Gender Discrimination Estimation in a Search Model with Matching and Bargaining

Luca Flabbi

September 2005

DISCUSSION PAPER SERIES

Gender Discrimination Estimation in a Search Model with Matching and Bargaining

Luca Flabbi
*Georgetown University
and IZA Bonn*

ABSTRACT

Gender Discrimination Estimation in a Search Model with Matching and Bargaining[*]

Gender wage differentials, conditional on observed productivity characteristics, have been considered a possible indication of prejudice against women in the labor market. However, there is no conclusive evidence on whether these differentials are due to labor market discrimination or to unobserved productivity differences. The objective of this paper is to propose a solution for this identification problem by developing and estimating a search model of the labor market with matching, bargaining and employers' taste discrimination. In equilibrium all types of employers wage discriminate women: prejudiced employers because of preference and unprejudiced employers because of spillover effects that worsen the bargaining position of women. Estimation is performed by maximum likelihood on Current Population Survey data for the year 1995. Results indicate that the productivity of women is 6.5% lower than the productivity of men and that about half of the employers are prejudiced against women. Three policy experiments are implemented using the estimated parameters: an equal pay policy, an affirmative action policy and a wage differential decomposition that takes into account equilibrium effects.

JEL Classification: C51, J7, J64

Keywords: gender differentials, discrimination, search models, maximum likelihood estimation, structural estimation, affirmative action

Corresponding author:

Luca Flabbi
Dept. of Economics
580 ICC
37th & O Streets, NW
Washington, DC 20057
USA
Email: lf74@georgetown.edu

---

1   Introduction

Widespread gender wage differentials,[1] persistent even after conditioning on observable productivity characteristics, have been considered a possible indication of prejudice against women in the labor market. The problem with this interpretation is that the observables used to proxy productivity - such as human capital characteristics, controls for industry and occupation, family and community background variables - are not an accurate description of actual productivity. In general, only half of the wage variability is explained by these wage regressions[2] and it is quite plausible that an unavoidable gap will always remain between the variables that are systematically observed by the researcher and the variables that are actually considered by an employer in hiring and promotion decisions. As a result, there is no conclusive evidence on whether gender wage differentials are due to labor market discrimination or to unobserved productivity differences.[3]

The objective of this paper is to propose a solution for the identification problem by developing a search model of the labor market with employer taste discrimination and then to determine how much of the observed wage differential is due to unobserved productivity and how much can be imputed to prejudiced behavior by estimating the model on standard Current Population Survey (CPS) data.

The model is characterized by four types of agents: two types of workers (male and female), and two types of employers (prejudiced and unprejudiced). Workers search for jobs and employers post vacancies. Upon meeting, they observe a match-specific value of productivity and engage in bargaining to determine wages. Following Becker's model of taste discrimination, prejudiced employers receive disutility from employing women. Matching and bargaining generate spillover effects, a crucial channel of the transmission of prejudiced behavior on labor market outcomes that has been neglected by the previous

---

[1]For surveys of this evidence and more see: Eckstein and Nagypal 2004, Blau and Kahn 2003, Blau and Kahn 2000, Altonji and Blank 1999, a special 1998 issue of the *Journal of Economic Perspective*, and Cain 1986. In the U.S., the unconditional gender wage differential, i.e. the ratio between average female earnings over average male earnings, is about 75%. This ratio has experienced a significant convergence in 1970s and 1980s, increasing from about 60% to about 75% and then remaining constant at this level during the 1990s (Eckstein and Nagypal 2004; Blau and Khan 2000). The U.S. rank average in comparison to other OECD countries: results for the late 1990s report Belgium at the highest level with a ratio equal to 90.1% in 1995, Japan at the lowest level with 63.6% in 1997 and the U.S. at 76.3% in 1996 from results in the same survey (Blau and Khan 2000). The most recent CPS release, 2004, reports a 80.36% ratio based on median weekly earnings for full-time workers.

[2]Results on differentials conditional on observable productivity characteristics are quite varied because they depend on the specification used and the decomposition implemented. These results are usually based on wage regressions with augmented human capital specifications. The differential may be captured by a simple gender dummy, by a decomposition based on estimated observable "returns," or by a residual. The general consensus is that a significant portion of the conditional differential remains "unexplained", even in the most sophisticated versions (Blau and Kahn 2003 and 2000; Altonji and Blank 1999).

[3]This point is generally acknowledged in the literature, see for example Blau and Kahn 2000 and Altonji and Blank 1999.

literature.[4] Spillover effects imply that the existence of a positive proportion of prejudiced employers lowers women's outside option at any type of employer. As a result, wage discrimination of women is present not only at prejudiced employers but also at unprejudiced employers.

Using data on accepted wages and unemployment durations, this model allows the separate identification of prejudice and unobserved productivity differences. The identification strategy exploits a distinctive feature of the observed earnings distribution of women with respect to men: female earnings are more concentrated in the left tail and the density to the left of the mode is quite flat.[5] The model generates this difference in shape because the observed female earnings distribution is a mixture between two earnings distributions: one of women working for unprejudiced employers and the other of women working for prejudiced employers. In equilibrium, the productivity reservation value in the second case is higher, generating matches with high productivity but low earnings. This additional mass of women earning low wages is able to generate the shape observed in the data. The difference between this earnings distribution and the shape of an earnings distribution implied by a model with only productivity differences allows for the separate identification and joint estimation of prejudice and gender productivity differences.

The model is estimated by maximum likelihood using CPS data for 1995. The year seems appropriate for this type of analysis because it is in a middle of a period of quite stable gender earnings differentials, following a period of fast convergence and before a period of higher variability in earnings differentials between men and women.[6] Results show that both discrimination and productivity differences are present in the labor market for white college graduates. Average female productivity is estimated to be about 6.5% lower than male productivity and explicit prejudice is estimated to involve about half of the employers.

Using these estimated structural parameters, it is possible to decompose the observed earnings differential taking into account equilibrium effects. The decomposition shows that prejudice is the most important factor in explaining the differential but productivity also plays a significant role. The simple presence of prejudice is able to generate about 2/3 of the observed earnings differential, while differences in productivity about 1/3. Then, two policy experiments are considered: an equal pay policy and an affirmative action policy. The equal pay policy imposes to pay the same wage at same productivity. At the estimated values, the policy significantly, but not completely, reduces the wage

---

[4]This is true also in the applied literature that explicitly models prejudice and the worker-employer relation such as Bowlus and Eckstein 2002 and Eckstein and Wolpin 1999.

[5]See Figure 1 and 2. The literature usually focuses on differences in means but this difference in shape is quite general and it is not specific to the year and sample used here. For example a similar shape is found on CPS data for 1985 and 2004 (the first and last year on which a similar sample can be extracted) and a similar pattern can be implied from the empirical cumulative density functions in Bowlus 1997 on data from NLSY.

[6]For recent earnings dynamic, in particular on the demographic group studied in this paper - white, College graduated in their mature working careers - see Eckstein and Nagypal 2004. For a study of more general women dynamics, see Blau 1998.

differentials, imposing a heavy welfare cost on employers. An affirmative action policy implemented as a quota system is shown to have no impact under the parameter estimates because a quite high proportion of women is already hired in the pre-policy equilibrium. Instead, an affirmative action policy defined as an employer's subsidy for hiring women is implemented. The policy implies a redistribution of welfare from men to women without significantly changing employers' welfare.

The paper is organized as follows. The next section briefly reviews some related literature. Section 3 presents the model and its main implications. Section 4 describes the data and the procedure to extract the estimation sample. Section 5 derives the likelihood function and discusses the identification strategy. Section 6 reports and discusses the estimation results. Section 7 contains the policy experiments based on the estimated parameters and section 8 draws some conclusions. An appendix contains all the proofs of propositions, a detailed description of the estimation sample extraction, the derivation of the likelihood functions and a more formal treatment of the identification.

2 Related literature

The search model with matching and bargaining utilized in this paper is a fairly standard framework to study labor market dynamics.[7] It is a very tractable improvement on partial job search models, allowing for a wider range of equilibrium effects to take place once major policy or structural changes are introduced. Search-matching-bargaining models have been estimated to study a variety of issues, such as: duration to first job and returns to schooling (Eckstein and Wolpin 1995); race discrimination (Eckstein and Wolpin 1999); the impact of mandatory minimum wage (Flinn 2005).

An alternative assumption to model labor market dynamics is to use an equilibrium search model. Equilibrium search models are based on the Burdett -Mortensen framework (Burdett and Mortensen 1998) and have the advantage of endogenously generating the dispersion in the wage distribution. Recently, the computational burden necessary to estimate these models has been overcome and structural estimates of equilibrium search model have been provided.[8] Still, the equilibrium search model generates some counterfactual empirical implications that can be partially solved only by imposing unmodelled workers and/or employers heterogeneity. This layer of heterogeneity on top of the form of heterogeneity studied in this paper, gender and prejudice, makes this approach a

---

[7] Jovanovic 1979 gives theoretical foundation to the importance of match-specific productivity in explaining labor market dynamics. Flinn and Heckman 1982 provide the basic theory for identification.

[8] Bowlus, Kiefer and Neumann 1995 use within market heterogeneity to estimate a Burdett-Mortensen model, while Van den Berg and Ridder 1998 use between market heterogeneity. Postel-Vinay and Robin 2002 estimate an equilibrium search model with both firm and worker heterogeneity, but they show that access to data on both sides of the market is essential to obtain identification.

less promising candidate to separately identify and jointly estimate productivity differences and prejudiced behavior.

A theory of taste discrimination was first proposed by Gary Becker in 1957.[9] The idea is to relate prejudiced behavior to preferences that economic agents may have with respect to clearly identified groups. Taste discrimination is still the most widespread, albeit debated, theory of prejudiced behavior and has already been combined with search models to study labor market discrimination. Search models are one of the most promising way to extend Becker's theory of discrimination (Altonji and Blank 1999) because the monopsony power induced by search frictions generate positive profits. Some employers can then choose to "indulge" in their prejudice according to their preferences, generating the persistent discrimination that we seem to observe (Heckman 1998). For example, Rosen 2003 develops a search and bargaining model showing that employers' taste discrimination generate persistent wage differentials. Black 1995 develops a search model where taste discrimination leads to complete segregation and shows that prejudiced employers survive in equilibrium if heterogeneity in entrepreneurial ability is introduced. Sasaki 1999 focuses on gender differences and shows that male welfare gains are enough to generate persistent discrimination in presence of coworker taste discrimination. Finally, Borjas and Bronas 1989 introduce consumer taste discrimination and asymmetric information in a search setting to generate empirical prediction about self-employment.

None of the previous papers, though, has the objective to separately identify the impact of prejudice and unobserved productivity differences. This objective is the focus of Bowlus and Eckstein 2002, an equilibrium search model with employer's taste discrimination against black workers. Their results, estimated on a sample of black and white high school graduates extracted from NLSY, indicate that the importance of productivity differences in explaining race wage differentials is significantly reduced when explicit discrimination is considered. In particular, about 56% of employers are estimated to be prejudiced and disutility from hiring black is estimated at about 31% of white productivity. The main objective of the paper, though, is more proposing an identification strategy than providing reliable estimates. Indeed, the model assumes no firm or job heterogeneity, generating counterfactual implications on wage distributions. As a result, maximum likelihood is not applicable and the estimates obtained by matching moments are not very robust.[10]

Bowlus 1997 is one of the very few papers that structurally estimates a search model focusing on gender differences in the labor market. The paper implements an equilibrium search model with firm heterogeneity to separately identify unobserved productivity differences and differences in behavior, but it does not assume any theory of discrimination. Differences in behavior are

---

[9]See Becker 1971.

[10]Estimates are not robust when both discrimination and productivity differences are assumed: in this case even bootstrap standard errors are problematic to obtain. As the authors comment on their results: *"For robust parameter estimates, one would need to modify the model by adding heterogeneity in firm productivity levels"* [pg. 1329, footnote 37, Bowlus and Eckstein 2002].

summarized by different rates of transition between employment, unemployment and non-participation. Results on NLSY data indicate that men have lower job separations rates and about 20% higher average productivity. The problem of this specification, as the author acknowledges, is that the large estimated difference in unobserved productivity is a sort of "catch all" variable, explaining all the residual differential in wages including, possibly, prejudiced behavior against women.

Finally, Eckstein and Wolpin 1999 are concerned with race discrimination and use a search-matching-bargaining model similar to the one used in this paper. They do not model prejudiced behavior but identify discrimination with differences in the Nash bargaining power coefficient. The main problem is that this coefficient is not identified unless some firm side data are available[11] and so they are forced to simply compute bounds for discrimination that end up being not informative on the estimation sample they work with.

3 The Model

3.1 Environment

The model is in continuous time, populated by four types of agents infinitely lived: two types of workers (Men and Women) and two types of employers (Prejudiced and Unprejudiced). The proportion of prejudiced employers is indicated with p and the proportion of male workers is denoted by m and they are both common knowledge to all the agents. Workers meet employers following a Poisson process with an instantaneous rate of arrival λ. The search process is random[12] and there is no on-the-job search. Once an employer and a worker meet, they observe a match-specific productivity value (x), modelled as a draw from an exogenous distribution denoted by the cdf G. Once a match is formed, it can be terminated following a Poisson process at an instantaneous rate η.

Wages are determined through wage bargaining between employers and workers upon observing the match value and their types. Hence, the wage schedule is a function of the match productivity, the threat points of the agents involved and their relative bargaining power coefficient α. Workers' utility functions are linear in wages and no disutility from working is assumed. While unemployed, workers receive an instantaneous utility flows b which can be interpreted as an unemployment benefit (if positive) or as the cost of searching or as any other disutility from being unemployed (if negative.) The other exogenous common knowledge parameter in the model is a discount rate ρ, assumed to be the same for employers and workers.

---

[11] They first discuss this point in Eckstein and Wolpin 1995, on which Eckstein and Wolpin 1999 is based. See also Flinn 2002b and 2005 for a discussion of the same identification issue.

[12] It may seem more realistic to introduce directed search in this context, as for example in Mailath, Samuelson and Shaked 2000. However, the focus of the paper is on the empirical prediction of the model and, given the data at hand, it does not make an identifiable difference to assume exogenously different arrival rates, as it is done in the empirical implementation of the model, or to model directed search behavior.

The workers' type is defined by an observable characteristic (gender) that induce a different behavior in the employer the worker is meeting. This different behavior is relevant in the bargaining process to determine wages and it happens conditionally on the realization of the meeting. Moreover, the two groups of workers may have ex-ante differences in some of the fundamental parameters that explain the labor market dynamic such as the arrival rate and the productivity distribution. In the theoretical presentation of the model, though, productivity and behavior are assumed to be the same for the two groups. This "homogenous" formulation is adopted to present the model more concisely and to clearly display that the presence of prejudiced employers is enough to replicate the standard descriptive empirical evidence we observe. Heterogeneity will then be introduced in the empirical section to estimate specifications that include both prejudice and gender differentials in productivity and behavior.

The employers' type is defined by a difference in preferences: prejudiced employers receive a disutility flow (d) from hiring women. This is a case of the taste discrimination model, first developed by Gary Becker in 1957.[13]

3.2 Value Functions

Unprejudiced and prejudiced employers are denoted by $I = N, P$ while male and female workers are denoted by $J = M, W$. The value of employment for a worker of type J working at an employer of type I at a wage $w_{JI}(x)$ is, for any $J = W, M$; $I = N, P$:

$$W_J[w_{JI}(x)] = \frac{w_{JI}(x) + \eta U_J}{\rho + \eta} \tag{1}$$

Equation (1) states that the value of employment is the current instantaneous value of the state for the worker ($w_{JI}(x)$) plus the value of the other possible state (unemployment, U_J) weighted by the probability associated to this event (η), all appropriately discounted by the instantaneous rates ρ and η. This expression results from the Poisson process assumption and the stationarity of the environment. The conjecture to solve for the policy rule is that the value of unemployment is constant with respect to wages, whereas the value of employment is clearly increasing in wages. It is therefore possible to find a reservation wage w_{JI}^* such that the values of the two states are equal. By equation (1) this value is: $w_{JI}^* = \rho U_J$. Since wages are determined by bargaining and productivity is match-specific, it will be necessary to find the reservation productivity value that corresponds to this reservation wage to have a complete description of the equilibrium.

While unemployed, a potential worker receives some instantaneous (dis)utility from unemployment (b) and as a result of the search activity three events may happen: not meeting any firm, meeting a prejudiced employer or meeting an

---

[13] The definition in Becker is: a discriminator is an employer who "*when faced with the money wage rate π [..] acts as if $\pi(1+d)$ were the net wage rate, with d being a discrimination coefficient measuring the intensity of his taste for discrimination*" (Becker 1971 [first. ed. 1957], pg. 39).

unprejudiced employer. By stationarity and by the Poisson processes governing meetings and terminations, the value of unemployment for a worker of type $J = W, M$ is then given by:

$$\rho U_J = b + \lambda \{ p \int \max[W_J[w_{JP}(x)] - U_J, 0]dG(x) + \qquad (2)$$

$$+ (1 - p) \int \max[W_J[w_{JN}(x)] - U_J, 0]dG(x) \}$$

Equation (2) confirms the conjecture of a constant value of unemployment and it has the usual interpretation: the reservation wage should compensate the state of unemployment with expected gains from matching with a prejudiced or an unprejudiced employer.

3.3 Wages

When employers and workers meet, the value of the match and the types are fully revealed. Common knowledge of the matching value is the usual practice in these models and it rules out the possibility of statistical discrimination.[14] But, while it seems pretty realistic to assume that the gender of a worker is fully observed,[15] it is less obvious to consider the employer's type fully revealed upon meeting. Yet, this is the usual assumption of equilibrium search models with workers and firms heterogeneity[16] and can be justified as follows. Ex-post distributions of female and male employees working for a given employer are different conditioning on the employer's type and they are observable once the meeting occurs. By observing these distributions, then, the worker may elicit the type of employers is meeting with. Moreover, if the information set of the two agents is enough to know the productivity value of the match then it is likely to be enough to reveal the types.

Since the axiomatic Nash bargaining solution is assumed, wage schedules are determined by choosing a wage that maximizes the product of the surplus in the match of the two agents, weighted by their relative bargaining power coefficient. The workers' surplus is given by the difference between the value of accepting the job, $W_J(w_{JI})$, and the value of the alternative, U_J. The employer's surplus is given by the discounted profit plus, in the case of prejudiced employers, the disutility from hiring women, as the following behavior implies.

Employers maximize utility, labor is the only factor of production, there are constant returns to scale and therefore the total output at a given employer is the sum of the productivity levels (x) of all his/her matched employees. Employers

---

[14]In a standard model of statistical discrimination it is exactly the asymmetric information over x that generates different labor market outcomes between the two groups. Ruling out statistical discrimination is a useful implication in this context since it would be very difficult to separately identify two types of discrimination on top of productivity differences.

[15]This is not always the case: in some context the gender of a potential employee is omitted intentionally to avoid potential prejudice. An example, studied in Goldin and Rouse 2000, are the blind auditions to hire musicians in some major US orchestras.

[16]See for example Postel-Vinay and Robin 2002; Van den Berg and Ridder 1998; Bowlus, Kiefer and Neumann 1995.

earn no revenues but make no payment if a match is not realized, therefore the value of their outside option is zero. Admittedly, the account of employers' behavior is very stylized but it seems to generate a model able to capture the essential features of the employer-worker search dynamic.

Using equations (1) and (2), Nash bargaining implies:

$$
\begin{aligned}
w_{JI}(x, U_J) &= \arg\max_w \left\{ \frac{[w - \rho U_J]^\alpha [x - d\mathbf{I}_{\{W,P\}} - w]^{(1-\alpha)}}{\rho + \eta} \right\} \qquad (3) \\
&= \rho U_J + \frac{\alpha}{1 - \alpha} \left[x - d\mathbf{I}_{\{W,P\}} - w \right]
\end{aligned}
$$

where $\mathbf{I}_{\{W,P\}}$ is an indicator function equal to one when the worker is female $(J = W)$ and the employer prejudiced $(I = P)$. Equation (3) states that the wage is equal to the threat point of the worker plus a portion of the surplus of the employer that is increasing in the relative bargaining power of the worker $(\frac{\alpha}{1-\alpha})$. A behavioral interpretation of this solution is that it is the unique subgame perfect equilibrium of the basic alternating offer game over dividing the surplus, in continuous time. Under this interpretation, the common discount value for all the agents implies that also the bargaining power coefficients are the same[17], that is $\alpha = 1/2$.

In terms of agents' types, four matches are possible: men with prejudiced or unprejudiced employers and women with prejudiced or unprejudiced employers. The wage schedules that corresponds to these matches can be obtained from equation (3).

A man matched with an unprejudiced or a prejudiced employer, $I = N, P$, will receive:

$$
\begin{aligned}
w_{MI}(x, U_M) &= \rho U_M + \alpha(x - \rho U_M) \qquad (4) \\
&= \alpha x + (1 - \alpha)\rho U_M
\end{aligned}
$$

The first line of equation (4) states that the wage of a realized match should guarantee the worker the reservation value ρU_M plus a portion α of the total surplus of the match, i.e. $(x - \rho U_M)$. Note that the wage schedule is independent of the employer's type and induces a reservation value on the productivity of the match. This is the truly relevant reservation value because both reservation wages and reservation profits depend on the match productivity value. The match reservation value is such that $W_M[w_{MJ}(x^*_{MJ}, U_M)] = U_M$ and $\pi_{JM}(x^*_{JM}, U_M) = 0$. Using (1)-(3), it is determined to be $x^*_{MJ} = x^*_{JM} = \rho U_M$. This is the value above which both the employer and the worker agree to enter the match. This non-disagreement result is an implication of the Nash bargaining assumption as shown by the first order condition of the maximization problem in equation (3). To summarize, the optimal decision rule is: a male worker and an employer both agree to accept the match, with wages governed by the wage schedule (4), if $x \geq \rho U_M$ and they both agree to reject the match otherwise.

---

[17]See Binmore, Rubinstein and Wolinsky 1986 and Binmore 1987.

A woman matched with an unprejudiced or a prejudiced employer, $I = N, P$, will receive:

$$w_{WI}(x, U_W) = \alpha \left(x - d\mathbf{I}_{\{W,P\}} \right) + (1 - \alpha)\rho U_W \tag{5}$$

Equation (5) has the same interpretation of equation (4) but now the wage schedule depends on the employer's type. Matching with a prejudiced employers will shift down the wage schedule thus creating wage discrimination with respect to women working for an unprejudiced employers. Proposition 2 will prove that also wage discrimination with respect to men is present due to the lower female value of unemployment in equilibrium. As before, the non-disagreement point is determined by equating the values of the two possible states for workers and employers, leading to $x_{WI}^* = x_{IW}^* = \rho U_W + d\mathbf{I}_{\{W,P\}}$. Therefore, even if the reservation wage is the same for all the female workers, the reservation productivity value is higher for women matched with prejudiced employers. A woman is more picky to accept a job from a prejudiced than from an unprejudiced employer. She will accept, though, if the value of the match is high enough because wages are increasing in productivity. A symmetric argument holds for the prejudiced employer since profits also are increasing in productivity. When female workers have all the bargaining power ($\alpha = 1$), they pay all the cost of discrimination as measured by the disutility d. If this is the case, a complete segregation result is likely to occur.

3.4 Equilibrium and Model Implications

To define the equilibrium we need to express the reservation values as function of the exogenous parameters. Using equations (1), (2) and (3) the reservation wage values ρU_J are implicitly determined by:

$$\rho U_J = b + \frac{\lambda \alpha}{\rho + \eta} \{ p \int_{\rho U_J + d\mathbf{I}_{\{W\}}} [x - d\mathbf{I}_{\{W\}} - \rho U_J] dG(x) + \tag{6}$$
$$+ (1 - p) \int_{\rho U_J} [x - \rho U_J] dG(x) \} \quad , \quad J = M, W$$

The equilibrium is therefore defined as:

Definition 1 *Given a vector $(\lambda, \eta, \rho, b, \alpha, d, p)$ and a probability distribution function for productivity of match values $G(x)$, an **equilibrium** is a vector of values of unemployment $U^* = (U_M^*, U_W^*)$ that solves equations (6) for $J = M, W$. The equilibrium vector U^* determines all the reservation values that constitute each agent's decision rules.*

An important implication of this equilibrium concerns the value of unemployment for the two types of workers. What we expect is a lower value of unemployment for women because a positive measure of prejudiced employers worsens their perspectives in the labor market. This result is stated in the following proposition.[18]

---

[18] All the proofs are in Appendix 9.1.

Proposition 2 *For any equilibrium previously defined such that $0 < p \leq 1$ and $d > 0$, the value of unemployment for women is lower than for men, i.e.*

$$U_W < U_M \qquad (7)$$

As discussed earlier, the main empirical motivation of the debate over labor market discrimination is the presence of gender differentials in unconditional or conditional means of some measure of wages. These differentials are crucial because they are supposed to proxy different wage rates for equally productive workers. This is also the definition of wage discrimination adopted in the literature. Another empirical evidence that is often considered is segregation, defined as the concentration of minority workers in relatively few sectors of the economy.[19] It is useful to define both concepts within the model to clarify the implications of the equilibrium for the broad empirical evidence at our disposal.

Definition 3 *In the economy defined so far, workers' type J suffers **wage discrimination** with respect to workers' type J' if and only if they are paid a lower wage conditioning on same productivity, i.e.:*

$$w_{JI}(x) - w_{J'I}(x) < 0 \quad \text{for any productivity value } x$$

Definition 4 *In the economy defined so far, **complete segregation** means that all workers of type J work for employers of type I; **partial segregation** means that workers of type J work in higher proportion for employers of type I.*

Another interesting impact of the presence of employers taste discrimination, often ignored by the literature, is the equilibrium effect of prejudice on employers that are not prejudiced. This seems a relevant issue and clarifies that prejudice and wage discrimination are two separate concepts, even if they quite often overlap. Wage discrimination may simply be a best response in a given environment, without any implication in terms of preferences. Moreover, wage discrimination disjoint from prejudice clearly has very different policy implications than a situation in which wage discrimination and prejudice coincide. In this respect, an interesting case described by the model is the behavior of unprejudiced employers that discriminate women simply because the presence of prejudiced employers worsen the bargaining position of women. This situation is summarized in the following definition.

Definition 5 *In the economy defined so far, **spillover effects** means that the presence of prejudiced employers induces wage discrimination also at unprejudiced employers, i.e.:*

$$w_{JN}(x) - w_{J'N}(x) < 0 \quad \text{for any productivity value } x$$

---

[19]For a review of the empirical evidence on segregation see for example Altonji and Blank 1999, Blau 1998.

Given these definitions, it is possible to summarize the implications of the model with respect to some widespread descriptive evidence on gender differentials in the labor market. This evidence can be summarized by male average earnings higher than female average earnings and by some degree of concentration of women in some sectors and occupations.[20] These implications can be directly derived from Proposition 2 and are summarized in the following:[21]

Corollary 6 *For any equilibrium previously defined such that $0 < p < 1$ and $d > 0$, women suffers wage discrimination and spillover effects with respect to men.*

Corollary 7 *For any equilibrium previously defined such that $0 < p < 1$ and $d > 0$, there is no complete segregation.*

The corollaries show the ability of the model to match the broad empirical evidence without introducing any source of heterogeneity on top of prejudice. Complete segregation is an outcome rarely found in the data, while some amount of partial segregation seems the consensus in the literature. Corollary 7 basically states that partial segregation is the most likely outcome in this economy. Wage discrimination, as implied by Corollary 6, generates wage differentials even if men and women are identical in terms of productivity and behavior. They arise from two channels. The first is standard: women working for a prejudiced employers will be paid less than a men equally productive. The second channel is the spillover effect that results from bargaining: the presence of a positive proportion of prejudiced employers implies in equilibrium a lower value of entering the labor market for women. This value is proportional to the threat point of women while bargaining with *any* employer. Therefore, even when working for an unprejudiced employer, women will receive lower wages at the same level of productivity because they are less able to extract rent when bargaining for wages.[22]

The relevant question is now empirical: how important is the impact of prejudice when other sources of heterogeneity are present? In particular, if men and women differs in ex-ante productivity and search behavior, to what extent prejudice is still a major factor in explaining gender differentials in the labor market? The result depends on parameter values and the objective of the estimation section is to obtain these values for a representative sample of U.S. workers. To answer these questions is also important to allow for spillover

---

[20]Results are different across countries and across time, see footnote 1 and the Introduction for more details and references. However, a positive gender differential in earnings and some degree of segregation are an evidence surprisingly common and persistent.

[21]All the proofs are in Appendix 9.1.

[22]Some episodic evidence is present in this respect. The academic market, for example, is a typical situation in which one to one bargaining seems a reasonable description of the actual wage setting. In this market, a report on faculty at a top institution concludes that *"women [are] receiving less despite professional accomplishments equal to those of their male colleagues"* (MIT, 1999 as reported in Blau and Kahn 2000) and one of the main reason of this outcome seems exactly the lower ability or willingness to bargain, as for example by using alternative offers as outside options.

effects since they magnify the impact of prejudice on worker welfare as proxied by accepted wage and unemployment dynamics. By the same reasoning, the simple observation of labor market differential may quite overestimate the amount of prejudice.

4   Data

The sample used in estimation is extracted from the *Annual Social and Economic Supplement* (March Supplement) of the Current Population Survey (CPS) for the year 1995. A more detailed description of the Survey and the sample extraction procedure can be found in Appendix 9.2. CPS is a very large nationally representative sample which permits disaggregation by relatively homogenous subgroups. The year 1995 is in the middle of a period of relative stability of gender differentials in the labor market. For example, Eckstein and Nagypal 2004 assess the US labor market dynamic over 1961-2002 using CPS and they show that the female-to-male wage ratio has increased substantially from the mid-seventies to the early nineties, starting to fluctuate after the late nineties. Blau and Khan 2000 show a significant convergence in the seventies and eighties, with the ratio increasing from about 60% to about 75%, and then a substantial stability at this level during the nineties.

The estimation sample is extracted among individuals that are:

- 30 to 55 years old (extremes included);

- employed or looking for a job;

- classified as *white*;

- holding a College degree or more.

These selection criteria are introduced to guarantee a degree of homogeneity to the sample. In the model workers are assumed to be homogenous with the only exception of gender. Selecting a sample homogenous with respect to some observables correlated with performance in the labor market is therefore the minimum requirement for a meaningful empirical application.

The age limitation is introduced to focus on individuals with mature working careers: they are reasonably homogenous in terms of working experience and are more likely to be in the steady state position assumed by the model. Then, only individuals employed or looking for a job are considered to exclude the category *layoffs* since layoffs are a labor market dynamic typically not related with the search process. Since race or ethnic group is an observable highly correlated with labor market performance, I will concentrate only on the most numerous ethnic group, i.e. whites. The classification *white* in the CPS is chosen by respondents out of the following alternatives: White, Black, American Indian-Eskimo, Asian/Pacific Islander, Other.[23]

---

[23]There is an issue about how individuals of Hispanic origin classify themselves and about

Education is a necessary homogeneity control since it is one of the main component of the individual human capital and schooling is highly correlated with labor market performance. The group of College graduated or more seems the most appropriate to fit the behavior assumed in the model since a one to one bargaining is more likely to occur for skilled labor. Moreover, employers are more likely to have a direct and frequent contact with workers in skilled job positions, therefore skilled jobs are a setting in which the taste discrimination argument seems more plausible.

The observed labor market variable used to estimate the contribution to the likelihood of unemployed individuals is the individual unemployment duration. No wage information for unemployed individuals is available. Unemployment duration is originally recorded in weeks[24] and in 1995 is top-coded at 99 weeks.[25] The estimation sample includes all the individuals satisfying the previous four criteria except the 3 top-coded observations. As standard in the literature, the weekly unemployment durations are transformed in monthly unemployment durations. The final sample contains 49 unemployed individuals, 28 women and 21 men. Some descriptive statistics are reported in Table 1. Women have on average lower unemployment duration but an higher unemployment rate.[26] Mean and standard deviation of unemployment durations are roughly equal, an indication that the exponential distribution implied by the model is not rejected by the data.

The labor market variable observed for employed individuals is earning at the date of the interview. Earnings are recorded before deductions on an hourly or weekly basis.[27] For some individuals both observations are available but usually only one of the two is present. I transformed all earnings in hourly earnings either using the recorded observation directly or by dividing weekly earnings for the number of hours worked per week reported by the individual. For 67 observations it is impossible to calculate hourly earnings because the number of hours worked is missing. Conditioning on the homogeneity controls, the sample contains 1,031 women and 1,244 men with valid observations on hourly earnings. Some descriptive statistics are in the top panel of Table 1.

---

the opportunity to distinguish them from other *white*. A recent question in the CPS allows these finer differences but they do not seem to influence the results presented in this paper.

[24] Weeks in uneployment are obtained by the CPS as answers to the following questions: (i) *As of the end of last week, how long had you been looking for work?* and (ii) *We would like to have that in weeks if possible. Exactly how many weeks (have/has) (name/you) been looking for work?*

[25] CPS documentation asserts this variable being topcoded at 999 weeks but in practice there are no observations higher than 99 weeks and exactly at 99 weeks there is a cluster of 10 observations in the raw data. Moreover the topcoding at 99 weeks is the one officially used from 1968 till 1993.

[26] Bowlus 1997, using a sample from NLSY, finds instead higher unemployemnt durations for women. This difference may be due to cohort effects and to the inclusion of transitions from non-participation. The participation rate differential for the demographic group considered in this paper is not too high: 96.8% for men and 85.2% for women.

[27] Earnings are obtained by the CPS as answers to the following questions: (i) *What is your best estimate of your hourly rate of pay?* and (ii) *What is your best estimate of your weekly earnings before taxes or other deductions?*

It is very often assumed that earnings or wages are measured with an error. Validation studies suggest that the two main sources of measurement errors in the CPS are overreporting of earnings by individuals at low income levels and underreporting by individuals at high income levels. Bollinger 1998 is a validation study that analyzes yearly earnings in the CPS concluding that the main problem is overreporting at low income, in particular for males. A test for equality between recorded CPS earnings and true earnings is rejected on the male sample but it is not rejected on the female sample. This difference seems due to a small subsample of men with very low income. Validation studies on hourly or weekly earnings are less common because it is more difficult to obtain the corresponding validation data.[28] The conclusion seems that earnings collected not on a yearly basis have larger measurement errors (Bound, Brown and Mathiowetz 2001).

A solution commonly used to take into account measurement errors assumes that observed wages are measured with an error generated from a parametric distribution.[29] Given the type of errors present in the CPS, this solution does not seem to solve the main measurement problems and it does not help to fit the model.[30] A more crude solution is adopted in this paper: the impact of measurement errors is reduced simply by trimming the sample. Underreporting in the top tail does not seem a major problem on the demographic group under consideration because the top-coding is not too high (1,923 dollars per week). In this respect, a problem could be the distortion in the shape of the observed earnings distribution due to the mass points typically created by top-coding. However, the problem is somewhat alleviated by variation in weekly hours worked[31] and in the end I have chosen just to drop the 6 observations above the top-coded value. The low tail is more problematic because there are some very low observations below the minimum wage[32] and because over-reporting in the low tail is expected to be asymmetric by gender (Bollinger 1998.) I have therefore chosen to trim some of the low tail. The amount of trimming implies a trade-off: the more observations are deleted, the higher is the cleaning but the higher is the distortion of the earnings distribution, in particular of the female distribution where many observations are clustered in the low tail. The compromise I have reached is to drop the bottom 5%, computing the percentile separately on the male and female earning distributions.[33] This threshold should be enough

---

[28] An example on CPS is Mellow and Sider 1983.

[29] For example, this assumption is used in the context of a search model by: Flinn 2002a; Eckstein and Wolpin 1999; Wolpin 1987.

[30] Very often this parametrization is used to take into account probability-zero events observed in the sample, such as job-to-job transitions associated with a wage loss in an on-the-job search model (see for example Wolpin 1987 and Flinn 2002a).

[31] Top-coding on hourly earnings is binding only on 5 observations. Top-coding on weekly earnings is more significant: 122 observations. Variation in hours worked somewhat reduces the impact of these observations on the shape of the distribution but still three mass points on the male distribution and one on the female distribution are generated on the top tail (see Figures 1 and 2). Mass points are in the order of a dozen observations each.

[32] There is a total of 28 observations below the minimum wage, with a minimum at 1$ per hour.

[33] The 5% percentile is a common bound on CPS, see for example Bowlus 1997. In terms of

to significantly reduce the difference in measurement errors between man and women by dropping the subsample of men with very low income responsible for most of the overreporting.

The final sample after the trimming and the unemployment durations data previously presented constitutes the estimation sample. Some descriptive statistics are reported in the lower panel of Table 1. The earning differential is basically unchanged by the trimming: the female-to-male earnings ratio is 79.4%, a value consistent with data on representative samples for the U.S.[34] This ratio is slightly higher compared to the whole population: using CPS, the ratio is 77.07% over the entire labor force and 74.53% over whites in the same age range. To compare over time, I have computed the ratio on CPS samples extracted following the same criteria in different years: for the mid-1980s, the ratio on 1985 is 73.86%; for the last year available, the ratio on 2004 is 76.85%.

In terms of second moments, male earnings have an higher standard deviation, mainly because female earnings tend to be more clustered in the left tail. If we look at the empirical earnings distributions reported in Figure 1 and 2 this difference in shape is quite clear: female earnings are highly concentrated between the minimum wage and 20 dollar per hour. Also the slopes of the two densities are quite different: the female density is sort of flat at the beginning, in particular if we ignore the spikes due to rounding, and then decreases quite fast after 20 dollar per hour; the male density has a more regular shape: with a less rapid increase and a smooth decrease after the 15/20 dollar per hour level.

5   Estimation: specification and identification

As discussed in the previous section, there are some data limitations in the CPS: for each respondent we can either observe the hourly earnings or the unemployment duration and no information is available from the employers' side so that it is impossible to directly identify prejudiced or unprejudiced employers. Given the data and the restrictions implied by the model, though, it is possible to estimate some crucial structural parameters if a parametric assumption on the productivity distribution is imposed.

The CPS sample can be described as a vector $\left(\{w_i\}_{i\epsilon E_J}; \{t_i\}_{i\epsilon U_J}\right)_{J=M,W}$ where w_i are hourly earning, t_i are on-going unemployment durations measured in months, E_J is the set of employed individuals of gender J, and U_J is the set of unemployed individuals of gender J. For clarity and comparative purposes,

---

parameters estimates, results do not change much for cutoffs in the range of the 1%-8%. If the cutoffs are set at higher percentiles (10% or 15%), estimates of the disutility from prejudice (d) in the specification that allows for productivity differences become more imprecise while the other parameters are relatively stable.

[34]Bowlus 1997 uses NLSY on white College graduates and finds a ratio equal to 0.815 on weekly wages. Without controlling for race, Eckstein and Nagypal 2004 use the annual wages and salary earnings for full-time full-year workers from CPS 1995 and compute a ratio equal to 0.662 on College graduates between the ages 22 and 65. Without controlling for race, Blau 1998 uses weekly wage for full-time workers on CPS 1994 and finds a ratio equal to 0.736 on 35-44 years old with 16 years of education or more.

it may be useful to start presenting the empirical specification in the model without prejudice, i.e. with $(p, d) = (0, 0)$.

5.1   Without Prejudice

In the theoretical model presented in section 3, men and women are identical if no prejudice is present. In the empirical specification, though, it is important to allow for behavioral and productivity differences between men and women. The final objective will be to estimate the presence and extent of prejudice once these other structural differences are accounted for. Productivity differences are introduced by letting the parameters of the productivity distribution to differ. Other behavioral differences are captured by allowing for gender specific arrival rates, termination rates and instantaneous values of unemployment. They describe in a reduced form fashion differences in the intensity of search or in other behaviors related to the labor market search process. The subscript $J = M, W$ is used to denote these differences in parameters.

Using equations (4), (5) and (6), the model under this specification is summarized by the wage equations

$$w_J(x, U_J) = \alpha x + (1 - \alpha)\rho U_J \tag{8}$$

and by the reservation wage equations

$$\rho U_J = b_J + \frac{\lambda_J \alpha}{\rho + \eta_J} \int_{\rho U_J} [x - \rho U_J] g_J(x) dx \tag{9}$$

where, as previously shown, the reservation wage is equal to the reservation match value, $w_J^* = x_J^* = \rho U_J$. To proceed with a maximum likelihood estimation, we need a parametric assumption on the productivity-match distribution that satisfies a recoverability condition.[35] The most common assumption is to consider the wage distribution or the match values distribution to be lognormally distributed.[36] A visual inspection of the empirical distributions of the accepted earnings (Figure 1 and 2) suggests that this is a sensible assumption on this sample. By assuming a lognormal distribution for productivity, the density in equation (9) becomes:

$$g_J(x) = \frac{1}{\sigma_J x}\phi[\frac{\ln(x) - \mu_J}{\sigma_J}], \, x > 0 \tag{10}$$

---

[35] A distribution is *recoverable* from a truncated distribution if knowledge of the point of truncation and of the distribution above the point of truncation are enough to uniquely determine it. Flinn and Heckman (1982) show that in a search model with match-specific productivity, it is impossible to determine the shape of the productivity distribution below the truncation point (the reservation value) without a parametric assumption. This knowledge is essential to incorporate equilibrium effects in evaluating policy experiments.

[36] The lognormal assumption is used, among others, in: Flinn 2002a and 2005; Dey and Flinn 2005; Eckstein and Wolpin 1999 and 1995. Flinn and Heckman 1982 use, instead, an exponential and a normal distribution.

where ϕ indicates the standard normal density. The corresponding cumulative distribution function is indicated by $G_J(x)$ and the survival function by $\widetilde{G}_J(x)$. The parameters that describe productivity are therefore two for each type: $(\mu_J, \sigma_J)_{J=M,W}$.

The likelihood of the sample is given by the contribution of unemployed and employed individuals. To obtain the unemployed contribution, consider the hazard rate out of unemployment:

$$h_J = \lambda_J \widetilde{G}_J(\rho U_J) \tag{11}$$

This constant hazard rate is implied by the time homogeneity of the environment, the Poisson process that governs the arrival of job offers and the optimal decision rule. It is given by two components: the arrival rate of offers, λ_J, and the probability that a match is formed once the meeting occurs, $\widetilde{G}_J(\rho U_J)$. The unconditional unemployment contribution over on-going durations is the density of an exponential random variable with coefficient equal to the hazard rate times the steady state probability of unemployment:

$$
\begin{aligned}
f_u(t_i, i\epsilon U | J) &= f_u(t_i | i\epsilon U, J) P(i\epsilon U | J) \\
&= h_J \exp(-h_J t_i) \frac{\eta_J}{\eta_J + h_J}, \quad t_i > 0
\end{aligned} \tag{12}
$$

The probability of unemployment takes into account that durations are observed only on unemployed individuals. The complete derivation of the density (12) is presented in Appendix 9.3.

The contribution of employed individuals is based on the mapping between wages and match values reported in equation (8). Starting with the unconditional cumulative distribution function of earnings, the contribution of employed individuals is obtained using the optimal decision rules, the parametric assumption on the match distribution and ergodic results on flows in and out employment. This derivation leads to the following distribution of observed earnings:

$$
\begin{aligned}
f_e(w_i, w_i > \rho U_J, i\epsilon E | J) &= \\
= f_e(w_i | w_i > \rho U_J, i\epsilon E, J) P(w_i > \rho U_J | i\epsilon E, J) &P(i\epsilon E | J) = \\
= \frac{\frac{1}{\alpha} g\left(\frac{w_i - (1-\alpha)\rho U_J}{\alpha}\right)}{\widetilde{G}_J(\rho U_J)} &\frac{h_J}{\eta_J + h_J}
\end{aligned} \tag{13}
$$

The derivation is presented in Appendix 9.3.

Finally, defining the following indicators:

$$
\begin{aligned}
s_i &= \begin{cases} 1 & \text{if } i\epsilon W \\ 0 & \text{if } i\epsilon M \end{cases} \\
\chi_i &= \begin{cases} 1 & \text{if } i\epsilon E \\ 0 & \text{if } i\epsilon U \end{cases}
\end{aligned} \tag{14}
$$

18

and the following likelihoods:

$$
\begin{aligned}
L_i^{00} &= f_u(t_i, i\epsilon U|M) \\
L_i^{10} &= f_u(t_i, i\epsilon U|W) \\
L_i^{01} &= f_e(w_i, I, w_i > \rho U_M, i\epsilon E|M) \\
L_i^{11} &= f_e(w_i, I, w_i > \rho U_W, i\epsilon E|W)
\end{aligned}
\tag{15}
$$

Then, the likelihood on the observed sample $\left(\{w_i\}_{i\epsilon E_J}; \{t_i\}_{i\epsilon U_J}\right)_{J=M,W}$ under a model without prejudice is:

$$
\ln L =
\tag{16}
$$

$$
\sum_{i=1}^{N} \left[(1-s_i)(1-\chi_i)\ln L_i^{00} + s_i(1-\chi_i)\ln L_i^{10} + (1-s_i)\chi_i \ln L_i^{01} + s_i\chi_i \ln L_i^{11}\right]
$$

Identification discussion

Conditioning on gender, the model without prejudice is identified following the strategy suggested in Flinn and Heckman 1982. A more detailed description of identification can be found in Appendix 9.4 but the main argument is as follows. Looking at (16), the first thing to notice is that the structural parameters ρ and b_J enter the log likelihood only through the reservation matching value $\rho U_J = x_J^*$. It is therefore possible to estimate x_J^* as a free parameter in the likelihood and then recover ρ and b_J using the reservation wage equation (9). This argument also implies that the discount rate ρ and the instantaneous value of unemployment b_J cannot be separately identified. The usual practice is to fix a reasonable value for the discount rate and then recover the instantaneous value of unemployment.

Conditioning on gender, the five parameters to be identified become:

$$
(\lambda_J, \eta_J, x_J^*, \mu_J, \sigma_J)
$$

By equation (11) and by first order conditions on (16), the maximum likelihood estimators (denoted by $\widehat{\ }$) for the hazard rate out of unemployment and the termination rate are:

$$
\widehat{h}_J = \frac{N_{UJ}}{\sum_{i\in U_J} t_i}
\tag{17}
$$

$$
\widehat{\eta}_J = \frac{N_{UJ}}{N_{EJ}}\widehat{h}_J
\tag{18}
$$

where N_{EJ} and N_{UJ} denote the number of employed and unemployed for each gender. By making a parametric assumption related to observed wages, it is possible to separately identify the two components of the hazard rate: arrival rate and probability to accept the match. Such parametric assumption is necessary to identify the shape of the distribution below the truncation point. Even with

19

a parametric assumption, though, the maximum likelihood estimator is non-regular in this setting because the lower limit of the distribution of observed wages is now a parameter to be estimated (x_J^*).

To solve this problem Flinn and Heckman use a two-step procedure. In the first step the reservation wage is estimated non-parametrically by the minimum observed wage in the sample,[37] shown to be a strongly consistent estimator. This procedure is readily applicable here since the reservation wage is the same for all the individuals of same type and it is equal to the reservation match value to be estimated $(x_J^* = w_J^*.)$ Since the obtained $\widehat{x_J^*}$ converges at rate N, in the second step the remaining parameters can be estimated by maximizing a concentrated likelihood in which the reservation value parameter is replaced by the first step estimate:

$$\ln L_J = \ln L(\lambda_J, \eta_J, \widehat{x_J^*}, \mu_J, \sigma_J) \tag{19}$$

On this likelihood, knowledge of unemployment durations and accepted wages is enough to identify the remaining parameters.

The literature has frequently pointed out the difficulty of identifying the Nash bargaining power parameter[38] α. In this paper, it is identified by assuming symmetric bargaining for each type and therefore fixing $\alpha = 0.5$. This assumption has a behavioral foundation if we see the Nash bargaining solution as the outcome of a Rubinstein alternating offer game, under the assumption that workers and employers share the same discount rate.[39]

5.2   With Prejudice

The estimation and specification of the model with prejudice follow exactly the same lines of the model without prejudice. The difference is the introduction of employers heterogeneity. Employers can be unprejudiced N or prejudiced P, following the definition reported in section 3. As before, the matching distribution is assumed to be lognormal. Employers heterogeneity is indicated by the subscript $I = N, P$, in addition to the subscript $J = M, W$ used to indicate workers heterogeneity.

Simply rewriting equations (5), (4) and (6) to allow for gender-specific parameters and employers heterogeneity, the four wages equations become:

$$w_{JI} = \alpha(x - d\mathbf{I}_{\{W,P\}}) + (1 - \alpha)\rho U_J \tag{20}$$

---

[37] This estimator is also implemented by Bowlus 1997, Bowlus, Kiefer and Neumann 1995, Kiefer and Neumann 1993.

[38] The main problem is that this coefficient is not identified unless some firm side data are available, see for example Flinn 2005 and Eckstein and Wolpin 1999 and 1995. In particular, Eckstein and Wolpin 1999 dicuss this issue in the context of race discrimination because they define discrimination as differences in the Nash bargaining power coefficient.

[39] See Binmore, Rubinstein and Wolinsky 1986 and Binmore 1987. Setting the Nash bargaining coefficient to one half is a common solution in applied work, see for example Eckstein and Wolpin 1993 and Flinn 2005. Also Flinn and Heckman 1982, without explicitly assuming a bargaining structure, generate a wage schedule that corresponds to setting $\alpha = 1/2$.

and the two reservation wage equations are:

$$\rho U_J = b_J + \frac{\lambda_J \alpha}{\rho + \eta_J} \{ p \int_{\rho U_J + d\mathbf{I}_{\{W\}}} [x - d\mathbf{I}_{\{W\}} - \rho U_J] g_J(x) dx + \quad (21)$$

$$+ (1-p) \int_{\rho U_J} [x - \rho U_J] g_J(x) dx \}$$

From the definition of equilibrium in the model, we expect the reservation wages to be the same at both employers for each worker's type. This is proved by equations (20) and (21) and it is important for identification because we only observe accepted wages conditional on worker's type but not on employer's type. Conditioning on gender, workers are identical while looking for jobs. When workers and employers meet, the productivity value of the match and the types are revealed, generating heterogeneity. As a result the reservation match value is higher for women meeting prejudiced employers than for women meeting unprejudiced employers. The difference between these two reservation values is exactly equal to the intensity of discrimination d. This implication is crucial to identify the prejudice parameters when productivity differences are present. Male workers have one reservation wage equal to the reservation match-value, equal to the discounted value of unemployment:

$$w_M^* = x_{MI}^* \equiv \rho U_M \qquad (22)$$

while female workers have one reservation wage but two reservation match values determined by the type of employer they are meeting:

$$w_W^* = \rho U_W \qquad (23)$$
$$x_{WI}^* = \rho U_W + d\mathbf{I}_{\{W,P\}} \qquad (24)$$

As before, the unemployed contribution is based on the hazard rates out of unemployment:

$$h_J = \lambda_J [(1-p)\widetilde{G}_J(\rho U_J) + p\widetilde{G}_J(\rho U_J + d\mathbf{I}_{\{W\}})] \qquad (25)$$

Hazard rates are conditional on worker's type and they depend on the exogenous meeting rate and on the probability to meet and accept the match with an employer of a given type. Therefore, men and women may have different hazard rates for exogenous reasons - such as a different arrival rate - and for endogenous reasons - such as the equilibrium impact of the presence of prejudice on the probability to accept the match.

As recalled before and derived in more detail in Appendix 9.3, the density of ongoing unemployment spells will be a negative exponential with parameter equal to the hazard rate, weighted by the corresponding probability of unemployment:

$$f_u(t_i, i\epsilon U | J) = f_u(t_i | i\epsilon U, J) P(i\epsilon U | J) \qquad (26)$$
$$= h_J \exp(-h_J t_i) \frac{\eta_J}{\eta_J + h_J}, t_i > 0$$

21

Employed contributions are based on the mapping between wages and match values determined by the wage equations. The result, conditioning on workers' type and using ergodic results on flows in and out employment, is:

$$f_e(w_i, w_i > \rho U_J, i\epsilon E | J) \tag{27}$$
$$= f_e(w_i | w_i > \rho U_J, i\epsilon E, J) P(w_i > \rho U_J | i\epsilon E, J) P(i\epsilon E | J)$$
$$= \left[\frac{\frac{(1-p)}{\alpha} g_J(\frac{w_i - (1-\alpha)\rho U_J}{\alpha})}{\widetilde{G}_J(\rho U_J)} + \frac{\frac{p}{\alpha} g_J(\frac{w_i + \alpha d \mathbf{I}_{\{W\}} - (1-\alpha)\rho U_J}{\alpha})}{\widetilde{G}_J(\rho U_J + d \mathbf{I}_{\{W\}})} \right] \frac{h_J}{h_J + \eta_J}, w_i > \rho U_J$$

This density is a distribution truncated at the reservation wage and obtained by mapping productivity values into observed wages. As detailed in Appendix 9.3, the mapping exploits the equilibrium wage equations and the probabilities to accept the match together with the exogenous proportion and intensity of discrimination.

The likelihood for the model with prejudice on the sample $\left(\{w_i\}_{i\epsilon E_J} ; \{t_i\}_{i\epsilon U_J}\right)_{J=M,W}$ is then obtained by using the indicators for gender and employment s_i and χ_i defined in equation (14) and by replacing the densities (12) and (13) with the densities (26) and (27) leading to an expression equivalent to (16).

Identification discussion

As in the case without prejudice, the parameters ρ and b_J enter the likelihood only through the reservation matching value. The reservation value ρU_J will then be treated as a primitive of the model in the estimation. The same two steps procedure described in the previous case can be used: the reservation wage, unique when conditioning on worker's type, can be estimated as the minimum observed wage in the male and female sample. In the second step, a concentrated likelihood is obtained where $(\lambda_J, \eta_J, \mu_J, \sigma_J)_{J=M,W}$ are identified. The real issue in the model with prejudice is how to identify two additional parameters: the proportion of prejudiced employers p and the disutility from hiring women d. A formal argument for the identification of (p, d) is given in Appendix 9.4 but an intuition can be given by analyzing the impact of these two parameters on the accepted wages distribution to see whether a model with productivity differences and prejudice generates a distinctively different distribution from a model with only productivity differences.

The result of this exercise[40] is reported in Figure 3. The top panel, labelled Productivity, shows the density of accepted wages under no prejudice, i.e. $(p, d) = (0, 0)$. As expected, the shape resembles a lognormal, with a truncation at the reservation wage value. The second panel shows the impact of prejudice: this is the shape of accepted wages in a model when all the employers are prejudiced, i.e. $(p, d) = \left(1, \widehat{d}\right)$. The shape is now quite different, with the increasing part of the distribution before the truncation point almost missing. It is a result of the fact that matches between women and prejudiced employers are

---

[40] All the estimated values used in this illustration are gender specific. The appropriate specifications reported in Table 2 are used.

acceptable only at relatively high productivity values, that is when $x \geq \rho U_W + \widehat{d}.$ These high values are in the decreasing part of the lognormal density with high probability, generating the observed decreasing earning density.

When the proportion of prejudiced employers is not forced to be either one or zero, the observed earning distribution is a mixture between earnings of workers employed at prejudiced employers and earnings of workers employed at unprejudiced employers. This situation is reported in the bottom panel of Figure 3 where $(p, d) = \left(\widehat{p}, \widehat{d}\right)$. The shape of the histogram is now more similar to a lognormal than the full prejudice case but the low tail between the truncation point and the mode is less steep. This is due to the mixture between the increasing density of women working for unprejudiced employers (top panel) and the decreasing density of women working for prejudiced employers (second panel). With this mixture model, it is therefore possible to distort the distribution of accepted wages in a different way than in a model with simple productivity differences or with full prejudice. The difference in shape between the top panel and the bottom panel is ultimately what allows the identification of (p, d) on top of $(\lambda_W, \eta_W, \mu_W, \sigma_W)$.

The comparison of Figures 1 and 2 shows that this is exactly the type of distortion that needs to be explained in the data for women. Observations are more concentrated in the left tail and the density on the left of the mode is much flatter than the density from a lognormal with this amount of skewness. In other words, a lognormal density assigns more mass on the left tail by shrinking to the left, leading to a steeper slope out of the truncation point. But this is not what we observe in the data. The empirical distribution of women earnings with respect to men has more mass on the left tail but the slope out of the truncation point is actually flatter. This shape is consistent with the presence of a mixture between two earnings distributions, one at prejudiced employers decreasing right after the truncation point and another at unprejudiced employers increasing above the truncation point.

A first word of caution about this identification strategy relates to the different contribution of p and d. There is clearly some substitution between these two effects: one is not possible without the other. For example, a very high disutility with a very low proportion of prejudiced employers will have no impact just as the exactly opposite case. Out of the boundaries, though, the two parameters seem to have a sufficiently different impact, as illustrated in Figure 4 and, more formally, in Appendix 9.4.. In the left column, p is fixed at 0.5 and d is allowed to vary from 5 to 20; in the right column, the disutility is fixed to 15 and p changes from 0.1 to 0.9. The proportion of prejudiced employers mainly affects the mode of the distribution: an increase in p (from top to bottom panel of the right column) moves the mode to the left. The disutility d, instead, marginally affects the mode but increasingly flattens the distribution on the left of the mode (d increases moving from top to bottom in the left column.)

The second word of caution relates to the more general problem of identification through functional form assumptions. This is generally the case in the identification of search model and this paper shares the same well-known

limitations. The additional step of the identification of prejudice is also through functional form assumptions. First, it assumes that men and women can have different productivity but only up to two parameters. In other words, the two distributions are both of the same family, even if the parameters can then be different. Second, it assumes that the productivity distribution is a location-scale parameters distribution: this is crucial to obtain the distortion in shape previously discussed. Consider for example a distribution without a location parameter, such as a negative exponential. In this case the density is monotone decreasing and the impact of productivity and prejudice is simply reflected in a different truncation point with no major difference in shape. Therefore, also the mixture will not have any major distortion in shape, failing to fit the flatter lower tail of the female earning distribution. Appendix 9.4 shows how the presence of a location parameter is necessary for identifying (p, d), but also shows that identification is not limited to lognormal distributions. It is valid for any distribution with at least one location and scale parameter that can be recovered by observing the finite mixture.

With these limitations in mind, the next section shows how it is possible to implement this identification strategy on standard labor market data. The result is a fairly precise estimate of the contribution of behavior, productivity and prejudice in explaining gender labor market differentials.

6   Estimation: Results

6.1   Maximum Likelihood Estimates

Estimation results are reported in Table 2. In the first three columns arrival rates and termination rates are constrained to be the same for men and women, while in the last three columns they are unconstrained. Specifications (3) and (6) estimate jointly prejudice and productivity differences. For comparison purposes and check robustness, specifications (1) and (4) estimate a model without prejudice and specifications (2) and (5) a model with prejudice and no productivity differences.

Consider first the model without prejudice. This is the specification more similar to previous works that estimate search models with heterogenous workers: under segmentation in the labor market, the same model is separately estimated on the different groups. Arrival rates (λ_J) and termination rates (η_J) are estimated to be higher for women and the estimated values for the location and scale parameters $(\mu_J$ and $\sigma_J)$ imply lower average productivity for women. Average productivity and other predicted values are reported in Table 3. They indicate a quite substantial gap in productivity, able to replicate the average earnings differential observed in the data. This specification is comparable to Bowlus 1997 and produces quite similar results, despite differences in the data set and in the model.[41] Bowlus also finds higher female arrival rates and ter-

---

[41] Bowlus 1997 uses NLSY instead of CPS and assumes an equilibrium search model with firm heterogeneity instead of a search-matching-bargaining model.

mination rates and lower average female productivity. Female productivity is reported 17% lower in Table 5 of Bowlus 1997, a value comparable with the 21% differential in average productivity found here under specification (1) and (4). As acknowledged by Bowlus and others, unobserved productivity difference is a "catch all" variable, fitting all the residual variation in the wage distributions not captured by the model. With the introduction of an explicit theory of discrimination, some of this differential is allowed to be related to prejudiced practices in the labor market. These practices turn out to have a significant impact, as specification (3) and (6) will show.

Consider now the model with only prejudice. Assuming no differences in productivity, most of the wage differential must be explained by the prejudice parameters (p, d). Under this model, the majority of the employers, about 81%, are estimated to be prejudiced against women, with a disutility value of about 19% of the estimated average productivity. As expected, coefficients on productivity and discrimination do not change much between columns (2) and (5) since, as described in section 5.2, arrival rates and termination rates are mainly identified by unemployment durations.

Finally, columns (3) and (6) report estimates of the specification in which productivity differences and explicit prejudice are jointly estimated. Results show that both components play a significant role in explaining the gender differentials we observe. Under specification (3) average female productivity is estimated 6.48% lower than male productivity and under specification (6) about 6.55% lower. The extent of explicit prejudice is significant in both specifications: about half of the employers are prejudiced and disutility from hiring women is about 36% of the average male productivity.

To give an idea of the order of magnitude of these point estimates, consider the result about race discrimination obtained by Bowlus and Eckstein 2002: they estimate the disutility from hiring black at about 31% of white productivity and the proportion of prejudiced employers at about 56%.[42] For a comparison across time, I estimate the same model on a similarly selected sample from CPS for 1985, obtaining an average productivity differential of more than 17%, a proportion of prejudiced employers of about 74% and a 30% ratio of disutility over average male productivity. On the most recent year available, 2004, I estimate that the average productivity differential decreases to about 13%, while the parameters that describe prejudiced are not significantly different from 1995.

However, to judge the real impact of these parameters on observables and welfare, we have to consider equilibrium effects. This will be the main focus of the remaining sections of the paper: first, by simply computing an earning differential decomposition and then by performing some policy experiments. Before this, an assessment of the fit of the model is provided.

---

[42]Bowlus and Eckstein 2002 also assume employers taste discrimination, but in a different model (homogeneous equilibrium search), on different data (a sample of high school graduates from NLSY) and implementing a different estimator (matching moments).

6.2   Fit and Specification Test

The bottom line of Table 2 reports the likelihood ratio specification test of any specification against the one in column (6). All the restrictions are rejected, leading to the conclusion that a model with only productivity differences or only prejudice performs worse than a model in which both are present. This conclusion is also confirmed by testing model (1) and (2) against model (3). Specification (6) is the one used to discuss the main implications of the model and to perform the policy experiments.

A first sense of how well specification (6) fits the data is obtained by looking at fit on first moments, reported in Tables 3 and 4, and by comparing the predicted and empirical density of accepted earnings, reported in Figures 1 and 2. The last two columns of Table 3 show that the fit on first moments is very good on unemployment variables and on female earnings while the model slightly overestimates average accepted male earnings. The last two rows of Table 4 compare the women/men ratio of average earnings, the measure commonly used to describe gender wage differentials. The ratios are computed on averages over the entire distribution, column 1, and over the bottom and top 25% quantiles, columns 2 and 3. The fit is quite good on the entire distribution since the model generates a ratio of 78.5%, quite close to the 79.4% of the sample. The model is also able to generate the relative gain of female earnings as we move up on the distribution: the top 25% has an higher ratio than the bottom 25%. Note that a model with only productivity differences is not able to generate this ranking, implying almost no difference in the bottom 25% and a significant differential in the top 25%.

Figures 1 and 2 compare the histogram of observed earnings with the densities predicted by the model: the fit is good, in particular on female earnings. The less precise fit on male earnings is mainly due to the three mass points on the right tail of the distribution generated by topcoding on the original data. In terms of shape of the distribution, the estimated model is able to generate the crucial features of the data that were discussed in the identification, in particular the flatter low tail of the predicted female distribution.

6.3   Earnings Differential Decomposition

The empirical literature on discrimination often employs wage regressions to decompose the observed wage differentials into different components, including discrimination or proxies for it. In a typical wage regression decomposition, estimated returns on productivity characteristics are separated from average endowments of these characteristics. The part of the wage differential due to difference in returns is often interpreted as discrimination.[43] Based on the estimates in Table 3 and the model used to generate them, it is possible to perform a similar wage differential decomposition with the advantage that the equilibrium effects of the counterfactual experiments can be taken into account.

---

[43]For surveys and references, see Blau and Kahn 2000 and Altnoji and Blank 1999.

Table 4 presents the results of this decomposition. The table shows ratios of female average earnings over male average earnings. Each cell of each column reports the contribution of the corresponding component. For example, the first cell in the first column, labelled *productivity,* reports the wage differential that would arise from an environment in which men and women are identical except for productivity and there is no prejudice. Conditioning on this environment, a new equilibrium is generated and earning distributions for each type of worker are obtained. The implied earning differential, as expressed by the ratio of the average female accepted earnings over the average male accepted earnings, is the value reported in the cell. The interpretation is that the implied earning ratio can be considered the contribution of differences in productivity to the overall observed differential, once equilibrium effects are taken into account. For comparison purposes, the last raw reports the ratio observed in the estimation sample. Results of the exercise are reported for ratios computed on the entire sample (first column) and on the bottom and top 25% quantiles (second and third column.) The parameter estimates utilized to perform these experiments are specification (6) in Table 3. When parameters are set equal for both groups, they are set at male values.

Focusing first on the entire distribution, results show that all groups of variables, except *behavior*, contribute to the negative female earning differential. The parameters labelled as behavior are the arrival rate of offer and the termination rate of the job contract. They favor women because the positive impact of the higher arrival rate more than compensate the negative impact of the higher termination rate. *Prejudice* generates the highest differential, even if below the value observed in the data. As expected, the lower estimated female *productivity* has also a significant negative impact. To summarize the results on the full sample: if all the difference between men and women was due to differences in productivity, the differential should be about 8%; if all the difference was due to prejudice, the differential should be about 18%; finally, when all the ingredients are blended together a differential of about 21% is generated, closely matching the one observed in the data.

Looking at results on ratios of quantiles we see that differences in productivity imply no negative differential for women at low levels of the distributions but a substantial differential as we move up in the accepted earnings distribution. On the contrary, prejudice implies that the differential decreases as wages increases. This implication matches the data[44] (last row) and it is not consistent with the so called *glass ceiling* hypothesis.[45] Women are in a relatively better position at the top of the distribution than at the bottom because the proportion of women working for prejudiced employers is lower at the top than at the

---

[44]CPS data suffers from top-coding and top-coding affect male wages in higher proportion. The amount of top-coding on the estimation sample, though, seems too small to fundamentally bias moments computed on the top 25% quantile.

[45]The glass ceiling hypothesis states that the most important asymmetry between men and women in the labor market is the low proportion of women that reach high level - high paying jobs. Albrecht, Bjorklund and Vroman 2003 also find that a *glass ceiling* does not seem the main determinant of the gender wage gap in the U.S.

bottom and women working for prejudiced employers are the ones suffering the largest wage discrimination.

These results on ratios of quantiles are also another way to look at the identification. They show that the crucially different implication of productivity with respect to prejudice is not on the conditional mean but on the shape of the distribution: productivity differences predict an higher differential as we move up in the distribution while prejudice predicts the opposite.

Looking at the ratio of reservation wages, reported in the last column, prejudice is the component that has the strongest impact on the negative female differential, lowering the female/male ratio from about 0.8 to about 0.57. Behavior, instead, would predict an higher outside option for women and productivity would predict a ratio only slightly higher than the one observed in the data.

7 Welfare Measures and Policy Experiments

The gender discrimination literature has devoted a lot of attention to average wage differentials. The overall welfare of labor market participants, though, is not only dependent on average wages but on the entire wage distribution and on the dynamic of the labor market, such as the transition probabilities between states and the durations in each state. It is therefore useful to define indicators that may give a more complete description of the overall workers' welfare taking into account at least some of the labor market dynamic. With these indicators in hand, it will then be possible to evaluate the welfare impact of some policy experiments.

7.1 Welfare Measures

The proposed welfare measures exploit the steady state equilibrium results of the model.[46]

In steady state workers occupy all the possible equilibrium states: the unemployment state and the employment state at each acceptable match value x. An average welfare measure should associate a value at each of these states and then weigh them according to some meaningful measure. To clarify the discussion, define the following labor market environment $\Gamma \equiv (\lambda, \eta, \rho, b, \alpha, G(x))$. This is the environment that generates the steady state equilibrium derived in section 3. A function that assigns welfare values to states is:

$$T(x) = U\left[1 - \mathbf{I}_{\{x \geq x^*\}}\right] + W\left[w(x,U)\right]\mathbf{I}_{\{x \geq x^*\}} \tag{28}$$

where U is the value of unemployment, $W[w(x,U)]$ is the value of employment at the wage that corresponds to the match value x and x^* is the reservation match value. This function assigns the value of unemployment to all the unemployed individuals ($x < x^*$) and the value of being employed at a wage $w(x,U)$ to all the employed individuals ($x \geq x^*$) .

---

[46]For a discussion and interpretation of this and other welfare measures in the context of a similar search model with minimum wage, see Flinn 2005 and Flinn 2002b.

A meaningful weighting function for $T(x)$ is the ex-post distribution of types in the population. Define with H the corresponding cumulative distribution function, then: $H(x) = H(x|u)P(u) + H(x|e)[1 - P(u)]$, where u indicates the state of unemployment, e the state of employment and $P(e) = [1 - P(u)]$. The distributions conditioning on the state are:

$$H(x|u) = \frac{G(x)}{G(x^*)}, x < x^*$$

$$H(x|e) = \frac{G(x) - G(x^*)}{\widetilde{G}(x^*)}, x \geq x^*$$

and the steady state probability to be in the unemployment state is:

$$P(u) = \frac{\eta}{\eta + h}$$

where $h = \lambda\widetilde{G}(x^*)$ is the hazard rate out of unemployment. The proposed welfare measure is therefore the average of the value of each state taken over the equilibrium distribution of types. Formally:

Definition 8 *The **average welfare** measure for workers in an environment $\Gamma \equiv (\lambda, \eta, \rho, b, \alpha, G(x))$ is defined as:*

$$E_H[T(x)] = \int_0^{x^*} U\frac{\eta}{\eta + h}\frac{g(x)}{G(x^*)}dx + \int_{x^*}^{+\infty} W[w(x, U)]\frac{h}{\eta + h}\frac{g(x)}{\widetilde{G}(x^*)}dx \quad (29)$$

This welfare measure is presented for simplicity in an homogenous environment. When heterogeneity is introduced, expression (29) must be specialized for each type of worker.

For men, in presence of heterogeneity, the average welfare will simply be (29) in an environment equal to $\Gamma_M \equiv \left(\widehat{\lambda}_M, \widehat{\eta}_M, \overline{\rho}, \widehat{b}_M, \overline{\alpha}, \widehat{G}_M(x)\right)$, where the estimated values correspond to the estimates of specification (6) in Table 2. For women, computation of $E_H[T(x)]$ under this specification is slightly more complicated since the differential impact of the two types of employers must be taken into account. The solution is highly simplified by the fact that the proportion of prejudiced and unprejudiced employers is fixed and does not depend on x. Therefore the welfare measure will simply be a linear combination of two quantities obtained by plugging in equation (29) the reservation values and wage schedules of women at the two types of employers in an environment $\Gamma_W \equiv \left(\widehat{\lambda}_W, \widehat{\eta}_W, \overline{\rho}, \widehat{b}_W, \overline{\alpha}, \widehat{G}_W(x), \widehat{d}, \widehat{p}\right)$.

To define employers' welfare, it is useful to recall the following. The model assumes search frictions, match specific productivity and bargaining. This characterization leads to positive profit and, given the assumption on preferences, to positive utility for each type of employer. A reasonable measure of welfare is then the steady state value of this quantity.

Denote the number of workers with N and the proportion of men with m; the number of employers with K and the proportion of prejudiced employers

with p. Then, the steady state number of male and female workers employed at each employer type are:

$$E_{MI} = \frac{h_{MI}}{\eta_M + h_M} mN \qquad ; \quad I = N, P \qquad (30)$$

$$E_{WI} = \frac{h_{WI}}{\eta_W + h_W}(1 - m)N$$

where h_{JI} denotes the hazard rate for a type J from being unemployed to being employed at an employer of type I. The average instantaneous utility per worker at each employers/workers match are:

$$AP_{JN} = \int_{\rho U_J} [x - w(x, U_J)] \frac{g_J(x)}{\widetilde{G}_J(\rho U_J)} dx \qquad ; \quad J = M, W$$

$$AP_{JP} = \int_{\rho U_J + d\mathbf{I}_{\{W\}}} \left[x - d\mathbf{I}_{\{W\}} - w\left(x, U_J, d\mathbf{I}_{\{W\}}\right)\right] \frac{g_J(x)}{\widetilde{G}_J\left(\rho U_J + d\mathbf{I}_{\{W\}}\right)} dx$$

Following the previous intuition, the proposed welfare measure is the average of the per-worker utility value times the proportion of that type of workers hired in steady-state.

Definition 9 *The **average welfare** measures for employers of types N, P in an environment $\Gamma \equiv (\lambda, \eta, \rho, b, \alpha, G(x), p, d)$ are:*

$$\Pi_N = AP_{MN}\frac{E_{MN}}{(1-p)K} + AP_{WN}\frac{E_{WN}}{(1-p)K} \qquad (31)$$

$$\Pi_P = AP_{MP}\frac{E_{MP}}{pK} + AP_{WP}\frac{E_{WP}}{pK}$$

The focus of interest will be the ratio Π_P/Π_N, so missing information on the number of employers K is innocuous.

The first column of Table 5 reports the agents' welfare measures when prejudice and productivity differences are jointly estimated. This specification is called Benchmark Model and corresponds to the estimates reported in column (6) of Table 2. The workers' welfare values are normalized with respect to the male value and the employers' welfare values are normalized with respect to the unprejudiced employers value. The lower panel in Table 5 shows some labor market variables related to workers' welfare.

On the workers' side, as expected, average welfare is about 24% lower for women than for men. This relative disadvantage is higher than the gap indicated by the indicators considered so far. For example, looking at the bottom panel of Table 5, average accepted earnings are about 21.5% lower for women than for men and the reservation wage about 19.9% lower. The ratio of reservation wages is interesting because it is also equal to the ratio of values of unemployment, that is the ex-ante value of participating in the labor market. On the employers' side, unprejudiced employers are better off than prejudiced employers. This result

was also somewhat expected: prejudiced employers pay in part their distaste for women but search frictions still guarantee them positive profit and utility.

Since both productivity differences and prejudice contribute to the worse performance of women in the labor market, it is interesting to isolate the impact of these two components on welfare. A rationale for this exercise can be to separate the impact of some pre-labor market factors - factors that are more likely to affect the unobserved productivity represented by the $G(x)$ distribution - from some specific labor market factors - such as the presence of prejudiced employers. The experiment is similar to the one reported in Table 4 but it focuses on the overall welfare impact instead of the simple wage differential.

To implement this policy it is necessary to obtain a new steady state equilibrium in which women are given the same productivity parameters of men. In terms of the previous notation, this experiment is equivalent to assuming an environment $\Gamma_W \equiv \left(\widehat{\lambda}_W, \widehat{\eta}_W, \overline{\rho}, \widehat{b}_W, \overline{\alpha}, \widehat{G}_M(x), \widehat{d}, \widehat{p}\right)$. The second column of Table 5 reports the corresponding welfare measures: women and both types of employers are better off while men's welfare is unchanged. As expected, the result is a Pareto improvement because we are exogenously increasing the productivity of a significant proportion of the worker's population. However, the difference in welfare between men and women is still sizeable, with female welfare about 10% lower than male welfare.

7.2 Equal Pay Policy

An Equal Pay policy is defined as a policy that imposes restrictions on the wage determination with the objective of equalizing differentials among groups. In this exercise, I will define the Equal Pay policy as requiring each employer to pay the same wage to workers with identical productivity.

The policy rises some clear problems of enforcement. The main issue is that an external agent, such as the public authority responsible of enforcing the policy, cannot directly observe the match-specific value of productivity and the measures used to proxy this productivity are often quite limited. Given that in practice the policy states that wages cannot be set conditional on appearance, a rule probably easier to enforce would be require that gender cannot be observed when wages and hiring are decided. This practice is theoretical appealing but its application has been very limited.[47] Moreover, it generates a slightly different policy because it directly reduces the amount of information available to employers.[48] Conditioning on these enforcement issues, I have anyway chosen to implement a very standard equal pay policy: each type of employer can

---

[47]An interesting example are the blind auditions to hire musicians implemented by some of the major US orchestras. Goldin and Rouse 2000 estimate that the introduction of blind hiring in the 1970's and 1980's explain about 30% of the big increase in the female proportion among musicians working at these orchestras.

[48]This also introduce some problems of comparability because the resulting asymmetric information bargaining problem has not an easy solution that generalizes the axiomatic Nash bargaining result used in the benchmark model. The problem is that employers will not know the outside option of the other agent, leading to delays difficult to embed in a dynamic search model. For a review, see for example Ausubel, Cramton and Deneckere 2002.

offer only one wage to workers with same productivity. Interpreting the Nash bargaining outcome as a reduced form sharing rule, offered wages are simply an average of the wages that would be offered without the policy. The average is over the expected probability that the worker is male or female and this probability is simply their (common knowledge) proportion in the population.

Using equation (5), we get the following wage schedules for each type of employer in presence of the policy:

$$w_N(x) = \alpha x + (1 - \alpha) [m\rho U_M + (1 - m) \rho U_W] \tag{32}$$

$$w_P(x) = \alpha x + (1 - \alpha) [m\rho U_M + (1 - m) \rho U_W] - \alpha (1 - m) d \tag{33}$$

Equation (32) is the unique wage, conditioning on x, paid at the unprejudiced employer, where m denotes the proportion of men in the population. Equation (33) is the corresponding expression for prejudiced employers: it shows that in a post-policy environment men also pay the cost of prejudice, given by the term $\alpha (1 - m) d$. This cost is increasing in the disutility coefficient, in the workers' bargaining power coefficient and in the proportion of women in the population.

This different wage determination has an impact on the equilibrium, generating the following new reservation values for each type of worker $J = M, W$:

$$
\begin{aligned}
x_{JN}^* &= \frac{1}{\alpha} \{\rho U_J - (1 - \alpha) [m\rho U_M + (1 - m) \rho U_W]\} \\
x_{JP}^* &= x_{JN}^* + (1 - m) d
\end{aligned}
$$

and the following values of unemployment for each type of worker $J = M, W$:

$$\rho U_J = b_J + \frac{\alpha \lambda_J}{\rho + \eta_J} \left\{ p \int_{x_{JP}^*} [x - x_{JP}^*] \, dG_J(x) + (1 - p) \int_{x_{JN}^*} [x - x_{JN}^*] \, dG_J(x) \right\}$$

The sign of the impact of the policy on the reservation values and values of unemployment is ambiguous and depends on parameter values. The reason is that men and women are not only heterogenous with respect to prejudice but also in terms of productivity and transition rates between states. At the estimated values, though, it is possible to compute the sign and the magnitude of these changes: results are reported in the third column of Table 5.

Female welfare increases by about 11.2% with respect to the benchmark model, while male welfare decreases by about 7.5%. Still, a 9% gender gap remains due to the lower productivity of women. The improvement in women's welfare is the result of a big increase in the value of the outside option - a 28.1% increase - that implies an increase in earnings at any productivity level. The increase is especially high at prejudiced employers where it is equal to 4 dollar an hour. In sum, results with respect to the benchmark model are: the male reservation value increases at prejudiced employers and decreases at unprejudiced employers, while male earnings decrease at both types of employers; the female reservation value and earnings increase at both types of employers; earning differentials are still present because women have lower productivity but they are significantly reduced.

The drawback of the policy is the large reduction in employers' welfare. This was expected since an additional constraint is imposed, but the size of the decrease depends on parameters. The decrease in average welfare with respect to the benchmark model is about 17% for unprejudiced employers and about 15.7% for prejudiced employers. This is due mainly to the increase in the outside option of women that eliminates the possibility of wage discriminating against them.

7.3 Affirmative Action Policy

An Affirmative Action policy is an anti-discrimination policy that requires pro-active steps (Holzer and Neumark 2000). In the economic literature, this broad definition is very often limited to that of a *quota system*.[49] A quota system is a specific policy where a system of numerical yardsticks for minority in hiring, federal contracts or school enrollment are exogenously imposed. The quota system definition was not explicitly used in the legislation that first introduced affirmative action in the U.S.[50] but it is commonly used by government agencies as a result of the Department of Labor regulations implementing the legislation.[51]

The difference between a quota system and a more general definition of affirmative action is considered crucially important by Holzer and Neumark 2000 in their review assessing the impact of affirmative action in the labor market. Donohue and Heckman 1991 also use a broad definition of affirmative action to conclude that the impact of Civil Rights policies were effective in improving the relative performance of blacks in the labor market in the late 1960s and early 1970s. A quota system, though, has the advantage of a clear quantitative implementation and some authors consider the difference between affirmative action and a quota system as merely semantic (Moro and Norman 2003). The difference between the two definitions is considered very relevant by the recent Supreme Court opinion about affirmative action.[52] The Court has considered admissible an affirmative action policy of the University of Michigan Law School but it is very careful in interpreting affirmative action not as a quota system but as "a narrowly tailored plan system" in which "race or ethnicity" may be

---

[49] Welch 1976 is one of the first to introduce and calibrate a model to study affirmative action and he defines the policy as a quota sytem. Other and more recent contributions using a quota system definition are: Coate and Loury 1993 and Moro and Norman 2003.

[50] This legislation starts with the 1961 Kennedy Executive Order #10925 that mandates "affirmative action" to avoid discrimination by race in the labor market. The 1965 Johnson Executive Order #11246 reiterates the Kennedy executive order and the 1967 Johnson Executive Order #11375 extends its application to cover women.

[51] The 1968 Department of Labor Regulations governing the Johnson executive orders require explicitly to identify "underutilization" of women and minority. The quota system is also the current definition implemented by the Equal Employment Opportunity Commission (EEOC). The EEOC is the Commission responsible to enforce all the federal statutes prohibiting discrimination.

[52] Grutter v. Bollinger and Gratz v. Bollinger, June 24, 2003. Both rulings are related to affirmative action policies at the University of Michigan but they are though to have a strong impact also on the labor market through the affirmative action policies that most of U.S. corporations implement.

considered "a 'plus' in a particular applicant's file."[53]

Whether a quota system has a significant impact in the model under consideration or not is mainly an empirical question. By looking at matching rates we expect that prejudiced employers will hire proportionally less women than unprejudiced employers but by Corollary 7 we know that a positive proportion of women will always be hired by both types of employers. Therefore, we have to look at the actual proportions implied by the estimated values to see if imposing a quota will be effective. The estimated parameters imply that the steady state proportions of women working at a prejudiced and unprejudiced employer are respectively 43.7% and 46.9%, out of a population where women count for 45.6% of the labor force. In this context, any quota policy that imposes a minimum proportion of women to be hired lower than 43.7% will have no impact. This level is quite high if compared to the usual level enforced by the Equal Employment Opportunity Commission (EEOC) and it shows an environment in which the numerical yardstick is met but prejudice and its impact are unaffected.

Since at the estimated values a quota policy has a very limited impact, it may be interesting to focus on an affirmative action policy that does not explicitly use quotas and that may capture the main indications of the Supreme Court rulings. An affirmative action policy defined as a subsidy received by employers for hiring women may constitute a crude model for such a policy. More precisely, assume that employers receive a flow subsidy γ for each woman employed, for all the time the employment relation lasts. The subsidy is paid by a lump-sum tax t on all the workers. This policy is a pro-active policy quite easy to implement and enforce. In spirit, it is similar to policy interventions that create incentives to hire first-seekers by lowering minimum wage requirement or other job related costs. It is also observational equivalent to other affirmative action policies implemented in practice, such as the 'plus-factor' idea supported by the Supreme Court ruling. Moreover, a subsidy policy is particularly interesting in the context of this model because the impact of the subsidy is magnified by the spillover effects, now working in favor of women.

The policy affects profits and wages schedules. Defining with γ the exogenously fixed employers' subsidy, the employers' utility will be:

$$
\begin{aligned}
\pi_{WI} &= x - d\mathbf{I}_{\{W,P\}} - w + \gamma, \quad I = N, P \\
\pi_{MI} &= x - w
\end{aligned}
\tag{34}
$$

The workers' utility will be equal to the wage net of the tax:

$$w_{JI}(x, U_J) - t(\gamma), \quad J = W, M; I = N, P$$

where $t(\gamma)$ is the endogenously determined lump-sum tax. The tax level t depends on all the parameters of the model but to simplify notation I simply emphasize the dependence on γ. The wage determination does not change.

---

[53]Excerpts from Justice O'Connor majority opinion on Grutter V. Bollinger (Law School Case), June 24, 2003. In the same opinion the Court explicitly states that "a race-conscious admission program cannot use a quota system."

Wages are determined by Nash bargaining upon observing types and productivity, leading to the following wage schedules:

$$w_{MI}(x, U_M, t) = \alpha x + (1 - \alpha)\left[t(\gamma) + \rho U_M\right], \quad I = N, P \qquad (35)$$
$$w_{WI}(x, U_W, t) = \alpha\left[x + \gamma - d\mathbf{I}(P)\right] + (1 - \alpha)\left[t(\gamma) + \rho U_W\right]$$

The reservation values that determine the decision rules[54] are:

$$x_{MI}^* = \rho U_M + t(\gamma), \qquad\qquad I = N, P \qquad (36)$$
$$x_{WI}^* = \rho U_W + t(\gamma) - \gamma + d\mathbf{I}(P)$$

from which we can obtain the equilibrium values of unemployment $(U_W(\gamma), U_M(\gamma))$. The (instantaneous) value of the tax t is determined by equating the total subsidy to the total tax. Formally, t is implicitly determined by the following equation:

$$t(\gamma) = \gamma \frac{\frac{h_W[t(\gamma),\gamma]}{h_W[t(\gamma),\gamma] + \eta_W}(1 - m)}{\left[\frac{h_W[t(\gamma),\gamma]}{h_W[t(\gamma),\gamma] + \eta_W}(1 - m) + \frac{h_M[t(\gamma),\gamma]}{h_M[t(\gamma),\gamma] + \eta_M}m\right]}$$

where the hazard rates $h_J[t(\gamma)]$ have the usual form and depend on $(t(\gamma), \gamma)$ through the reservation values (36).

The subsidy is paid by both men and women but only benefits women. Then, we expect the value to participate in the labor market to increase for women and decrease for men once the policy is implemented. This result is stated in the following Proposition.[55]

Proposition 10 *For any positive subsidy γ previously defined, the women's outside option increases and the men's outside option decreases, i.e.*

$$\frac{\partial U_W(\gamma)}{\partial \gamma} > 0$$
$$\frac{\partial U_M(\gamma)}{\partial \gamma} < 0$$

Even if by Proposition 10 we know the impact of the policy on the values of unemployment, we cannot sign the impact on the reservation values (36). This is due to the fact that both the values of unemployment and the lump-sum tax are endogenous and they do have an ambiguous impact on x_{MI}^* and x_{WI}^* as we increase the subsidy γ. What is unambiguous, instead, is the impact on the wage schedules: men's wages will be lower at any productivity values and women's wages will be higher. With respect to a pre-policy environment,

---

[54]It is possible that, if the subsidy is high enough, the expressions $\left(x_{WN}^*, x_{WP}^*\right)$ becomes negative. If this is the case, there is no truncation and all the matches are acceptable because the support of x is R_+. I ignore this case in the text because I will only consider small enough subsidies.

[55]The proof is in Appendix 9.1.

women's earnings, conditioning on same x and same employer type $I = N, P$, increase by:

$$w_{WI}(x, \gamma) - t - w_{WI}(x) = \alpha (\gamma - t) + (1 - \alpha) [\rho U_W(\gamma) - \rho U_W] > 0$$

where the amount is positive since $\gamma > t$ and $\rho U_W(\gamma) > \rho U_W$. The second term on the RHS is the spillover effect, now favoring women and magnifying the effect of the policy: women get higher wages not simply because they are the only beneficiary of a subsidy that men also pay (the $(\gamma - t)$ term) but also because their bargaining position has improved (the $[\rho U_W(\gamma) - \rho U_W]$ term.) For men the opposite is true: their lower outside option induced by the lump-sum tax reinforces the decrease in earnings.

The policy implemented on the estimated model sets the subsidy at one dollar. This is a reasonably large subsidy because it corresponds to one dollar an hour more for each women employed for all the time the employment relation lasts. It also corresponds to about 10% of the disutility parameter. Results are reported in the last column of Table 5. The net increase in women's earning is 42 cents an hour at any level of x. In terms of equilibrium effects, 64% of the impact is due to sharing the tax with men and 36% is due to the spillover effect. Once again, spillover effects play a significant role.

Women's welfare increases by about 1.8 percentage points and men's welfare decreases by about 1.5 percentage points. A welfare gap remains even after this relatively generous subsidy: the reason is not simply a difference in productivity but also the presence of prejudice that still has an impact. In particular, wage discrimination is still present at prejudiced employers but it is not at unprejudiced employers. Conditioning on same productivity, women's earnings are 6.7 dollars lower than men's earning at prejudiced employers and 4 cents higher than men's earning at unprejudiced employers.

Employers' welfare is almost unaffected because both a positive and a negative impact are present. The positive impact is due to the presence of the subsidy and the lower outside option for men which allows employers to pay them lower wages; the negative impact is due to the presence of the tax that increases the reservation values and the higher outside option for women which reduces wage discrimination. This policy has also an impact on the quota of women hired at the two types of employers: prejudiced employers hire an higher proportion of women than in the pre-policy setting. As expected, there is a tendency to converge to the proportion of men and women in the labor force but changes are very small.[56]

To summarize, this affirmative action policy implies a redistribution of welfare from men to women, leaving employers' welfare almost unaffected. The impact is relatively modest for a subsidy of one dollar an hour but it implies that wage discrimination is eliminated at unprejudiced employers.

---

[56] The changes in the proportion of women hired by both employers' types are in the order of 0.002 percentage points.

8 Conclusion

By developing a search model of the labor market with matching, bargaining and employers' taste discrimination, this paper shows that it is possible to separately identify and estimate gender discrimination and unobserved productivity differences. The result is obtained by exploiting the markedly different shape of the accepted earnings distribution of women with respect to the one of men. The crucial assumptions are a standard parametric assumption on the productivity distribution and a parsimonious parametrization of explicit prejudiced behavior. Prejudice is summarized by the disutility that a proportion of employers receives when hiring women. The bargaining setting generates *spillover effects*: the presence of some prejudiced employers lowers women's outside option with respect to men, generating wage discrimination also at unprejudiced employers. These effects have been neglected by the previous literature and in this paper they are estimated to have a sizable impact.

Maximum likelihood estimates on Current Population Survey data show that both discrimination and productivity differences are present in the labor market for white College graduates. Average female productivity is estimated to be about 6.5% lower than male productivity and the extent of explicit prejudice is significant: about half of the employers are prejudiced and the disutility from hiring a woman is about one third of the average male productivity. These values may overestimate the presence of explicit prejudice for the following reasons: the employers side of the model is highly stylized and labor market participation is exogenous; homogeneity controls are limited to race, human capital and age; the higher intermittency of women in the labor market is not explicitly modelled. Still, some crucial and peculiar features of the women's labor market dynamic, as summarized by the accepted earnings distribution and unemployment durations, are well described by the model and lead to rather precise and robust estimates.

The estimated structural parameters allow us to decompose the observed earning differential and to perform policy experiments. The earning differential decomposition shows that prejudice is the most important factor in generating the 20.6% difference between average accepted male and female earnings. If the difference between man and women were due exclusively to prejudice, we should observe about 2/3 of this gap; if the difference were due exclusively to productivity, only about 1/3 of this gap would remain. Both results are based on counterfactual experiments that take into account equilibrium effects.

Two policy experiments are performed: an Equal Pay policy and an Affirmative Action policy. The equal pay policy requires employers to offer the same wage to men and women with same productivity. The policy increases female welfare more than it decreases male welfare but imposes an heavy welfare cost on employers. An affirmative action policy implemented as a strict quota policy is shown to be not binding. More in line with the recent Supreme Court ruling,[57] an alternative pro-active policy is implemented. The policy is defined as

---

[57]I am referring to Grutter v. Bollinger and Gratz v. Bollinger, June 24, 2003. The Court has considered admissible an affirmative action policy of the University of Michigan Law School but it is very careful in interpreting affirmative action not as a quota system.

an employer's subsidy for hiring women and implies a redistribution of welfare from men to women, leaving employers' welfare almost unaffected.

9 Appendix

9.1 Proofs of Propositions and Corollaries

Proposition (2)

Proof. Rewrite the reservation equations (6) as:

$$x^*(d,p) = b + \frac{\lambda\alpha}{\rho+\eta} \left\{ p \int_{x^*(d,p)+d}^{+\infty} \widetilde{G}(x)dx + (1-p) \int_{x^*(d,p)}^{+\infty} \widetilde{G}(x)dx \right\} \quad (37)$$

where: $\rho U_W \equiv x^*(d,p)$ for $0 < p < 1$ and $d > 0$; and $\rho U_M \equiv x^*(0,0)$. I have also exploited integration by parts to rewrite: $\int_{x^*}[x - x^*]dG(x) = \int_{x^*} \widetilde{G}(x)dx$. Given the definition of ρU_W and ρU_M, and given that the discount rate ρ is always positive, to prove the proposition is enough to show that:

$$\frac{\partial x^*(d,p)}{\partial d} < 0 \quad \text{if} \ \ 0 < p < 1 \quad (38)$$

and

$$\frac{\partial x^*(d,p)}{\partial p} < 0 \quad \text{if} \ \ d > 0 \quad (39)$$

To obtain the first claim, total differentiate (37) with respect to d:

$$\frac{\partial x^*(d,p)}{\partial d} = \frac{-\frac{\lambda\alpha}{\rho+\eta}p\widetilde{G}(x^*(d,p)+d)}{1 + \frac{\lambda\alpha}{\rho+\eta}\left[p\widetilde{G}(x^*(d,p)+d) + (1-p)\widetilde{G}(x^*(d,p))\right]}$$

and observe that all the parameters are positive, the survival function assumes only positive values and the numerator is negative due to the minus sign.

To obtain the second claim, total differentiate (37) with respect to p:

$$\frac{\partial x^*(d,p)}{\partial p} = \frac{-\frac{\lambda\alpha}{\rho+\eta}\left[\int_{x^*}^{x^*(d,p)+d}\widetilde{G}(x)dx\right]}{1 + \frac{\lambda\alpha}{\rho+\eta}\left[p\widetilde{G}(x^*(d,p)+d) + (1-p)\widetilde{G}(x^*(d,p))\right]}$$

and observe that all the parameters are positive, the survival function assumes only positive values and again the numerator is negative due to the minus sign.
∎

Corollary (6)

Proof. By wage schedules (4) and (5) we know:

$$\begin{aligned} w_{WP}(x) - w_{MP}(x) &= -\alpha d + (1-\alpha)\rho(U_W - U_M) < 0 \\ w_{WN}(x) - w_{MN}(x) &= (1-\alpha)\rho(U_W - U_M) < 0 \end{aligned}$$

38

where both differentials are negative independently from x because $d > 0$ and, by proposition (2), $U_W < U_M$. ∎

Corollary (7)
Proof. Men are indifferent between working for the two types of employers and therefore they will work for both as long as both are present. Women ex-ante prefer to work for unprejudiced employers but once they meet an employer of any type they will accept to enter the match if the wage is high enough. In equilibrium, the proportion of women working for unprejudiced employers is given by the ratio of the hazard rates:

$$P_{WN} = \frac{(1-p)\,\widetilde{G}\,(\rho U_W)}{(1-p)\,\widetilde{G}\,(\rho U_W) + p\widetilde{G}\,(\rho U_W + d)}$$

where $\widetilde{G}\,(x)$ denotes the survival function $[1 - G\,(x)]$. Complete segregation arises only if:

$$P_{WN} = 1 \Longleftrightarrow p = 0 \text{ since } U_W < +\infty$$
$$P_{WN} = 0 \Longleftrightarrow p = 1 \text{ since } U_W > 0$$

therefore no complete segregation arises when $0 < p < 1$. Partial segregation of women in the unprejudiced sector arises if:

$$P_{WN} > \frac{1}{2} \Longleftrightarrow \frac{1-p}{p} > \frac{\widetilde{G}\,(\rho U_W + d)}{\widetilde{G}\,(\rho U_W)}$$

that is if the proportion of prejudiced employers is not so high to offset the lower acceptance probability induced by the higher reservation value. Partial segregation in the prejudiced sector arises if the opposite is true. ∎

Proposition (10)
Proof. By total differentiating with respect to γ the female value of unemployment, we get:

$$\rho\frac{\partial U_W}{\partial \gamma} = \frac{\lambda\alpha}{\rho+\eta}\{-p\widetilde{G}(\rho U_W + t - \gamma + d)\frac{\partial(\rho U_W + t - \gamma + d)}{\partial \gamma}$$
$$-(1-p)\widetilde{G}(\rho U_W + t - \gamma)\frac{\partial(\rho U_W + t - \gamma)}{\partial \gamma}\}$$

collecting terms:

$$\rho\frac{\partial U_W}{\partial \gamma}\left\{1 + \frac{\lambda\alpha}{\rho+\eta}\left[p\widetilde{G}(\rho U_W + t - \gamma + d) + (1-p)\widetilde{G}(\rho U_W + t - \gamma)\right]\right\}$$
$$= -\frac{\lambda\alpha}{\rho+\eta}\left(\frac{\partial t}{\partial \gamma} - 1\right)\left\{p\widetilde{G}(\rho U_W + t - \gamma + d) + (1-p)\widetilde{G}(\rho U_W + t - \gamma)\right\}$$

every terms is positive except:

$$\left(\frac{\partial t}{\partial \gamma} - 1\right) < 0$$

This term is negative because the tax is on both men and women so that the increase in t is always smaller than the increase in γ. This negative term cancels out with the minus sign of the RHS and we get the result. By total differentiating with respect to γ the female value of unemployment, we get:

$$\rho \frac{\partial U_M}{\partial \gamma} = \frac{\lambda \alpha}{\rho + \eta} \left\{ -\widetilde{G}(\rho U_M + t) \frac{\partial (\rho U_M + t)}{\partial \gamma} \right\}$$

collecting terms:

$$\rho \frac{\partial U_M}{\partial \gamma} \left[1 + \frac{\lambda \alpha}{\rho + \eta} \widetilde{G}(\rho U_M + t) \right] = -\frac{\lambda \alpha}{\rho + \eta} \widetilde{G}(\rho U_M + t) \left[\frac{\partial t}{\partial \gamma} \right]$$

where the claim is proven because all the terms are positive and there is a minus sign in front of the RHS. ∎

9.2 Data Appendix

The estimation sample is extracted from the *Annual Social and Economic Supplement* (March Supplement) of the Current Population Survey (CPS) for the year 1995. The raw data files were provided by Unicon. The CPS is organized around monthly interviews with different content. The March survey focuses on work experience, income sources and amounts, noncash benefits, health insurance, and migration and it is the one that allows for the construction of unemployment durations. Information about weekly and hourly pay is collected each month on a random subset of respondents. They constitute the *Earner Study* and they are individuals in the last month of their four-month participation period (i.e. they are in *rotation groups* 4 and 8).

This background information helps understand the extraction process presented in Table A.1. We first have to consider only the relevant individuals: individuals in the labor force with recorded wages if employed or recorded unemployed durations if unemployed. Then an homogenous sample with respect to some demographic and human capital characteristics is selected. Finally, some data cleaning leads to a final sample of 2,324 observations. Thanks to the high number of observations in the raw sample, the estimation sample after the homogeneity controls still contains more than one thousand observations for both men and women.

9.3 Derivation of Likelihood Contributions

9.3.1 Model without prejudice

The subscript J denotes the worker's type: $J = W, M$. The time homogeneity of the environment, the Poisson process that governs the arrival of job offers and

the optimal decision rule imply a constant hazard rate out of unemployment. The hazard rate is given by the probability to meet an employer times the probability to accept the match:

$$h_J = \lambda_J \widetilde{G}_J(\rho U_J) \tag{40}$$

This hazard function uniquely determine the distribution of complete unemployment durations: it is exponential with parameter equal to the hazard rate. The corresponding density function is:

$$f_c(t_i|J) = h_J \exp(-h_J t_i), \quad t_i > 0 \tag{41}$$

Unemployment durations in the sample have two limitations: they are the time in unemployment up to the sampling date (on-going unemployment durations) and they are observed only for individuals currently unemployed. Since the distribution of complete spells is exponential, on-going spells are also exponential:[58]

$$f_u(t_i|i\epsilon U, J) = h_J \exp(-h_J t_i), \quad t_i > 0 \tag{42}$$

The intuition is that the underestimation due to right censoring is compensated by the overestimation due to length bias. The second limitation is taken into account using ergodic results to weight the density by the probability of being unemployed,[59] leading to the following unconditional unemployment contribution:

$$
\begin{aligned}
f_u(t_i, i\epsilon U|J) &= f_u(t_i|i\epsilon U)P(i\epsilon U) \\
&= h_J \exp(-h_J t_i)\frac{\eta_J}{\eta_J + h_J}, \quad t_i > 0
\end{aligned}
\tag{43}
$$

To consider the employed contributions, start with the unconditional cumulative distribution function of wages:

$$
\begin{aligned}
F_e(w_i|J) &= P(W \leq w_i|J) = P(\alpha\Theta + (1-\alpha)\rho U_J \leq w_i|J) \\
&= P(\Theta \leq \frac{w_i - (1-\alpha)\rho U_J}{\alpha}|J) = G_J(\frac{w_i - (1-\alpha)\rho U_J}{\alpha})
\end{aligned}
\tag{44}
$$

---

[58] The result is obtained by imposing the constant hazard function on the general characterization for the density of on-going unemployement spells:

$$f_u(t) = \frac{\exp\left[-\int_0^t h(x)\,dx\right]}{\int_0^\infty s f_c(s)ds}$$

[59] The model without prejudice implies the following flows in and out unemployment:

$$\frac{\partial u}{\partial t} = \eta_J(1 - u_t) - \lambda_J \widetilde{G}(\rho U_J)u_t$$

therefore in steady state:

$$\frac{\eta_J}{\eta_J + \lambda_J \widetilde{G}_J(\rho U_J)} = u = P(i\epsilon U)$$

The corresponding density function is:

$$f_e(w_i|J) = \frac{dF_e(w_i|J)}{dw_i} = \frac{1}{\alpha}g_J(\frac{w_i - (1-\alpha)\rho U_J}{\alpha}) \tag{45}$$

therefore the conditional distribution of observed wages will be:

$$f_e(w_i|w_i > \rho U_J, i\epsilon E, J) = \frac{\frac{1}{\alpha}g_J(\frac{w-(1-\alpha)\rho U_J}{\alpha})}{\widetilde{G}_J(\rho U_J)}, \quad w_i > \rho U_J \tag{46}$$

Finally, the unconditional distribution of observed wages is obtained as:

$$f_e(w_i, w_i > \rho U_J, i\epsilon E, J) = \tag{47}$$
$$= f_e(w_i|w_i > \rho U_J, i\epsilon E, J)P(w_i > \rho U_J|i\epsilon E, J)P(i\epsilon E|J) =$$
$$\frac{\frac{1}{\alpha}g_J(\frac{w-(1-\alpha)\rho U_J}{\alpha})}{\widetilde{G}_J(\rho U_J)}\frac{h_J}{\eta_J + h_J}$$

9.3.2 Model with prejudice

In addition to the worker's type J, now the subscript I denotes the employer's type: $I = N, P$. The derivation of the unconditional unemployment contribution follows the previous steps, once recognized that the constant hazard property holds and that the hazard rates are:

$$h_J = \lambda_J[(1-p)\widetilde{G}(\rho U_J) + p\widetilde{G}(\rho U_J + d\mathbf{I}_{\{W,P\}})] \tag{48}$$

The derivation of the unconditional employment contribution starts considering the cumulative distribution function conditional to agents' types:

$$\begin{aligned} F_e(w_i|J, I) &= P(W \le w_i|J, I) \tag{49}\\ &= P(\alpha(X - d\mathbf{I}_{\{W,P\}}) + (1-\alpha)\rho U_J \le w_i|J, I)\\ &= P(X \le \frac{w_i + \alpha d\mathbf{I}_{\{W,P\}} - (1-\alpha)\rho U_J}{\alpha}|J, I)\\ &= G_J(\frac{w_i + \alpha d\mathbf{I}_{\{W,P\}} - (1-\alpha)\rho U_J}{\alpha}) \end{aligned}$$

The corresponding density will be:

$$f_e(w_i|J, I) = \frac{dF_e(w_i|J, I)}{dw_i} = \frac{1}{\alpha}g_J(\frac{w_i + \alpha d\mathbf{I}_{\{W,P\}} - (1-\alpha)\rho U_J}{\alpha}) \tag{50}$$

To move to the conditional density, just recall that the reservation wage for each workers' type is the same at each employers' type. Therefore we get:

$$f_e(w_i|w_i > \rho U_J, i\epsilon E, J, I) = \frac{\frac{1}{\alpha}g_J(\frac{w+\alpha d\mathbf{I}_{\{W,P\}}-(1-\alpha)\rho U_J}{\alpha})}{\widetilde{G}_J(\rho U_J + d\mathbf{I}_{\{W,P\}})}, w_i > \rho U_J \tag{51}$$

and removing conditioning on employers' type:

$$f_e(w_i|w_i > \rho U_J, i\epsilon E, J) = \tag{52}$$

$$= f_e(w_i|w_i > \rho U_J, i\epsilon E, J, N)P(N) + f_e(w_i|w_i > \rho U_J, i\epsilon E, J, P)P(P) =$$

$$= \frac{\frac{(1-p)}{\alpha}g_J\left(\frac{w_i-(1-\alpha)\rho U_J}{\alpha}\right)}{\widetilde{G}_J(\rho U_J)} + \frac{\frac{p}{\alpha}g_J\left(\frac{w_i+\alpha d\mathbf{I}_{\{W\}}-(1-\alpha)\rho U_J}{\alpha}\right)}{\widetilde{G}_J(\rho U_J + d\mathbf{I}_{\{W\}})}, w_i > \rho U_J$$

Finally, conditioning only on workers' type and using ergodic results on flows in and out employment, we get:

$$f_e(w_i, w_i > \rho U_J, i\epsilon E|J) = \tag{53}$$

$$= f_e(w_i|w_i > \rho U_J, i\epsilon E, J)P(w_i > \rho U_J|i\epsilon E, J)P(i\epsilon E|J) =$$

$$= \left[\frac{\frac{(1-p)}{\alpha}g_J\left(\frac{w_i-(1-\alpha)\rho U_J}{\alpha}\right)}{\widetilde{G}_J(\rho U_J)} + \frac{\frac{p}{\alpha}g_J\left(\frac{w_i+\alpha d\mathbf{I}_{\{W\}}-(1-\alpha)\rho U_J}{\alpha}\right)}{\widetilde{G}_J(\rho U_J + d\mathbf{I}_{\{W\}})}\right]\frac{h_J}{h_J + \eta_J}, w_i > \rho U_J$$

9.4   Identification

Write the log likelihood of the model with prejudice as:

$$\ln L\left(\Omega; w, t\right) = \tag{54}$$

$$N_M \ln \frac{h_M}{h_M + \eta_M} - N_{UM}\ln \eta_M - h_M \sum_{i\in U_M} t_i + \sum_{i\in E_M} \ln \frac{\frac{1}{\alpha}g\left(\frac{w_i-(1-\alpha)\rho U_M}{\alpha}\right)}{\widetilde{G}\left(\rho U_M\right)} +$$

$$+N_W \ln \frac{h_W}{h_W + \eta_W} - N_{UW}\ln \eta_W - h_W \sum_{i\in U_W} t_i + \sum_{i\in E_W} \ln \left[\begin{array}{c}\frac{\frac{(1-p)}{\alpha}g\left(\frac{w_i-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}(\rho U_W)} \\ +\frac{\frac{p}{\alpha}g\left(\frac{w_i+\alpha d-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}(\rho U_W+d)}\end{array}\right]$$

where w_i and t_i are observations on accepted wages and unemployment durations and Ω is the vector of parameters to be estimated. Before defining the set Ω, it is useful to notice the following.

First, we will just discuss here the hardest case to identify: complete heterogeneity between man and women. Complete heterogeneity means that the only parameter men and women have in common is the discount rate ρ. The additional assumption that this discount rate is also shared by employers lead to symmetric Nash-bargaining[60], that is to $\alpha = 0.5$. If identification is proved under complete heterogeneity, then it is proved for specifications where men and women have other parameters in common.

Second, the structural parameters ρ and b_J enter the log likelihood (54) only through the reservation matching value ρU_J. Following Flinn and Heckman 1982, it is therefore possible to estimate ρU_J as a free parameter in the likelihood

---

[60] This assumption basically identifies the Nash bargaining coefficient α. This parameter cannot be identified without employer's side information, as shown in Flinn 2005 and Eckstein and Wolpin 1995.

and then recover ρ and b_J using the reservation wage equation (9). This also shows that ρ and b_J are only jointly identified. Since ρU_J is equal to the reservation wage for each type of worker, the following is a strongly consistent estimator:

$$\widehat{\rho U_J} = \min_{w_i} (w_i, i\epsilon E_J)$$

Third, even if the primitive parameter is the exogenous arrival rate λ_J, we can reparametrize the model considering the hazard rate h_J as the parameter to be estimated since:

$$h_J = \lambda_J[(1-p)\widetilde{G}_J(\rho U_J) + p\widetilde{G}_J(\rho U_J + d\mathbf{I}_{\{W\}})]$$

implies that knowledge of the hazard rate and the probability to accept the match determine a unique value for the arrival rate.

Fourth, we make the following parametric assumption on the productivity distribution:

$$g_J(x;\mu_J,\sigma_J) = \frac{1}{\sigma_J x}\phi[\frac{\ln(x) - \mu_J}{\sigma_J}],\ x > 0$$

where μ_J and σ_J are respectively the location and scale parameter of a lognormal distribution and ϕ is the standard normal density. It is easy to show that information on x is enough to separately identify them.

In the end, we obtain the following set of parameters to be identified:

$$\Omega = \left\{ \begin{array}{c} h_M,\eta_M,\mu_M,\sigma_M \\ h_W,\eta_W,\mu_W,\sigma_W,p,d \end{array} \right\}$$

First consider the hazard rates and termination rates. By first order condition on the log likelihood, we obtain the following Maximum Likelihood Estimators:

$$\widehat{h}_J = \frac{N_{UJ}}{\sum_{i\in U_J} t_i}$$

$$\widehat{\eta}_J = \frac{N_{UJ}}{N_{EJ}}\widehat{h}_J$$

Note that no information from wages is used and therefore their identification is secured independently from the other six parameters (but the other parameters are necessary to recover the primitive λ_J.) The opposite is also true.

Second, consider the parameters of the male productivity distribution (μ_M,σ_M). They play a role only on the term:

$$\sum_{i\in E_M} \ln \frac{\frac{1}{\alpha}g_M\left(\frac{w_i-(1-\alpha)\rho U_M}{\alpha}\right)}{\widetilde{G}_M(\rho U_M)} \tag{55}$$

which only involves male wages and $(\alpha,\rho U_M)$. Wages are truncated at the reservation wage $w_J^* = \rho U_J$ and are the following function of the lognormal random variable x :

$$w = \alpha x + (1-\alpha)\rho U_M$$

44

Therefore we can rewrite (55) as the sum over the following truncated lognormal:

$$\frac{\frac{1}{\alpha}g\left(\frac{w_i-(1-\alpha)\rho U_M}{\alpha}\right)}{\widetilde{G}\left(\rho U_M\right)} = \frac{\frac{1}{s_M w_i}\phi\left[\frac{\ln(w_i)-l_M}{s_M}\right]}{\widetilde{\Phi}\left[\frac{\ln(\rho U_M)-l_M}{s_M}\right]} \tag{56}$$

$$l_M = \alpha\mu_M + (1-\alpha)\rho U_M$$

$$s_M = \alpha\sigma_M$$

from which we know the location and scale parameters l_M and s_M are identified.But since $(\alpha, \rho U_M)$ are known at this stage, this also shows that μ_M and σ_M are identified.

Finally, consider the last term of (54):

$$\sum_{i\in E_W}\ln\left[\frac{\frac{(1-p)}{\alpha}g\left(\frac{w_i-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}\left(\rho U_W\right)} + \frac{\frac{p}{\alpha}g\left(\frac{w_i+\alpha d-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}\left(\rho U_W+d\right)}\right] \tag{57}$$

It is the term that contains the last four parameters to be identified: (μ_W, σ_W, p, d). Each term of (57) can be rewritten as a sum over two truncated lognormal distributions:

$$\frac{\frac{(1-p)}{\alpha}g\left(\frac{w_i-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}\left(\rho U_W\right)} + \frac{\frac{p}{\alpha}g\left(\frac{w_i+\alpha d-(1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}\left(\rho U_W+d\right)} \tag{58}$$

$$= (1-p)\frac{\frac{1}{s_{WN}w}\phi\left[\frac{\ln(w_i)-l_{WN}}{s_{WN}}\right]}{\widetilde{\Phi}\left[\frac{\ln(\rho U_W)-l_{WN}}{s_{WN}}\right]} + p\frac{\frac{1}{s_{WP}w}\phi\left[\frac{\ln(w_i)-l_{WP}}{s_{WP}}\right]}{\widetilde{\Phi}\left[\frac{\ln(\rho U_W)-l_{WP}}{s_{WP}}\right]}$$

where:

$$l_{WN} = \alpha\mu_W + (1-\alpha)\rho U_W$$

$$l_{WP} = \alpha\mu_W + (1-\alpha)\rho U_W - \alpha d$$

$$s_{WN} = s_{WP} = \alpha\sigma_W \equiv s_W$$

This model constitutes a mixture of two truncated lognormal distributions that share the same scale parameter. Therefore p, the proportion in the mixture, and l_{NW}, l_{PW}, s_W, the location and scale parameters, are identified (Teicher 1963.) From l_{WN} we can then recover μ_W since $(\alpha, \rho U_W)$ are known at this stage. From l_{WP} we can secure d since μ_W is fixed by l_{WN}. Finally, s_W recovers σ_W, completing the identification. Note that $s_{WN} = s_{WP}$ is an overidentifying restriction implied by the model. This identification strategy is more general than the lognormal case shown here. It holds for any distribution with a location and scale parameter and that can be recovered by observing the finite mixture.

Even if identification is proved in theory, it is hard to precisely estimate p if the two densities are not well separated (Hill 1963). In this case the separation, as measured by $\frac{l_{WN}-l_{WN}}{s_W} = \frac{\alpha d}{s_W}$, is entirely due to d. Therefore, if d is too close

to zero, p cannot be recovered. On the other side if $p = 0$, then d is trivially not identified. But also if $p = 1$, then d is not identified because knowledge of l_{WP} only determines $(\mu_W - d)$. In sum, it is hard to estimate p and to identify d if they are too close to the boundaries of the parameter space.

The dependence of the identification on functional form assumptions is shown by considering a distribution without a location parameter. Suppose we assume the productivity distribution equal to a negative exponential, then:

$$g_J(x) = \frac{1}{\sigma_J} \exp[-\frac{x}{\sigma_J}],\ x > 0$$

where σ_J is the scale parameter. Identification of σ_M is easy to show. The interesting case is identification of (σ_W, p, d). Rewrite (58) under the new assumption:

$$\frac{\frac{(1-p)}{\alpha} g\left(\frac{w_i - (1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}(\rho U_W)} + \frac{\frac{p}{\alpha} g\left(\frac{w_i + \alpha d - (1-\alpha)\rho U_W}{\alpha}\right)}{\widetilde{G}(\rho U_W + d)}$$

$$= (1-p)\frac{\frac{1}{\sigma_M}\exp[-\frac{w_i - (1-\alpha)\rho U_W}{\sigma_M}]}{\exp[-\frac{\rho U_W - (1-\alpha)\rho U_W}{\sigma_M}]} + p\frac{\frac{1}{\sigma_M}\exp[-\frac{w_i + \alpha d - (1-\alpha)\rho U_W}{\sigma_M}]}{\exp[-\frac{\rho U_W + \alpha d - (1-\alpha)\rho U_W}{\sigma_M}]}$$

$$= \frac{1}{\sigma_M}\exp\left[\frac{\rho U_W - w_i}{\sigma_M}\right]$$

The expression is now independent from both p and d and therefore shows they are not identified under this parametric assumption.

References

[1] Albrecht, J., A. Bjorklund and S. Vroman (2003), "Is There a Glass Ceiling in Sweden?", *Journal of Labor Economics,* 21(1), 145-177.

[2] Altonji, J.G. and R.M. Blank (1999) "Race and Gender in the Labor Market", in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3, Amsterdam: Elsevier Science.

[3] Ausubel, L., P. Cramton and R. Deneckere (2002), "Bargaining with Incomplete Information", in: R. Aumann and S. Hart (eds.) *Handbook of Game Theory*, Vol. 3, Amsterdam: Elsevier Science.

[4] Becker, G. (1971), *The economics of discrimination*, Chicago: The University of Chicago Press (first ed. 1957).

[5] Binmore, K. (1987), "Perfect Equilibria in Bargaining Models" in K. Binmore and P. Dasgupta (eds.) *The Economics of Bargaining*, Basil Blackwell, Oxford.

[6] Binmore, K.; A. Rubinstein and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modelling", *The RAND Journal Of Economics*, 17(2), 176-188.

[7] Black, D. (1995), "Discrimination in an Equilibrium Search Model", *Journal of Labor Economics,* 13(2), 309-334.

[8] Blau, F. (1998), "Trends in the Well-Being of American Women, 1970-1995", *Journal of Economic Literature,* 36(1): 112-165.

[9] Blau. F. and L. Kahn (2003), "Understanding International Differences in the Gender Pay Gap", *Journal of Labor Economics*, 21(1): 106-144.

[10] Blau. F. and L. Kahn (2000), "Gender Differences in Pay", *Journal of Economic Perspectives*, 14(4), 75-99.

[11] Bollinger, C. (1998), "Measurement Error on the Current Population Survey: A Nonparametric Look", *Journal of Labor Economics*, 16(3): 576-594.

[12] Borjas, G.J. and S.G. Bronars (1989), "Consumer Discrimination and Self-Employment", *Journal of Political Economy*, 97(3), 581-605.

[13] Bound, J., C. Brown and N. Mathiowetz (2001), "Measurement Errors in Survey Data", in: J.J. Heckman and E.E. Leamer (eds.), *Handbook of Econometrics*, Volume 5, Amsterdam: Elsevier Science.

[14] Bowlus, A. (1997), "A Search Interpretation of Male-Female Wage Differentials", *Journal of Labor Economics*, 15(4), pg. 625-657.

[15] Bowlus, A. and Z. Eckstein (2002), "Discrimination and Skill Differences in an Equilibrium Search Model", *International Economic Review*, 43(4), pg. 1309-1345.

[16] Bowlus, A., N. Kiefer and G. Neumann (2001), "Equilibrium Search Models and the Transition from School to Work", *International Economic Review*, 42(2), 317-343.

[17] Bowlus, A., N. Kiefer and G. Neumann (1995), "Estimation of Equilibrium Wage Distribution with Heterogeneity", *Journal of Applied Econometrics*, 10: S119-131.

[18] Burdett, K. and D. Mortensen (1998), "Wage Differentials, Employer Size, and Unemployment", *International Economic Review*, 39(2), 257-273.

[19] Cain, G.C. (1986), "The Economic Analysis of Labor Market Discrimination", in: Ashenfelter O. and Layard R. (ed.) *Handbook of Labor Economics*, Amsterdam: North Holland.

[20] Coate, S. and G.C. Loury (1993), "Will Affirmative-Action Policies Eliminate Negative Stereotypes?", *American Economic Review*, 83(5), 1220-1240.

[21] Dey, M. and C. Flinn (2005), "An Equilibrium Model of Health Insurance Provision and Wage Determination", *Econometrica*, 73: 571-627.

[22] Donohue, J. and J. Heckman (1991), "Continuos Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks", *Journal of Economic Literature, 29(4), 1603-1643.*

[23] Eckstein, Z. and É. Nagypál (2004), "US Earnings and Employment Dynamics 1961-2002: Facts and Interpretation", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 28, No. 2, pp. 10-29.

[24] Eckstein, Z. and K. Wolpin (1995), "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model", *The Review of Economic Studies*, 62(2): 263-286.

[25] Eckstein, Z. and K. Wolpin (1999), "Estimating the Effect of Racial Discrimination on First Job Wage Offers", *The Review of Economics and Statistics*, 81(3): 384-392.

[26] Edgeworth, F.Y. (1922), "Equal Pay to Men and Women for Equal Work", *The Economic Journal*, 32(128), 431-457.

[27] Flinn, C. (2002a), "Labor Market Structure and Inequality: A Comparison of Italy and the U.S.", *Review of Economic Studies*, 69(3), 611-645.

[28] Flinn, C. (2002b), "Interpreting Minimum Wage Effects on Wage Distributions: A Cautionary Tale", *Annales d'Economie et de Statistique*, 67/68, 309-355.

[29] Flinn, C. (2005), "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining and Endogenous Contact Rates", mimeo, New York University.

[30] Flinn, C. and J. Heckman (1982), "New Methods in Analyzing Structural Models of Labor Market Dynamics", *Journal of Econometrics*, 18: 115-168.

[31] Goldin, C. and C. Rouse (2000), "Orchestrating Impartiality: the Impact of "Blind" Auditions on Female Musicians", *American Economic Review*, 90(4), 715-741.

[32] Heckman, J. J. (1998), "Detecting Discrimination", *Journal of Economic Perspective* 12(2): 101-116.

[33] Hill, B.M. (1963), "Information for Estimating the Proportions in Mixtures of Exponential and Normal Distributions", *Journal of the American Statistical Association*, 58(304), 918-932.

[34] Holzer, H. and D. Neumark (2000), "Assessing Affirmative Action", *Journal of Economic Literature*, 37, 483-568.

[35] Jovanovic B. (1979), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, 87(5), 973-90.

[36] Kiefer, N.M. and G.R. Neumann (1993), "Wage Dispersion with Homogeneity: the Empirical Equilibrium Search Model", in: H. Bunzel, P. Jensen and N. Westergard-Nielsen (eds.) *Panel Data and Labor Market Dynamics*, Amsterdam: Elsevier Science.

[37] Mailath, G., L. Samuelson and A. Shaked (2000), "Endogenous inequality in Integrated Labor Markets with Two-Sided Search", *American Economic Review*, V. 90, N. 1, Pp. 46-72.

[38] Moro, A. and P. Norman (2003), "Affirmative action in a competitive economy", *Journal of Public Economics*, 87, 567-594.

[39] Neumark, D. (1996), "Sex Discrimination in Restaurant Hiring: an Audit Study", *Quarterly Journal of Economics*, 111(3), 915-941.

[40] Postel-Vinay, F. and J-M Robin (2002). "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70(6), 2295-2350.

[41] Rosen, A. (2003), "Search, Bargaining and Employer Discrimination", *Journal of Labor Economics,* 21(4), 807-829.

[42] Sasaki, M. (1999), "An Equilibrium Search Model with Coworker Discrimination", *Journal of Labor Economics,* 17(2), 377-407.

[43] Teicher, H. (1963), "Identifiability of Finite Mixtures", *The Annals of Mathematical Statistics*, 34(4), 1265-1269.

[44] Van den Berg, G. and G. Ridder (1998), "An Empirical Equilibrium Search Model of the Labor Market", *Econometrica*, 66, 1183-1221.

[45] Welch, F. (1976), "Employment Quotas for Minorities", *Journal of Political Economy*, 84(4), S105-S139.

[46] Wolpin, K.I. (1987), "Estimating a Structural Search Model: the Transition from School to Work", *Econometrica*, 55(4), 801-817.

Table 1: Descriptive statistics.

Sample Moments	N	$P(i\epsilon U)$	$E(w_i|i\epsilon E)$	$SD(w_i|i\epsilon E)$	$E(t_i|i\epsilon U)$	$SD(t_i|i\epsilon U)$
			Without Trimming			
All	2,324	0.0211	19.16	9.51	4.53	4.13
Women	1,059	0.0264	16.78	8.46	3.72	3.33
Men	1,265	0.0166	21.13	9.87	5.59	4.88
Women/Men Ratio	0.837	1.590	0.794	0.857	0.665	0.682
			With Trimming			
All	2,213	0.0226	19.89	9.16	4.53	4.13
Women	1,009	0.0277	17.41	8.17	3.72	3.33
Men	1,204	0.0174	21.94	9.44	5.59	4.88
Women/Men Ratio	0.838	1.592	0.794	0.865	0.665	0.682

Note: Data extracted from the *Annual Social and Economic Supplement* (March Supplement) of the CPS for the year 1995. Variables definition: $t_i=$ monthly unemployment duration; $w_i=$hourly earnings in dollars; $i\epsilon U$ individual i is unemployed; $i\epsilon E$ individual i is employed.

Table 2: Maximum Likelihood Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
λ	0.2220	0.2247	0.2280			
	(0.0317)	(0.0322)	(0.0327)			
λ_M				0.1795	0.1792	0.1795
				(0.0392)	(0.0391)	(0.0392)
λ_W				0.2700	0.2777	0.2862
				(0.0510)	(0.0528)	(0.0545)
η	0.0050	0.0050	0.0050			
	(0.0010)	(0.0010)	(0.0010)			
η_M				0.0032	0.0032	0.0032
				(0.0010)	(0.0010)	(0.0010)
η_W				0.0077	0.0077	0.0077
				(0.0021)	(0.0021)	(0.0021)
μ		3.4338			3.4338	
		(0.0154)			(0.0154)	
σ		0.5231			0.5231	
		(0.0096)			(0.0096)	
μ_M	3.4563		3.4563	3.4563		3.4563
	(0.0165)		(0.0165)	(0.0165)		(0.0165)
σ_M	0.5578		0.5579	0.5578		0.5578
	(0.0123)		(0.0123)	(0.0123)		(0.0123)
μ_W	3.2119		3.4575	3.2119		3.4546
	(0.0187)		(0.0589)	(0.0187)		(0.0456)
σ_W	0.5719		0.4222	0.5719		0.4232
	(0.0140)		(0.0264)	(0.0140)		(0.0235)
d		6.6777	13.4555		6.6786	13.5940
		(3.1068)	(2.6082)		(3.1054)	(2.3602)
p		0.8113	0.5160		0.8113	0.5065
		(0.3640)	(0.1483)		(0.3633)	(0.0988)
N	2,213	2,213	2,213	2,213	2,213	2,213
lnL	-7898.945	-7901.449	-7880.325	-7896.602	-7899.115	-7877.978
$\chi^2_{(df)}$	41.93	46.94	4.69	37.25	42.27	

Notes: Asymptotic standard errors in parentheses. Data from CPS 1995. Sample: College graduated or more; 30-55 years old; white. The reservation values are estimated by the minimum observed earning in the distribution of each group and they are, with bootstrap standard errors in parentheses: $w^*_W = 5.750$ (0.0911) and $w^*_M = 7.175$ (0.0438). $\chi^2_{(df)}$ is the value of the statistic for a likelihood ratio specification test against specification (6).

	(1)	(2)	(3)	(4)	(5)	(6)	Sample
			Productivity				
$E(x)$		35.54			35.54		
		(0.538)			(0.537)		
$V(x)$		397.39			397.38		
		(21.591)			(21.582)		
$E_M(x)$	37.04		37.04	37.04		37.04	
	(0.645)		(0.645)	(0.645)		(0.645)	
$V_M(x)$	500.76		500.79	500.77		500.77	
	(35.430)		(35.434)	(35.432)		(35.432)	
$E_W(x)$	29.24		34.70	29.24		34.61	
	(0.575)		(1.733)	(0.575)		(1.320)	
$V_W(x)$	330.75		234.95	330.74		234.92	
	(26.306)		(20.827)	(26.304)		(20.799)	
			Earnings				
$E_M(w|E)$	22.17	21.39	22.17	22.17	21.39	22.17	21.94
	(0.322)	(0.267)	(0.323)	(0.323)	(0.266)	(0.323)	
$E_W(w|E)$	17.56	18.37	17.40	17.56	18.37	17.40	17.41
	(0.288)	(0.280)	(0.295)	(0.288)	(0.280)	(0.273)	
			Unemployment				
u_M	0.022	0.022	0.022	0.017	0.017	0.017	0.017
	(0.0031)	(0.0031)	(0.0031)	(0.0038)	(0.0038)	(0.0038)	
$E_M(t|U)$	4.523	4.462	4.403	5.593	5.593	5.593	5.593
	(0.6461)	(0.6389)	(0.6313)	(1.2205)	(1.2205)	(1.2204)	
u_W	0.022	0.023	0.023	0.028	0.028	0.028	0.028
	(0.0031)	(0.0032)	(0.0032)	(0.0052)	(0.0052)	(0.0052)	
$E_W(t|U)$	4.529	4.602	4.668	3.724	3.723	3.725	3.725
	(0.6470)	(0.6594)	(0.6706)	(0.7036)	(0.7035)	(0.7040)	

Notes: Predicted values from specifications (1)-(6) reported in Table 2. Asymptotic standard errors by Delta method in parentheses.

Table 4: Earnings Differential Decomposition

Women/Men Ratio generated by:	Entire Distribution	Bottom 25%	Top 25%	Reservation Values
Productivity (μ, σ)	.917	1.058	.884	.842
Prejudice (d, p)	.819	.640	.905	.573
Behavior (λ, η)	1.168	1.351	1.084	1.832
All $(b, \lambda, \eta, \mu, \sigma, d, p)$	.785	.765	.813	.801
Sample	.794	.784	.800	.801

Notes: Women/Men Ratio on average accepted earnings computed over the entire distribution or over the bottom and top 25% quantiles. The last column report the ratio of the two reservation wages. Results based on specification (6), Table 2.

Table 5: Agents' Average Welfare and Policies

	Benchmark Model	Same Productivity	Equal Pay	Affirmative Action
Welfare Measures:				
Workers:				
Men	100	100	92.46	98.53
Women	75.85	89.11	84.36	77.65
Overall	88.99	95.04	88.77	89.01
Employers:				
Unprejudiced	100	103.36	84.03	99.84
Prejudiced	79.01	82.77	66.59	78.88
Overall	89.37	92.93	75.19	89.22
Workers' Labor Market Variables:				
w_M^*	7.175	7.175	6.017	7.418
w_W^*	5.750	7.365	7.366	6.507
$E_M(w|E)$	22.17	22.17	20.53	21.84
$E_W(w|E)$	17.40	20.40	19.23	17.79
$E_M(t|U)$	5.593	5.593	5.676	5.597
$E_W(t|U)$	3.725	3.964	3.551	3.714
u_M	0.017	0.017	0.018	0.018
u_W	0.028	0.030	0.027	0.028

Notes: The *Benchmark Model* is specification (6), Table 2. *Same productivity* means women at men productivity. *Equal Pay* means each employer must pay one wage at same productivity. *Affirmative Action* means employer receive a flow subsidy of 1\$/h when hiring a woman and the subsidy is financed by a lump-sum tax on workers; the earnings reported in the lower part of the Table are net of this tax. The top panel reports average welfare normalized with respect to men (workers) and with respect to unprejudiced employers (employers) in the Benchmark model.

Table A.1: Estimation sample selection.

CPS raw data sample	149,642
Only relevant observations:	
In the 4th and 8th in-month-sample	-111,784
In the Labor Force	-19,238
Eligible for *Earner Study*	-3,243
Homogeneity criteria:	
Employed or looking for a job	-192
Mature working career (30-55 years old)	-5,932
White	-1,372
College or more	-5,481
Data cleaning:	
Top-coded durations	-3
Impossible to obtain hourly earnings	-67
Above top-coded earnings	-6
Final sample	**2,324**
Women	1,059
Men	1,265

Note: Data from CPS - March, 1995.
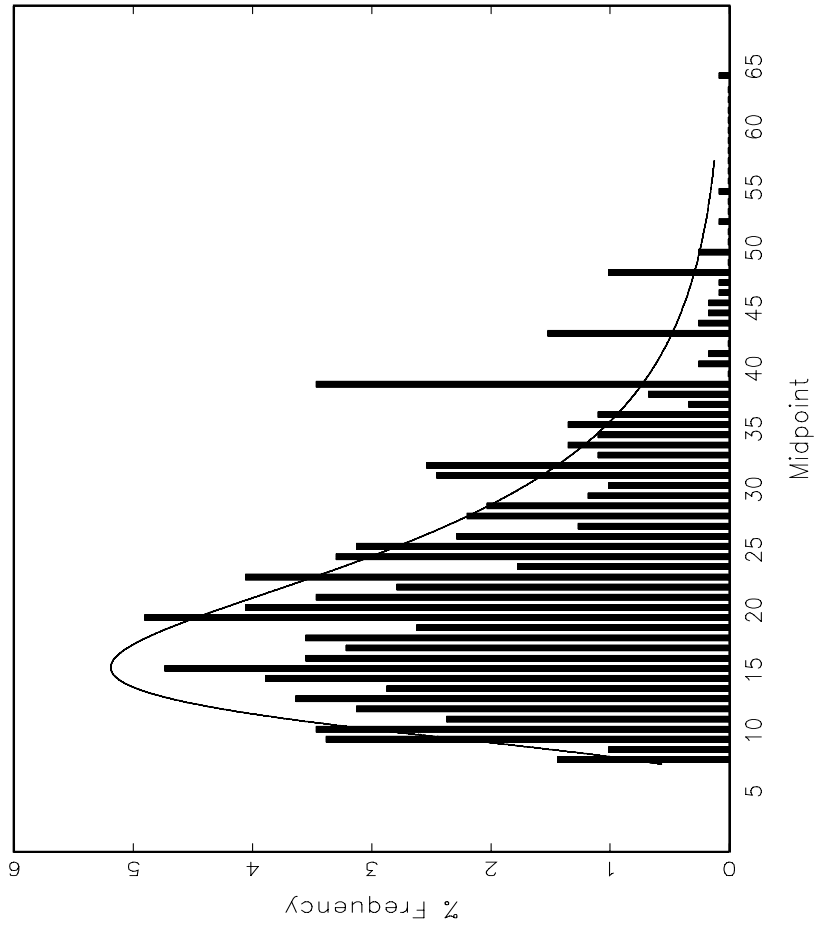
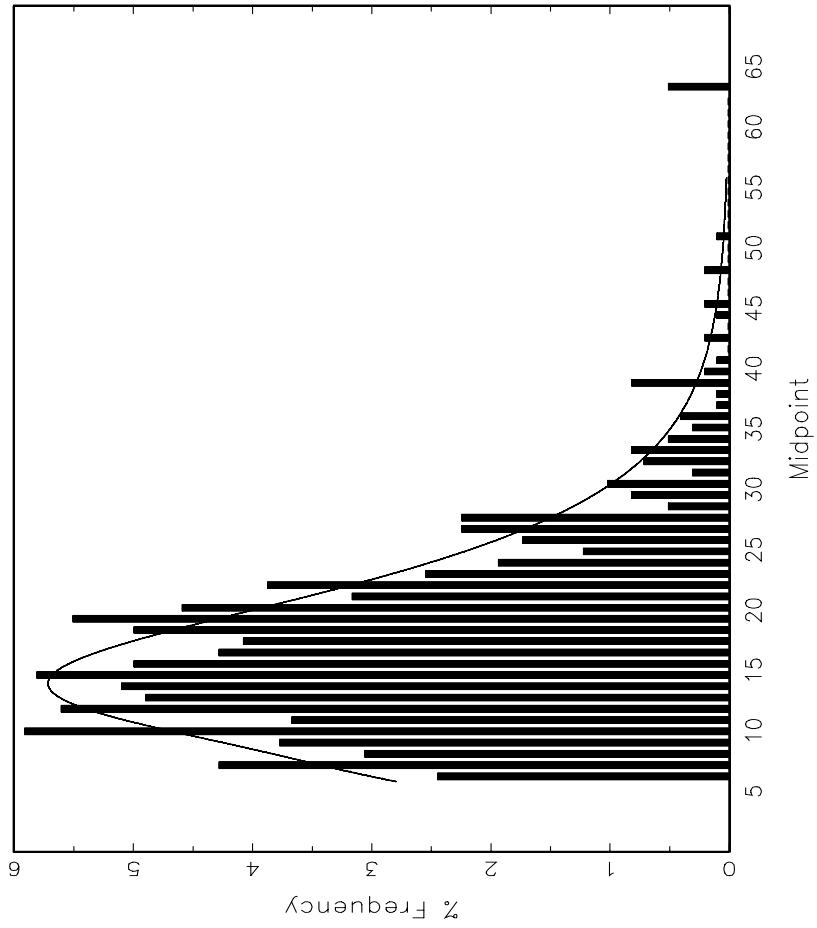Figure 1: Empirical and Predicted Earnings Distribution - Men

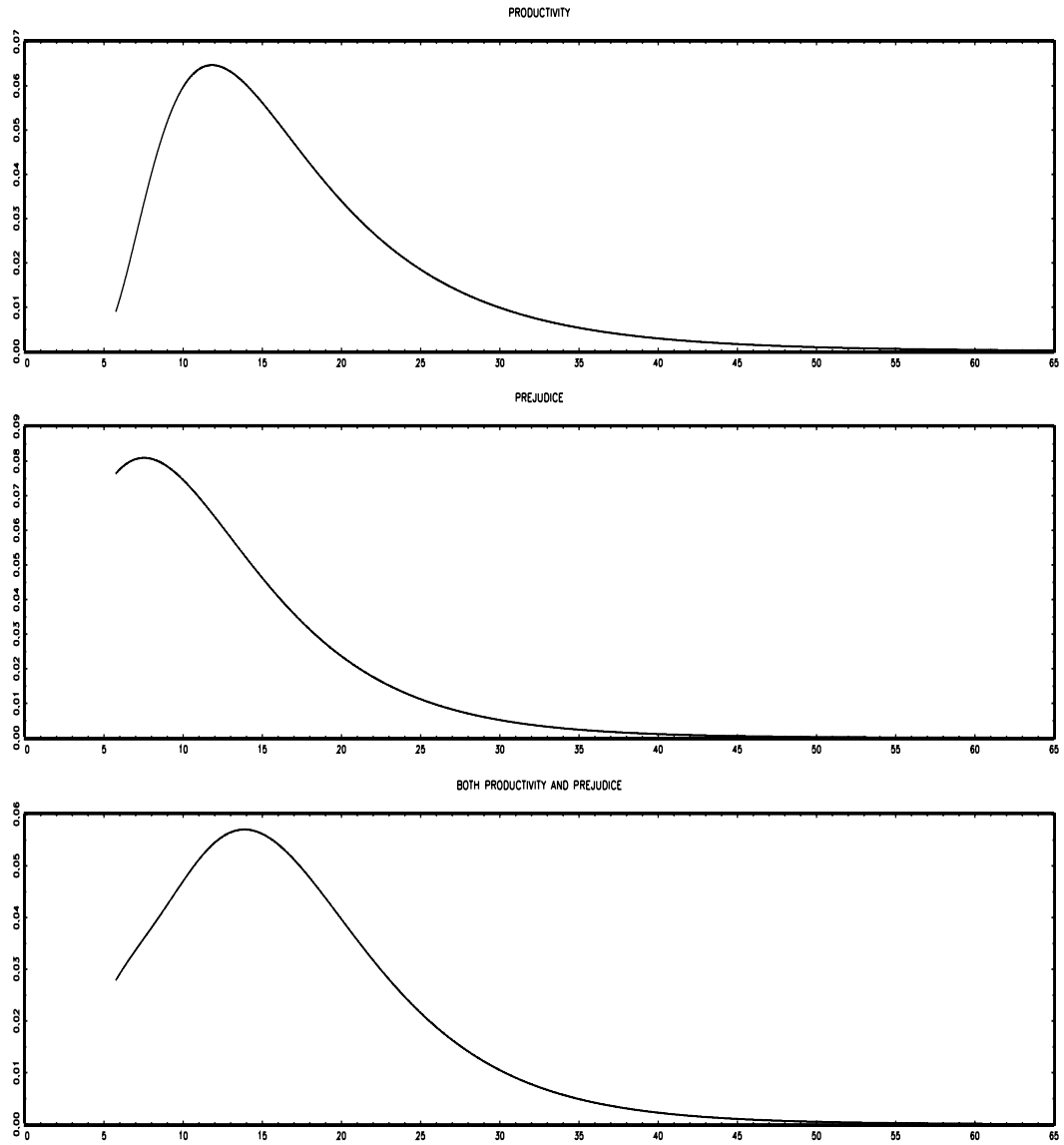Figure 2: Empirical and Predicted Earnings Distribution - Women

Figure 3: Differential Impact of Productivity and Prejudice on the Accepted
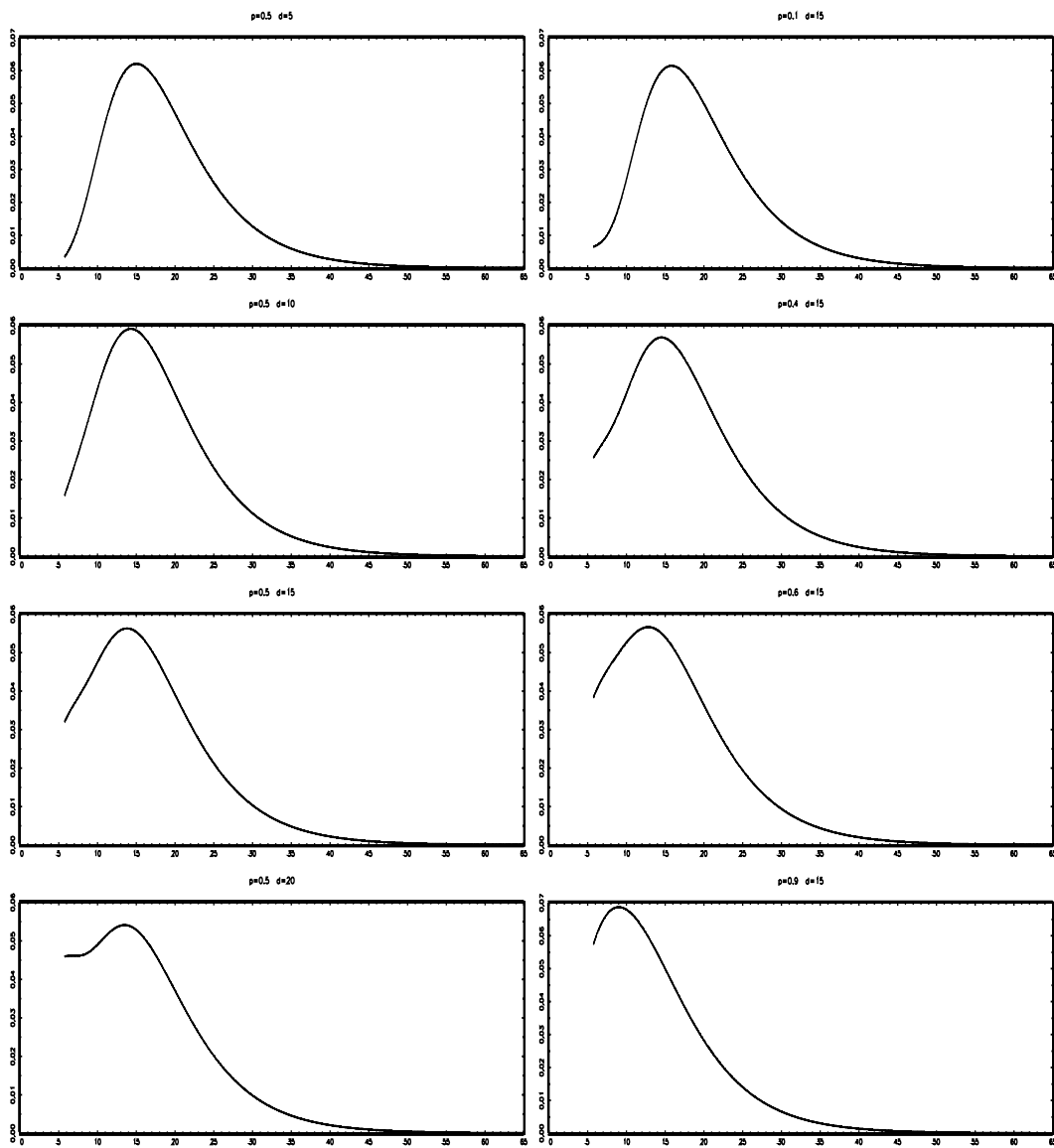Earnings Distribution

Figure 4: Impact of Change in Disutility (Left) and Change in Proportion of Prejudiced Employers (Right) on the Accepted Earnings Distribution

60