034.

BECKER, G. S., *The Economics of Discrimination* (University of Chicago Press, 1971).

———, "A theory of marriage: Part I," *Journal of Political economy* 81 (1973), 813–846.

BERTRAND, M. AND S. MULLAINATHAN, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *The American Economic Review* 94 (2004), pp. 991–1013.

BLACK, D. A., "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics* 13 (April 1995), 309–33.

BOLLINGER, C., "Measurement Error in the CPS: A Nonparametric Look," *Journal of Labor Economics* 16 (July 1998), 576–594.

BOROWCZYK-MARTINS, D., G. JOLIVET AND F. POSTEL-VINAY, "Accounting for endogeneity in matching function estimation," *Review of Economic Dynamics* 16 (2013), 440–451.

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

BRADLEY, J. AND A. KUEGLER, "Labor Market Reforms: An Evaluation of the Hartz Policies in Germany," *mimeo* (2017).

CAHUC, P., F. POSTEL-VINAY AND J.-M. ROBIN, "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica* 74 (2006), 323–364.

CAIN, G. G., "The economic analysis of labor market discrimination: A survey," *Handbook of labor economics* 1 (1986), 693–785.

CARNEIRO, P., J. J. HECKMAN AND D. V. MASTEROV, "Labor Market Discrimination and Racial Differences in Premarket Factors," *Journal of Law and Economics* 48 (2005), 1–39.

CAVOUNIDIS, C. AND K. LANG, "Discrimination and Worker Evaluation," *mimeo* (2016).

CHARLES, K. K. AND J. GURYAN, "Prejudice and wages: an empirical assessment of Becker's The Economics of Discrimination," *Journal of political economy* 116 (2008), 773–809.

——, "Studying Discrimination: Fundamental Challenges and Recent Progress," *Annual Review of Economics* 3 (September 2011), 479–511.

DAVIS, S. J., R. J. FABERMAN AND J. C. HALTIWANGER, "The establishment-level behavior of vacancies and hiring," *The Quarterly Journal of Economics* 128 (2013), 581–622.

DECREUSE, B. AND L. TARASONIS, "Estimating a Search Equilibrium Model of Statistical Discrimination: Racial Wage and Employment Disparities in the US," *mimeo* (2016).

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

EDERINGTON, J. AND J. SANDFORD, "Employer discrimination and market structure: Does more concentration mean more discrimination?," *International Journal of Industrial Organization* 48 (2016), 1–33.

EECKHOUT, J. AND P. KIRCHER, "Identifying Sorting — In Theory," *The Review of Economic Studies* 78 (2011), 872.

FLABBI, L., "Gender Discrimination Estimation In A Search Model With Matching And Bargaining," *International Economic Review* 51 (August 2010), 745–783.

FLINN, C., "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates," *Econometrica* 74 (2006), 1013–1062.

FRYER, R. G., D. PAGER AND J. L. SPENKUCH, "Racial disparities in job finding and offered wages," *The Journal of Law and Economics* 56 (2013), 633–689.

FRYER JR, R. G., "Racial Inequality in the 21st Century: The Declining Significance of Discrimination," *Handbook of Labor Economics* 4 (2011), 855–971.

HAGEDORN, M., T. H. LAW AND I. MANOVSKII, "Identifying equilibrium models of labor market sorting," *Econometrica* 85 (2017), 29–65.

HOPENHAYN, H., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica* 60 (1992), 1127–50.

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

LANG, K. AND M. MANOVE, "Education and Labor Market Discrimination," *American Economic Review* 101 (2011), 1467–96.

LISE, J., C. MEGHIR AND J.-M. ROBIN, "Matching, sorting and wages," *Review of Economic Dynamics* 19 (2016), 63–87.

LOPES DE MELO, R., "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy* (forthcoming).

MILLER, C., "The Persistent Effect of Temporary Affirmative Action," *American Economic Journal: Applied Economics* (forthcoming).

NEAL, D., "Why has black–white skill convergence stopped?," *Handbook of the Economics of Education* 1 (2006), 511–576.

NEAL, D. A. AND W. R. JOHNSON, "The Role of Premarket Factors in Black-White Wage Differences," *The Journal of Political Economy* 104 (1996), 869–895.

PETRONGOLO, B. AND C. A. PISSARIDES, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* 39 (2001), 390–431.

PHELPS, E. S., "The Statistical Theory of Racism and Sexism," *The American Economic Review* 62 (1972), pp. 659–661.

RITTER, J. A. AND L. J. TAYLOR, "Racial disparity in unemployment," *The Review of Economics and Statistics* 93 (2011), 30–42.

ROSÉN, A., "Search, Bargaining, and Employer Discrimination," *Journal of Labor Economics* 21 (October 2003), 807–830.

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

SHIMER, R. AND L. SMITH, "Assortative matching and search," *Econometrica* 68 (2000), 343–369.

Appendix

A   Proofs of theoretical results

In all the proofs presented in this section we assume the production function $f_i(h, x)$ 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 $f_i(h, x)$ is nonnegative, symmetric, continuous, and twice differentiable, with uniformly bounded first partial derivatives on $[0, 1]^2$. The production function $f_i(h, x)$ is strictly supermodular.*

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

Part 1 of Lemma 1 in Shimer and Smith (2000) states that, for any worker of any skill 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.[40] In particular, it

---

[40]The proof is available from the authors upon request.

implies that, for any type-1 worker with skill h:

$$\rho U_1(h) \geq b + \lambda^W \sum_{j=P,N} \int \alpha_2^j(h,x) \left[W_1^j(h,x) - U_1(h)\right] \frac{v^j(x)}{v^P + v^N} \, dx. \qquad (22)$$

Subtracting $\rho U_2(h)$ to both sides of this inequality, substituting in the bargaining solution (equation (8)) and rearranging one obtains the following inequality:

$$U_1(h) - U_2(h) \geq \frac{d}{\rho} \times \frac{\frac{\lambda^W \beta}{\rho+\delta} \int \alpha_2^P(h,x) \frac{v^P(x)}{v^P+v^N} \, dx}{1 + \frac{\lambda^W \beta}{\rho+\delta} \sum_{j=P,N} \int \alpha_2^j(h,x) \frac{v^j(x)}{v^P+v^N} \, dx}. \qquad (23)$$

If $\pi \in (0,1)$ then all workers face a positive probability of meeting a prejudiced firm due to random matching frictions, i.e. $\frac{v^P(x)}{v^P+v^N} > 0$, $\forall x$. 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 :

$$\begin{aligned}
w_2^P(h,x) =& \beta \left[f_i(h,x) - d - \rho V^P(x)\right] + (1-\beta)\rho U_2(h) < \\
& \beta \left[f_i(h,x) - \rho V^P(x)\right] + (1-\beta)\rho U_2(h) < \\
& \beta \left[f_i(h,x) - \rho V^P(x)\right] + (1-\beta)\rho U_1(h) = w_1^P(h,x).
\end{aligned} \qquad (24)$$

We have proven that $w_1^P(h,x) > w_2^P(h,x)$, $\forall(h,x)$. *Mutatis mutandis*, one can prove that $w_1^N(h,x) > w_2^N(h,x)$, $\forall(h,x)$.

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

Using equation (8) one can write $S_1^P(h,x)$ as a function of $S_2^P(h,x)$:

$$S_1^P(h,x) = S_2^P(h,x) + \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta}.$$

First, proving $d + \rho(U_2(h) - U_1(h)) > 0$ will imply $S_1^P(h,x) > S_2^P(h,x)$, $\forall(h,x)$. 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,

$$U_1(h) - U_2(h) \leq \frac{d}{\rho} \times \frac{\frac{\lambda^W \beta}{\rho+\delta} \int \alpha_1^P(h,x) \frac{v^P(x)}{v^P+v^N} \, dx}{1 + \frac{\lambda^W \beta}{\rho+\delta} \sum_{j=P,N} \int \alpha_1^j(h,x) \frac{v^j(x)}{v^P+v^N} \, dx}. \qquad (25)$$

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. $S_2^P(h, x) \leq 0$. She will suffer hiring discrimination by prejudiced firms for all combinations of (h, x) such that $0 < S_1^P(h, x) < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta}$.

Recall that we are characterizing equilibria where some but not all matches of every kind are feasible, which implies $\exists(h, x) : S_1^P(h, x) > 0$ and $\exists(h', x') : S_1^P(h', x') \leq 0$. 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.[41] In particular, it implies that $S_i^j(h, x)$ is continuous with respect to both h and x. Hence, by the intermediate value theorem, $\exists(h'', x'') : 0 < S_1^P(h'', x'') = \epsilon < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta}$.

We have shown that there exists at least one combination of (h'', x'') between a type-1 worker and a prejudiced firm such that $0 < S_1^P(h'', x'') < \frac{d + \rho(U_2(h) - U_1(h))}{\rho + \delta}$. Since the support of workers' skill is the same, $S_2^P(h'', x'')$ is well-defined and we know that $S_2^P(h'', x'') \leq 0$, i.e. a type-2 worker of the same skill level suffers hiring discrimination by the very same firm.

Proof of Corollary 3. **Hiring Discrimination by Nonprejudiced Firms**

$S_2^N(h, x)$ can be expressed as a function of $S_1^N(h, x)$:

$$S_2^N(h, x) = S_1^N(h, x) + \frac{\rho(U_1(h) - U_2(h))}{\rho + \delta}.$$

Proposition 1 implies $S_2^N(h, x) > S_1^N(h, x)$, $\forall(h, x)$. The rest of the proof is analogous to the proof of Corollary 2.

B Equilibrium Definition

Definition 3 (EQUILIBRIUM): *Given exogenous parameters L, m, d, π, ρ, β, b, κ, δ, the production function $f_i(h, x)$, a meeting 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 vector*

$$(\alpha_i^j(h, x), u_i(h), v^j(x), U_i(h), V^j(x), G)$$

that solves the system of equations composed of the value functions of unmatched agents (equations (26) and (27)), the free-entry condition (equation (5)), the measures of unmatched agents (equations (28) and (29)) and the matching indicator functions of all the agents participating in the economy $\alpha_i^j(h, x)$ (equation (9)):

---

[41]The proof is available from the authors upon request.

$$U_i(h) = \frac{b + \frac{\lambda^W \beta}{\rho+\delta} \sum_{j=\{N,P\}} \int \alpha_i^j(h,x) \left[f_i(h,x) - d\mathbf{1}_{[(i,j)=(2,P)]} - \rho V^j(x) \right] \frac{v^j(x)}{v^P+v^N}\, \mathrm{d}x}{\rho + \frac{\rho\lambda^W \beta}{\rho+\delta} \sum_{j=\{N,P\}} \int \alpha_i^j(h,x) \frac{v^j(x)}{v^P+v^N}\, \mathrm{d}x}, \quad (26)$$

$$V^j(x) = \frac{\frac{\lambda^F(1-\beta)}{\rho+\delta} \sum_{i=\{1,2\}} \int \alpha_i^j(h,x) \left[f_i(h,x) - d\mathbf{1}_{[(i,j)=(2,P)]} - \rho U_i(h) \right] \frac{u_i(h)}{u_1+u_2}\, \mathrm{d}h}{\rho + \frac{\rho\lambda^F(1-\beta)}{\rho+\delta} \sum_{i=\{1,2\}} \int \alpha_i^j(h,x) \frac{u_i(h)}{u_1+u_2}\, \mathrm{d}h}, \quad (27)$$

$$u_i(h) = \frac{l_i(h)}{1 + \frac{\lambda^W}{\delta} \sum_{j=\{N,P\}} \int \alpha_i^j(h,x) \frac{v^j(x)}{v^P+v^N}\, \mathrm{d}x}, \quad (28)$$

and

$$v^j(x) = \frac{g^j(x)}{1 + \frac{\lambda^F}{\delta} \sum_{i=\{1,2\}} \int \alpha_i^j(h,x) \frac{u_i(h)}{u_1+u_2}\, \mathrm{d}h}. \quad (29)$$

C Simulation Algorithm

The computational burden of our estimation protocol is considerably reduced by treating G as an exogenous parameter and backing out κ after the estimation is completed. Recall that, conditional on values for all other parameters, equation (5) provides a one-to-one mapping between κ and G. Using this this approach spares us from solving the free-entry condition every time we simulate the model. Hence, to estimate the model we implement the following simulation algorithm:

1. We set initial values for all the equilibrium objects (values, densities, matching sets and the total number of firms in the economy) and parameters of the model. We discretize the supports and measures of the skill distributions of workers and firms in the economy.

2. Using the initial values we iterate over equations (26) and (27) to determine $U_i(h)$ and $V^j(x)$, at each stage updating the region of feasible matches determined by equation (9).

3. New values for $u_i(h)$ and $v^j(x)$ are obtained from equations (13) and (14).

4. Given the values determined in step 3, new values of u_i, v^j and λ^W are determined.

5. Steps 2 through 4 are updated until the endogenous distributions $u_i(h)$ and $v^j(x)$ converge.

6. Once the model converges we evaluate the left-hand side of equation (5) to determine the value of κ.

D   Asymptotic Standard Errors

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

The solution $\hat{\boldsymbol{\theta}}$ can be solved by differentiating the criterion above, giving the first order condition.

$$\boldsymbol{M}^S(\hat{\boldsymbol{\theta}})' \boldsymbol{\Omega}^{-1} \left(\hat{\boldsymbol{m}}_N - \boldsymbol{m}^S(\hat{\boldsymbol{\theta}}) \right) = \boldsymbol{0}$$

$\boldsymbol{M}^S(\hat{\boldsymbol{\theta}})$ is a Jacobian matrix of moment conditions with respect to the parameters.

$$\boldsymbol{M}^S(\boldsymbol{\theta}) = \frac{\partial \left(\hat{\boldsymbol{m}}_N - \boldsymbol{m}^S(\boldsymbol{\theta}) \right)}{\partial \boldsymbol{\theta}'} = -\frac{\partial \boldsymbol{m}^S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$$

Taking a Taylor series approximation of the score condition, where $\boldsymbol{\theta_0}$ is the true value of the parameters, $\tilde{\boldsymbol{\theta}}$ is some point between the estimate $\hat{\boldsymbol{\theta}}$ and the true value and for ease of notation $\boldsymbol{M} \equiv \boldsymbol{M}^S(\boldsymbol{\theta_0})$. After, the Taylor series approximation, the subsequent lines in the expression below impute the score condition and scale by a factor $\sqrt{N}$, respectively. Also, note the N subscript has been dropped from the moment conditions, this implies population, rather than sample moments.

$$
\begin{aligned}
\boldsymbol{0} =& \boldsymbol{M}^S(\hat{\boldsymbol{\theta}})' \boldsymbol{\Omega}^{-1} \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\hat{\boldsymbol{\theta}}) \right) \\
\approx& \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \left(\boldsymbol{m}^S(\boldsymbol{\theta_0}) - \hat{\boldsymbol{m}} \right) + \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \frac{\partial \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\tilde{\boldsymbol{\theta}}) \right)}{\partial \boldsymbol{\theta}'} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0} \right) \\
\approx& \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \left(\boldsymbol{m}^S(\boldsymbol{\theta_0}) - \hat{\boldsymbol{m}} \right) + \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0} \right) \\
\sqrt{N} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0} \right) \approx& - \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \left[\sqrt{N} \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\boldsymbol{\theta_0}) \right) \right]
\end{aligned}
$$

Assuming normality of every element in $\hat{\boldsymbol{m}}$, the above expansion shows that the parameters in $\boldsymbol{\theta}$ will also be normally distributed. Given that all moments are means, subject to some weak regularity conditions, the central limit theorem guarantees normality. Thus to compute the asymptotic variance one needs to take the variance operator through the above expression.

$$
\begin{aligned}
V \left[\sqrt{N} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0} \right) \right] =& \\
=& \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \boldsymbol{M}' \boldsymbol{\Omega}^{-1} V \left[\sqrt{N} \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\boldsymbol{\theta_0}) \right), \sqrt{N} \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\boldsymbol{\theta_0}) \right)' \right] \\
& \times \boldsymbol{\Omega}'^{-1} \boldsymbol{M} \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \\
=& \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \boldsymbol{M}' \boldsymbol{\Omega}^{-1} V \left[\sqrt{N} \hat{\boldsymbol{m}}, \sqrt{N} \hat{\boldsymbol{m}} \right] \boldsymbol{\Omega}'^{-1} \boldsymbol{M} \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \\
=& \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1} \boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{S} \boldsymbol{\Omega}'^{-1} \boldsymbol{M} \left(\boldsymbol{M}' \boldsymbol{\Omega}^{-1} \boldsymbol{M} \right)^{-1}
\end{aligned}
$$

Since the weighting matrix is specified to approximate the asymptotic standard errors

one needs to approximate the Jacobian $\boldsymbol{M}$ and the variance-covariance matrix of the empirical moments $\boldsymbol{S}$. $\boldsymbol{M}$ is approximated using two sided finite differences and $\boldsymbol{S}$ is approximated by bootstrapping the empirical moments, resampling the data 200 times. A slight complication to the bootstrap procedure is that our data comes from three different sources. We abstract away from this problem by bootstrapping each data source assuming independence. Clearly some of the empirical moments coming from different sources will depend on one another, for example the job finding rates and the market tightness. However, we find imputing arbitrary covariances into the matrix $\boldsymbol{S}$ has little implication on the size of the standard errors.

E   Separate Identification of Prejudice Parameters

In this section we show evidence to argue that, in our estimation, π (the proportion of prejudice firms) and d (the level of prejudice) are separately identified. To demonstrate this, we fix all elements of $\boldsymbol{\theta}$ to the values that minimize the criterion in Table 5.3 with the exception of π and d. With a slight abuse of notation, we can define the criterion as a two-dimensional object, $L(\pi, d)$ (viz. the other elements of $\boldsymbol{\theta}$ are fixed to their optimal values). Panel (a) in figure 5 shows a contour plot of the criterion in the π-d space. The criterion appears well behaved and seems to have a global minimum at the values of parameter estimates.

To better see that the criterion has a global minimum, we take the (lower) envelope of the criterion $L(\pi, d^\star(\pi))$. $d^\star(\pi)$ is defined as a solution to the first-order condition $L_d(\pi, d) = 0$. In practice, $d^\star(\pi)$ is found as the level of d for a given π that minimizes the criterion. Panel (b) plots $L(\pi)$ against π. One can clearly see that the function is strictly concave.

F   Results with Heterogeneous Job-destruction Rates

G   Counterfactuals

G.1   Model specification

In a model with a nontransferable cost of prejudice, the expressions for the value functions remain unaltered and are given by equations (1), (2), (3) and (4). However, as the cost of prejudiced is no longer transferable from the employer to the worker, it is not present in the match surplus expression or in the wage equation, which are now written as follows:

No College

(a) Contour of Criterion

(b) Criterion against Pi

Some College

(c) Contour of Criterion

(d) Criterion against Pi
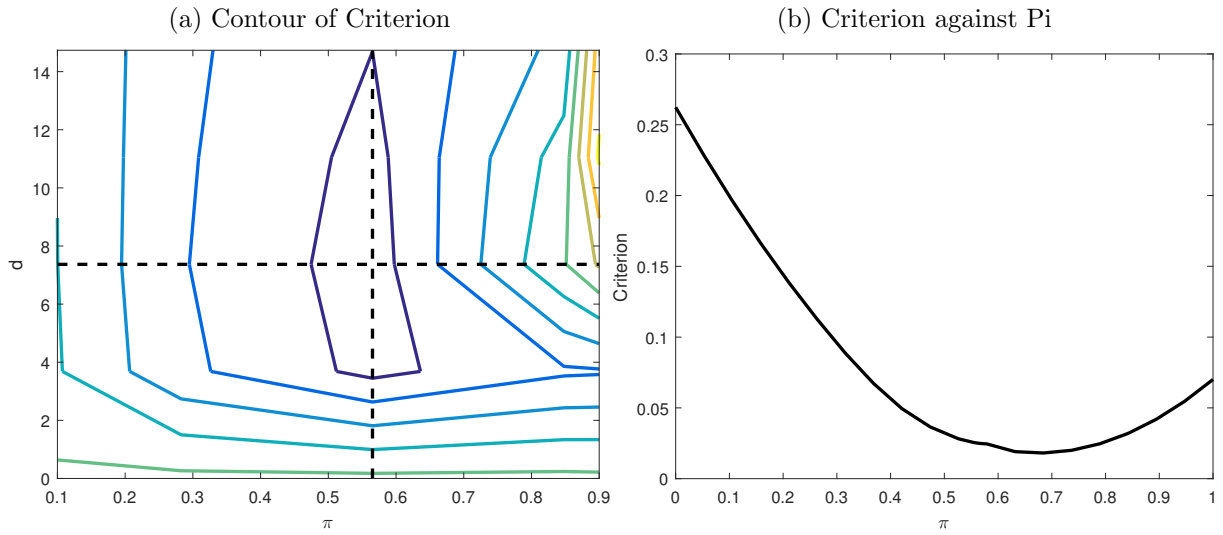
Figure 5: Separate Identification of π and d

Table 5: Fit of the moments

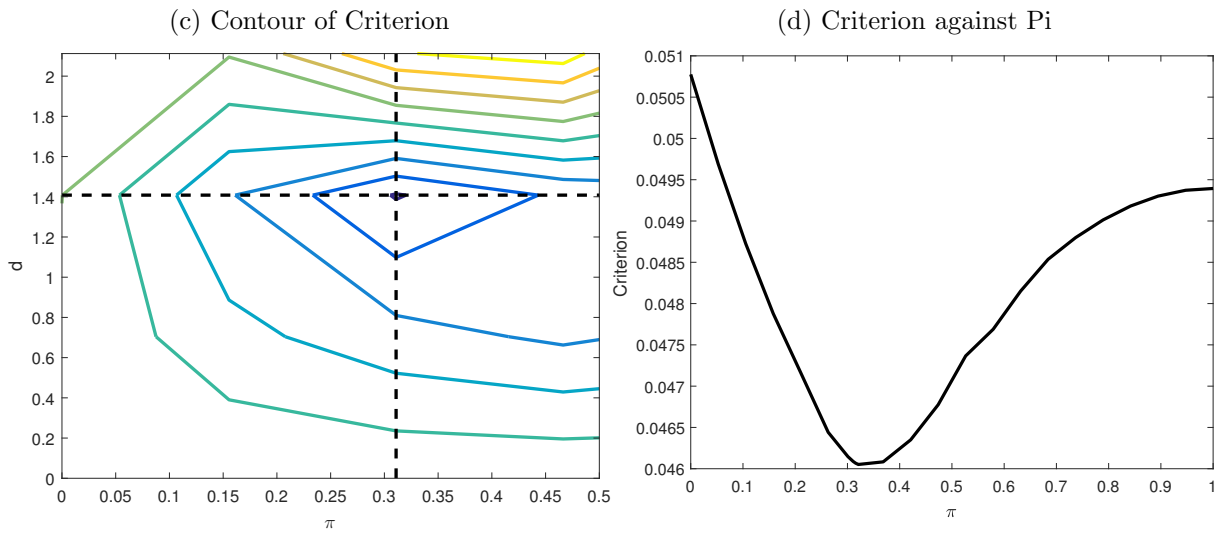| | | A. No college | | | | | | B. Some college | | | | | |
| | | Model specification | | | | Data | | Specification | | | | Data | |
| | | Deltas | Discrim. | Skills | Bench. | Mean | Sd. | Deltas | Discrim. | Skills | Bench. | Mean | Sd. |
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transitions													
UTJ	Whites	.253	.414	.371	.414	.402	(.011)	.173	.356	.371	.376	.379	(.041)
	Blacks	.279	.243	.276	.242	.243	(.036)	.200	.162	.169	.167	.177	(.143)
JTU	Whites	.041	.039	.040	.039	.039	(.001)	.018	.018	.018	.017	.017	(.001)
	Blacks	.074	.071	.069	.071	.072	(.008)	.046	.047	.045	.045	.044	(.008)
Wages													
Mean	Whites	2.57	2.64	2.70	2.64	2.54	(.005)	2.81	2.79	3.01	3.03	2.89	(.008)
	Blacks	2.51	2.45	2.54	2.45	2.47	(.011)	2.71	2.52	2.50	2.46	2.71	(.024)
Sd.	Whites	.291	.305	.322	.306	.299	(.003)	.414	.438	.445	.428	.407	(.004)
	Blacks	.259	.259	.272	.258	.273	(.009)	.369	.389	.338	.354	.395	(.018)
Skewness	Whites	.282	.251	.238	.251	.256	(.026)	.120	.110	.108	.107	.106	(.024)
	Blacks	.316	.417	.482	.415	.409	(.082)	.125	.482	.558	.582	.651	(.110)
Kurtosis	Whites	2.22	2.23	2.19	2.23	2.42	(.041)	2.11	2.26	2.26	2.26	2.16	(.030)
	Blacks	2.33	2.53	2.35	2.54	3.00	(.169)	2.15	2.45	2.57	2.55	2.72	(.242)
Value-added													
Mean		3.63	3.87	3.68	3.87	4.08	(.041)	4.04	4.02	4.24	4.32	4.08	(.041)
Sd.		.620	.576	.493	.576	.548	(.038)	.539	.500	.542	.540	.548	(.038)
Tightness		.485	.490	.497	.491	.494	(.001)	.507	.490	.504	.505	.494	(.002)
Criterion		**.157**	**.021**	**.075**	**.021**			**.519**	**.058**	**.035**	**.025**		

Notes.—Model fit of the three specifications.

52

Table 6: Parameter Estimates

Parameter	A. No college				B. Some college			
	Deltas (1)	Discrim. (2)	Skills (3)	Bench. (4)	Deltas (5)	Discrim. (6)	Skills (7)	Bench. (8)
π		.646		.648		.899		.678
		(.147)		(.067)		(.060)		(.255)
d		8.96		9.24		8.27		3.23
		(1.92)		(1.48)		(1.32)		(1.64)
s			-.068	.000			-.171	-.137
			(.038)	(.024)			(.020)	(.004)
λ	.613	1.20	1.23	1.20	.422	1.44	1.26	1.21
	(.061)	(.070)	(.111)	(.047)	(.460)	(.431)	(.226)	(.200)
μ	3.62	3.87	3.68	3.87	4.03	4.02	4.25	4.33
	(.062)	(.166)	(.061)	(.089)	(.306)	(.107)	(.096)	(.275)
σ_h	.131	.106	.151	.106	.160	.153	.149	.134
	(.012)	(.019)	(.015)	(.008)	(.144)	(.019)	(.012)	(.034)
σ_x	.613	.572	.441	.572	.494	.432	.474	.485
	(.037)	(.029)	(.039)	(.039)	(.019)	(.023)	(.038)	(.036)
b	7.83	6.93	8.27	6.93	7.44	6.32	7.29	6.72
	(.694)	(.708)	(.909)	(.746)	(1.788)	(.426)	(.465)	(1.593)
G	.932	.957	.957	.957	.954	.974	.976	.976
	(.064)	(.008)	(.021)	(.053)	(.146)	(.153)	(.030)	(.078)
δ_1	.041	.040	.040	.039	.018	.018	.018	.018
	(.002)	(.001)	(.001)	(.001)	(.002)	(.001)	(.001)	(.001)
δ_2	.076	.074	.071	.074	.047	.048	.046	.046
	(.008)	(.007)	(.007)	(.008)	(.006)	(.008)	(.007)	(.007)

Notes.—Estimation by SMM of the three specifications.

53

$$S_i^j(h, x) = \frac{f_i(h, x) - \rho U_i(h) - \rho V^j(x)}{\rho + \delta}, \tag{30}$$

$$w_i^j(h, x) = \beta \left[f(h, x) - \rho V^j(x) \right] + (1 - \beta) \rho U_i(h). \tag{31}$$

Because prejudice employers have to incur the total cost of prejudice when they match with a black worker, a prejudiced employer will only match with a black worker if his flow surplus in that match is greater than the psychic cost d:[42]

$$\alpha_i^j(h, x) = \mathbf{1} \left\{ (1 - \beta)[f_i(h, x) - \rho U_i(h) - \rho V^j(x)] > d\mathbf{1}_{[(i,j)=(2,P)]} \right\}. \tag{32}$$

Finally, in equilibrium, the reservation values of prejudiced employers and black workers are given by the following equations:

$$V^P(x) = \frac{-\kappa + \frac{\lambda^F(1-\beta)}{\rho+\delta} \sum_{i=\{1,2\}} \int \alpha_i^j(h, x) \left[f_i(h, x) - \frac{d}{1-\beta} \mathbf{1}_{[i=2]} - \rho U_i(h) \right] \frac{u_i(h)}{u_1+u_2} \, \mathrm{d}h}{\rho + \frac{\rho\lambda^F(1-\beta)}{\rho+\delta} \sum_{i=\{1,2\}} \int \alpha_i^j(h, x) \frac{u_i(h)}{u_1+u_2} \, \mathrm{d}h}, \tag{33}$$

$$U_2(h) = \frac{b + \frac{\lambda^W \beta}{\rho+\delta} \sum_{j=\{N,P\}} \int \alpha_2^j(h, x) \left[f_i(h, x) - \rho V^j(x) \right] \frac{v^j(x)}{v^P+v^N} \, \mathrm{d}x}{\rho + \frac{\rho\lambda^W \beta}{\rho+\delta} \sum_{j=\{N,P\}} \int \alpha_2^j(h, x) \frac{v^j(x)}{v^P+v^N} \, \mathrm{d}x}. \tag{34}$$

---

[42]Note that this is the relevant match feasibility condition since $(1 - \beta)S_2^P(h, x) > d \Rightarrow S_2^P(h, x) > d/(1 - \beta) > 0.$ for $d > 0$.