Collegio Carlo Alberto

Prejudice and Gender Differentials in the U.S. Labor Market in the Last Twenty Years

Luca Flabbi

Working Paper No. 57
September 2007
www.carloalberto.org

Prejudice and Gender Differentials in the U.S. Labor Market in the Last Twenty Years

Luca Flabbi*
Georgetown University

September 14, 2007

*Georgetown University, Department of Economics, 37th & O Streets, NW, Washington, DC 20057, E-mail: lf74@georgetown.edu, URL: www.georgetown.edu/faculty/lf74/. Other affiliations: Visiting Fellow Collegio Carlo Alberto, Torino; IZA, Bonn; CHILD-Collegio Carlo Alberto, Torino. I would like to thank Marjorie McElroy and the referees and editors of the *Journal of Econometrics* volume on the estimation of Structural Models for useful and detailed comments; Chris Flinn for crucial and encouraging comments on the research project within which this paper has been developed; Anna Fruttero, Bob Miller, and Ken Wolpin for very useful insights; and the participants and organizers of the Duke Conference on Structural Models for the productive and friendly environment. I also greatly enjoyed the hospitality of Collegio Carlo Alberto during a revision of this paper.

Abstract

Earnings differentials between men and women have experienced a stable convergence during the 1980s, following a process started in the late 1970s. However, in the 1990s the convergence has almost stopped. The first objective of the paper is to evaluate if discrimination, defined as explicit prejudice, may have a role in explaining this slowdown in the converge. The second objective is to assess whether the prediction of a decrease in the proportion of prejudiced employers implied by the Becker's model of taste discrimination is taking place and if so at what speed. These objectives are achieved by developing and estimating a search model of the labor market with matching, bargaining, employer's prejudice and worker's participation decisions. The results show that the proportion of prejudiced employers is estimated to be decreasing at an increasing speed, going from about 69% in 1985 to about 32% in 2005. Therefore prejudice does not seem a relevant factor in explaining the slower convergence between male and female earnings in the 1990s. The results are consistent with the Becker's model of taste discrimination if one is willing to assume a very slow adjustment process.

JEL Classification: C51; J7; J64
Keywords: Gender Differentials; Discrimination; Search Models; Maximum Likelihood Estimation; Structural Estimation..

1 Introduction

Gender differentials in the labor market have experienced a stable reduction in the 1980s, following a process started in the late 1970s. Figure 1 shows this trend on hourly earnings, one of the main indicators used to evaluate gender differentials in the labor market. In 1981 women were earning on average 37.9% less than men, compared to 29.2% in 1990 and to about 20.7% in 1993. The convergence between male and female earnings persists after controlling for standard human capital characteristics: the conditional differential decreased from more than 36.9% in 1981 to 27.8% in 1991.

However, during the 1990s this convergence has almost stopped. The evidence is quite robust and is documented in the literature: Blau and Kahn 2004 use PSID data and a careful treatment of labor market experience to find a similar result over the 1979-1998 period; Eckstein and Nagypal 2004 find it using yearly income data from CPS over the 1961-2002 period; Fortin and Lemieux 2000 use CPS data and relate this trend with trends in the wage distribution. Also the simple male/female median earnings ratio, regularly published by the Bureau of Labor Statistics, produces a similar picture.

The first objective of this paper is to evaluate whether discrimination, defined as explicit prejudice, may have a role in explaining this slowdown in the converge between male and female earnings. The question is relevant because the convergence in the 1980s has been interpreted by some authors as evidence that gender discrimination in the labor market was becoming irrelevant (O'Neill 2003). Policy implications are quite different depending on whether the slowdown is due to barriers that do not allow for an efficient use of female labor, such as prejudice, or to other labor market institutions and out of labor market factors.

The second objective of the paper is theoretical and closely related to one of the most common definition of discrimination used in the literature: discrimination defined as explicit prejudice of employers against women due to preference. This is the *taste discrimination* idea developed by Gary Becker (1957, 1971) and is the form of discrimination that will be studied in this paper. Becker's model of taste discrimination has been very influential but also strongly criticized because, as Arrow 1973 puts it, the "model predicts the absence of the phenomenon it was designed to explain." The issue is that only the least discriminatory employers survive in the long run, eventually leading to the disappearance of discrimination. Becker's answer to this criticism has been that there is not an infinite supply of "entrepreneurial ability", so there is not really free-entry of employers with unprejudiced tastes. Moreover, labor markets are not necessarily perfectly competitive and, as for example Heckman 1998 points out, search frictions may give employers monopsony power slowing down this adjustment process. The second objective of this paper is then to provide an estimate of the proportion of prejudiced employers over time to determine whether the prediction of the model, a reduction in the proportion of prejudiced employers, is actually taking place and at what speed.

The paper develops a search model with matching, bargaining and labor market participation decisions in presence of two types of employers (prejudiced and unprejudiced) and two types of workers (men and women). Additional heterogeneity is introduced by match-specific productivity. This structure solves the fundamental identification problem of the empirical literature on discrimination: how to separate the impact of prejudice from the impact of other group-specific characteristics, such as unobserved productivity differences. This identification strategy is detailed and proved in Flabbi 2005, which builds on Flinn and Heckman 1982. The main intuition is to exploit the peculiar impact of prejudice - parametrized by the proportion of prejudiced employers and by their disutility from hiring women - on the shape of the accepted earnings distribution. The model also allows for a variety of equilibrium effects, in particular the bargaining structure generates spillover effects. Spillover effects imply that in equilibrium unprejudiced employers wage discriminate women despite not receiving any disutility from hiring them.

The model is estimated on hourly earnings and monthly unemployment durations extracted from Current Population Survey (CPS) data. Some restrictions on race, education and age are imposed on the estimation sample to guarantee the ex-ante homogeneity assumed by the theoretical model. Various specifications are estimated to account for family characteristics like marital status and presence of young children.

The proportion of prejudiced employers is estimated to be decreasing at an increasing speed: from 69% in 1985 to 57% in 1995 to 32% in 2005. A counterfactual experiment is implemented in order to isolate the impact of prejudiced behavior over time on relevant labor market variables. The main conclusions of the paper are that prejudice does not seem a major factor in the slowdown of the convergence of earnings between men and women and that the estimates are broadly consistent with the predictions of the taste discrimination model if one is willing to allow for a very slow adjustment process over time.

In terms of the empirical search literature that focuses on discrimination, the paper is most closely related to Flabbi 2005, Bowlus and Eckstein 2002 and Bowlus 1997. The first contribution develops the methodology to separately identify productivity differences and discrimination utilized in this paper, provides a first set of estimates on one cross-section (CPS 1995), and implements policy experiments about affirmative action and equal pay policies. Differently from the current paper, it does not focus on over time changes and does not allow for a labor market participation decision. Bowlus and Eckstein 2002 focuses on race discrimination and it is the first contribution which exploits the structure of a search model to propose an identification strategy to jointly estimate productivity differences and prejudiced behavior. Finally, Bowlus 1997 separately estimates a search model on male and female data allowing for a labor market participation decision. Her model, though, does not allow for the presence of discrimination and as a result most of the wage differential is estimated to be due to productivity differences.

The paper is organized as follows. Section 2 describes the model, section 3 specifies the likelihood function and discusses the identification, section 4 presents the data, section 5 comments on the estimation results and the counterfactual experiments, sections 6 concludes.

2 The Model

The model is characterized by four main ingredients: a labor market with search frictions; an employment relation based on employer-employee match-specific productivity; a wage setting obtained by bargaining; and workers and employers' types defined by prejudiced preferences. There are two types of workers - Women (W) and Men (M) - and two types of employers - Prejudiced (P) and Unprejudiced (N). Following Becker's taste discrimination idea, prejudiced employers receive a disutility d from hiring female workers.

The search-matching-bargaining framework is an empirically tractable way to characterize labor market dynamics and has been used by, among others, Eckstein and Wolpin 1995 to study internal rate of return to schooling and Flinn 2005 to study the impact of mandatory minimum wage policies. Types defined by prejudice have been used in a huge number of applications and theoretical contributions. In the context of the estimation of a search model, they have been used by Bowlus and Eckstein 2002 and Flabbi 2005.

2.1 The Environment

The model is in continuous time and agents are infinitely lived. Workers can be in one of three different states: employment, unemployment and nonparticipation in the labor market. The introduction of the nonparticipation state is not standard in the literature: for example, all the works previously mentioned assume only two states, employment and unemployment. However, the participation rate of women has changed much more than the participation rate of men in the last twenty years. Figure 2 reports some statistics about this dynamic: the gender differentials in the participation and employment rates. Conditioning on participation, there are no significant differences in the employment rates between men and women. On the contrary, participation rates were quite different at the beginning of the 1980s (about a 30% differential) and despite converging over the whole period a 17.7% differential persists in 2005. A meaningful comparison across years should therefore take into account, at least in part, this phenomenon.

The following notation is used to indicate the proportion of the population in each state (Rate) and the present discounted value corresponding to each state (Value):

State	Rate	Value
Employment	el	$V[w(x)]$
Unemployment	ul	U
NonParticipation	$(1 - l)$	NP

where $w(x)$ denotes the flow wage of a match that generates a flow productivity x. Each rate is computed with respect to the entire population, therefore employment and unemployment rates conditioning on labor market participation are e and u. A subscript will denote types: for example the value for a woman working at an unprejudiced employer is denoted by $V_W[w_{WN}(x)]$.

Workers receive a flow utility value $z \sim Q(z)$ when non-participating in the labor market.[1] The assumption is that leisure, public goods, or other factors generate positive utility for individuals that voluntarily decide not to supply labor. This implies that individuals with a sufficiently low value of nonparticipation will enter the labor market, the others will stay out. The model is quite stylized in this respect because of the lack of transition to (and duration in) nonparticipation.[2] However, the presence of a nonparticipation state allows for some interesting equilibrium effects.

Once individuals decide to supply labor, they randomly meet employers following a Poisson process characterized by an exogenous arrival rate λ. Once a worker and an employer meet, their types and the productivity value specific to the match are revealed. Productivity is denoted by $x \sim G(x)$. Upon observing types and productivity, the two agents engage in bargaining to determine wages and then they decide whether to accept the match. If they do, the employment relation may be terminated by shocks modelled as a Poisson process with exogenous rate η. If they do not, agents go back to the previous state and search continues.

With a constant and common intertemporal discount rate ρ, the stationarity of the environment generates the following value for an individual deciding not to participate in the labor market:

$$NP_J = \frac{z}{\rho} \tag{1}$$

for $J = W, M$.

Once an agent decides to participate, the value of participating in the labor market is the value of the unemployment state since search while unemployed is a necessary step to meet employers and receive offers. In each instant while unemployed, the worker may receive no offer or an offer from a prejudiced or unprejudiced employer. The proportion of prejudiced employers is denoted by p.[3] The unemployment state generates a flow value b which describes the flow of utility or disutility from the search process. Therefore, by stationarity and by property of the Poisson processes, the value of being unemployed for each

[1] Capital letters function of a random variable denote the cumulative distribution function, i.e. $G(x)$. Moreover, given a cdf $G(x)$ I will define $\widetilde{G}(x) \equiv 1 - G(x)$.

[2] Bowlus 1997 develops and estimates a search model that allows for a full dynamic treatment of the non-participation state to study gender differentials in the labor market.

[3] A more precise definition of p would be as the *measure* of offer that the worker is receiving coming from prejudiced employers, with a single employer contributing in a higher or smaller proportion to this measure. However, employers are highly stylized here and the difference becomes simply semantic. To simplify exposition, I will then refer to p simply as the proportion of prejudiced employers.

type of worker is:

$$\rho U_J = b + \lambda\{p \int \max[V_J [w_{JP} (x)] - U_J, 0]dG(x)+ \\ +(1 - p) \int \max[V_J [w_{JN} (x)] - U_J, 0]dG(x)\} \tag{2}$$

for $J = W, M$. The expression is written in recursive form: the *max* operator indicates that workers accept the match if and only if the value of filling that job is higher than the outside option, i.e. unemployment. Notice that the value of unemployment is constant with respect to both wages and productivity because all the agents, conditioning on gender, are *ex-ante* identical: only when a match is realized the *ex-post* individual heterogeneity will be generated.

A worker of type J filling a job at an employer of type I and receiving a wage w is characterized by:

$$V_J [w_{JI} (x)] = \frac{w_{JI} (x) + \eta U_J}{\rho + \eta} \tag{3}$$

for $J = W, M$ and $I = N, P$.

The last element necessary to fully characterize the decision rule and the equilibrium is the wage setting process. The axiomatic Nash-bargaining solution is assumed to solve the bargaining game. The wage schedule is therefore obtained by maximizing with respect to w_{JI} the product of the surpluses for the worker and the employer (the Nash bargaining product):

$$[V_J (w_{JI}) - U_J]^\alpha \left[x - d\mathbf{I}_{\{W,P\}} \square\right]^{(1-\alpha)} \tag{4}$$

generating:

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

for $J = W, M$ and $I = N, P$, where $\alpha \in [0, 1]$ denotes the workers' weight in the Nash bargaining product and $\mathbf{I}_{\{W,P\}}$ denotes an indicator function equal to one when the worker is female and the employer prejudiced. As a result the wage of the worker is a linear combination of the worker's outside option in the bargaining process (the discounted value of unemployment ρU_J) and the total yield generated in the match (x or $(x - d)$ in the case of a women matching with a prejudiced employer). Notice that the productivity x is reduced by the discrimination intensity parameter d when a prejudiced employer is matching with a woman because this is a utility cost that the employer is bearing in that particular type of match.

2.2 The Equilibrium

The main implications of the previous section are that wages are increasing in productivity (equation (5)), that the value of being employed is increasing in wages (equation (3)) while the value of unemployment is constant with respect to both productivity and wages (equation (2)). Hence the optimal decision rule is characterized by a reservation value property defined over productivity. That is, the optimal decision rule to accept a match has the following form: accept the match if the productivity draw is higher than the reservation value x^*_{JI} and reject otherwise. The reservation value is common to both the worker of type J and the employer of type I: this is an agreement result that derives directly from the assumed axiomatic Nash solution for the bargaining problem.[4] The reservation value is defined as the productivity value that makes agents indifferent between accepting or rejecting the match. Imposing this condition on equations (3) and (5) leads to:

$$\begin{aligned} x^*_{JI} &= \rho U_J + d\mathbf{I}_{\{W,P\}} && (6) \\ w^*_J &= \rho U_J && (7) \end{aligned}$$

for $J = W, M$ and $I = N, P$ and where w^*_J defines the reservation wage. Notice that the reservation productivity value depends on both the worker's and employer's type, while the reservation wage depends only on the worker's type. The intuition is that employers have different preferences over workers' types while workers base their decisions only on wage considerations.

The optimal decision rule to participate in the labor market also has a reservation value property: an individual will enter the labor market when the value of nonparticipation is lower than the value of participation. As pointed out before, the value of participating in the labor market is the value of the unemployment state. The reservation value z^*_J is then obtained by equation (1):

$$z^*_J = \rho U_J \tag{8}$$

for $J = W, M$.

The unique steady state equilibrium will: (i) uniquely define the values of unemployment as a function of the exogenous parameters; (ii) equalize flows in and out of unemployment; and (iii) determine the measure of labor market participants. This is stated in the following definition.

Definition 1 *Given a vector $(\lambda, \eta, \rho, b, \alpha, d, p)$ and the probability distribution functions $G(x)$ and $Q(z)$, an* ***equilibrium*** *is two vectors $(U_J, u_J, l_J)_{J=W,M}$ such that:*

$$\rho U_J = b + \frac{\lambda\alpha}{\rho+\eta}\left\{ \begin{array}{l} (1-p)\int_{\rho U_J}[x-\rho U_J]dG(x)+ \\ +p\int_{\rho U_J + d\mathbf{I}_{\{W,P\}}}[x - d\mathbf{I}_{\{W,P\}} - \rho U_J]dG(x) \end{array} \right\} \tag{9}$$

$$u_J = \eta\left\{\eta + \lambda[(1-p)\widetilde{G}(\rho U_J) + p\widetilde{G}(\rho U_J + d\mathbf{I}_{\{W,P\}})]\right\}^{-1} \tag{10}$$

$$l_J = Q\left(\rho U_J\right) \tag{11}$$

for $J = W, M$.

[4] Formally, it results from the first order condition of the maximization involving equation 4.

The main implication of the equilibrium in terms of prejudice and gender differentials is that the value of unemployment is lower for women than for men, that is $U_W < U_M$.[5] This is a direct consequence of the presence of prejudiced employers. Even if men and women are ex-ante identical (same distribution of productivity and nonparticipation values; same arrival and termination rates; same discount rate and Nash-bargaining coefficient), women are disadvantaged because with positive probability they can meet prejudiced employers that will pay them a lower wage at same productivity and will hire them with a lower probability.

This equilibrium effect alone can broadly explain the empirical evidence presented in the Introduction, that is a negative wage differential for women and a lower female participation rate. The negative wage differential is driven by two components, easy to identify by looking at equation (5). First, women working for prejudiced employers receive lower wages than men with the same productivity because of the disutility component d. Second, both women working for prejudiced and unprejudiced employers receive lower wages than men because of the lower outside option ($\rho U_W < \rho U_M$). This second "spillover" effect (Flabbi 2005) implies that wage discrimination is present also at unprejudiced employers. It is an equilibrium effect that derives from the endogenously weaker bargaining position of women at any type of employers.

Implications for the unemployment rate differential are ambiguous since the female unemployment rates can be both higher, lower or equal to male unemployment rates depending on the intensity of prejudice.

These results establish the ability of the model to broadly match the evidence. The task of the identification strategy will be to show that the separate identification of the different components of the model is possible even when worker heterogeneity is introduced; in particular when we allow for different productivity distributions between men and women.

3 Empirical Specification and Identification

The minimum data requirements necessary to identify the fundamental parameters of the model are fairly standard: accepted wages for individuals currently employed (w_i), on-going unemployment durations for individuals currently unemployed (t_i), knowledge of gender and employment status for each individual.[6] As it will be clear in the identification discussion, some parametric assumptions are necessary. The productivity distribution $G(x)$ is assumed to be a location-scale distribution, with location parameter denoted by μ and scale parameter denoted by σ, while the nonparticipation value distribution $Q(z)$ is assumed to

[5] The proof is obtained by totally differentiating the implict function $\rho U(d, p)$ with respect to d and p.

[6] No employers' side information is necessary if symmetric Nash bargaining is assumed.

be an invertible one parameter distribution, with parameter denoted by :

$$
\begin{aligned}
G(x) &= G(x;\mu,\sigma) && (12) \\
Q(z) &= Q(z;\ ) && (13)
\end{aligned}
$$

The likelihood accounts for three different contributions: one for unemployed individuals defined over the density of unemployment duration; one for employed individuals defined over the density of accepted wages; and one for individuals that are not participating in the labor market. For each of these labor market states, individuals contribute differently depending on their gender and the year in which they are observed. As it will be clarified in the empirical section, each year represents a potentially different steady state. To simplify the notation, the identification discussion will not include all this heterogeneity, which is however fully exploited in the estimation. In this respect, it is useful to introduce the following definitions.

Definition 2 *Given a labor market environment characterized by the parameters*

$$ {}_0 \equiv \{\lambda, \eta, \alpha, b, \rho, \mu, \sigma, p, d, \ \} $$

the following cases are defined:

Homogenous case*: no employers' types and no workers' types*

$$ {}_1 \equiv \{\lambda, \eta, \alpha, b, \rho, \mu, \sigma, \ \} $$

Heterogenous case*: employers' types defined by prejudice and workers' types defined by gender*

$$ {}_2 \equiv \{\lambda, \eta, \alpha, b, \rho, \mu, \sigma, p, d, \ \} $$

Completely Heterogenous case*: types defined by prejudice and gender; heterogeneity in other labor market parameters over types*

$$ {}_3 \equiv \{\lambda_J, \eta_J, \alpha, b_J, \rho, \mu_J, \sigma_J, p, d, \ {}_J\}_{J=W,M} $$

Completely Heterogenous case with Over-Time Heterogeneity*: types defined by prejudice and gender; heterogeneity in other labor market parameters over types and over time*

$$ {}_4 \equiv \{\lambda_{Jt}, \eta_{Jt}, \alpha, b_{Jt}, \rho, \mu_{Jt}, \sigma_{Jt}, p_t, d_t, \ {}_{Jt}\}_{J=W,M;t=1,..,T} $$

These definitions are useful to understand which parameters are allowed to vary and in which context. For example, the discount rate is common to all agents in the economy and it is always assumed constant; the arrival rate instead can be assumed to be different for men and women, and to vary over time. The *Heterogenous case* has been used to illustrate the theoretical model in the previous section, while a *Completely Heterogenous case with Over-Time Heterogeneity* is the specification that will be used in estimation.

3.1 Likelihood Function

To specify the likelihood and discuss the identification strategy I will use the *Homogenous case*: it greatly simplifies the notation while allowing for a sufficiently detailed treatment. Each individual in the population is denoted by i and belongs to either one of three possible states: nonparticipation (NP), unemployment (U) and employment (E).

The first contribution concerns non-labor market participants $\{i \in NP\}$ and, as implied by the optimal decision rule, is equal to:

$$P\left(i \in NP\right) = \widetilde{Q}\left(\rho U\right) \tag{14}$$

The second contribution concerns labor market participants currently unemployed $\{i \in U, i \notin NP\}$. Their contribution is a density defined over unemployment durations. By the stationarity of the environment, the hazard rate out of unemployment is constant:

$$h = \lambda \widetilde{G}(\rho U) \tag{15}$$

implying a negative exponential density for on-going unemployment durations:

$$f_u\left(t_i, i \in U, i \notin NP\right) = h \exp(-h t_i) \frac{\eta}{\eta + h} Q\left(\rho U\right) \tag{16}$$

The last two terms in equation (16) are introduced to take into account that we observe unemployment duration only for individuals that participate in the labor market and are currently unemployed.

Finally, the third contribution concerns employed individuals $\{i \in E, i \notin NP\}$ and it is defined over accepted wages. The density is obtained by exploiting the equilibrium mapping between match-specific productivity and wages:

$$f_e(w_i, w_i > \theta^*, i \in E, i \notin NP) = \frac{\frac{1}{\alpha} g\left(\frac{w - (1-\alpha)\rho U}{\alpha}\right)}{\widetilde{G}(\rho U)} \frac{h}{\eta + h} Q\left(\rho U\right) \tag{17}$$

The joint likelihood combines this information leading to:

$$\ln L\left(w_i, t_i, U, E, NP; \;{}_1\right) = \tag{18}$$

$$= \sum_{i \in U} \ln\left[f_u(t_i, i \in U, i \notin NP)\right] + \sum_{i \in E} \ln\left[f_e(w_i, w_i > \theta^*, i \in E, i \notin NP)\right] + \sum_{i \in NP} \ln P\left(i \in NP\right)$$

$$= (N_U + N_E) \ln\left(\frac{h}{\eta + h}\right) + N_U \ln \eta + (N_U + N_E) \ln\left[Q\left(\rho U\right)\right] + N_{NP} \ln\left[\widetilde{Q}\left(\rho U\right)\right]$$

$$- h \sum_{i \in U} t_i + \sum_{i \in E} \ln\left[\frac{\frac{1}{\alpha} g\left(\frac{w - (1-\alpha)\rho U}{\alpha}\right)}{\widetilde{G}(\rho U)}\right]$$

where $N_{E,U,NP}$ denotes the numerosity of the corresponding set. A similar procedure can be used to extend the derivation of the likelihood for more heterogenous cases.

3.2 Identification

Conditioning on participation, the identification of the *homogenous case* is based on Flinn and Heckman 1982, the identification of the *heterogenous* and *completely heterogenous cases* is developed in Flabbi 2005.

Starting with the *homogenous case*, the parameters of the model are $_1 \equiv \{\lambda, \eta, \alpha, b, \rho, \mu, \sigma, \ \}$. First notice that the symmetric bargaining assumption implies $\alpha = 0.5$. This assumption is consistent with the common discount rate shared by workers and employers but it is admittedly motivated by the difficulty of identifying this coefficient using only labor market supply information.[7] The second observation is that b and ρ enter the likelihood (18) only through the reservation value ρU. This fact implies the two parameters can only be jointly identified but has the advantage of generating a useful reparametrization: directly identify ρU in place of (b, ρ). Also notice that λ enters the likelihood only through h. As a first step, then, consider the identification of the parameters $\{h, \eta, \rho U, \mu, \sigma, \ \}$.

The reservation value may be identified and non-parametrically estimated by the minimum observed wage:

$$\widehat{\rho U} = \min_{i \in E} \{w_i\} \tag{19}$$

This estimator is strongly consistent[8] allowing for the estimation of the other parameters on the concentrated likelihood using $\widehat{\rho U}$ in place of ρU.

The hazard rate and the termination rate are identified by unemployment durations information. In particular, the maximum likelihood estimators are:

$$\widehat{h} = \left(\frac{\sum_{i \in U} t_i}{N_U} \right)^{-1} \tag{20}$$

$$\widehat{\eta} = \frac{N_U}{N_E} \widehat{h} \tag{21}$$

The nonparticipation distribution parameter is identified and easily estimated if $Q(z; \ )$ is invertible with respect to $\ $. In this application, I will assume a negative exponential distribution:

$$Q(z; \ ) = 1 - \exp(- \ z) \ , \ z > 0 \tag{22}$$

leading to the following maximum likelihood estimator:

$$\widehat{\ } = \frac{\ln (N/N_{NP})}{\widehat{\rho U}} \tag{23}$$

The estimator is also equivalent to the sample analog estimator implied by (14). Notice that under the current parametrization the only link between the

[7] Eckstein and Wolpin 1995 and Flinn 2005 discuss the issue in the context of a similar search-matching-bargaining model.

[8] Flinn and Heckman 1982 prove strong consistency of the estimator.

participation decision parameter and the other structural parameters is the reservation value $\widehat{\rho U}$. This implies that with respect to a model without a participation decision, such as Flabbi 2005 or Flinn 2006, there are no gains in terms of the identification of the other structural parameters by introducing the participation decision information. In practice, one piece of information is added, the participation rate, and one additional parameter is identified, . Still, the presence of the nonparticipation state generates gains in terms of the equilibrium effects that can be taken into account while performing policy experiments, effects that are particularly relevant when comparing over time.

Two additional parameters are left to be identified: μ and σ. Flinn and Heckman 1982 shows that no additional progress can be done in this direction without assuming a recoverable distribution for $G(x;\mu,\sigma)$: recoverability guarantees that the location and scale parameters of the distribution are identified by observing only a truncation of the distribution. In this application, I will assume lognormality

$$g\left(x;\mu,\sigma\right)=\frac{1}{\sigma x}\phi[\frac{\ln(x)-\mu}{\sigma}],\ x>0 \tag{24}$$

where ϕ and Φ denote the standard normal pdf and cdf. The lognormal distribution is recoverable and constitutes the most common assumption in the literature, since the empirical distribution of wages does actually resemble a truncated lognormal. Given the distributional assumption, it is possible to reparametrize the density (17) as:

$$\begin{aligned}
\frac{\frac{1}{\alpha}g\left(\frac{w-(1-\alpha)\rho U}{\alpha}\right)}{\widetilde{G}\left(\rho U\right)} &= \frac{\frac{1}{\sigma_0 w}\phi[\frac{\ln(w)-\mu_0}{\sigma_0}]}{\widetilde{\Phi}\left[\frac{\ln(\rho U)-\mu_0}{\sigma_0}\right]} && (25)\\
\mu_0 &= \alpha\mu+(1-\alpha)\,\rho U && (26)\\
\sigma_0 &= \alpha\sigma && (27)
\end{aligned}$$

where μ_0 and σ_0 are the location and scale parameters of the lognormal wage offers distribution. However, the only observed distribution is the accepted wage distribution that is actually the truncation reported in (25). By knowledge of the truncation point $\left(\widehat{\rho U}\right)$ and recoverability, μ_0 and σ_0 are identified. With the additional knowledge of α, it is then possible to identify μ and σ by solving the two linear equations (26) and (27).

At this stage, it is also possible to separately identify the two components of the hazard rate: the arrival rate and the probability to accept the match. This is relevant because it allows for the identification of the parameter λ that constitutes the real primitive parameter of the model in place of the reduced form parameter h. By equation (15), the arrival rate is identified by:

$$\widehat{\lambda}=\frac{\widehat{h}}{\widetilde{G}(\widehat{\rho U};\widehat{\mu},\widehat{\sigma})} \tag{28}$$

This step concludes the identification of the parameters $\{\lambda, \eta, \rho U, \mu, \sigma, \ \}$ under the *homogenous case*.

The *heterogenous case* implies the identification of two additional parameters p and d. Flabbi 2005 provides the proof and a detailed discussion. The main result from this analysis is that beside the recoverability condition defined by Flinn and Heckman 1982, two additional requirements must be imposed on the match-specific productivity distribution: it should be characterized by a location and a scale parameters and should allow for identification of its finite mixture. These restrictions rule out some fairly common distributions, as the Pareto and the Exponential, but still allow for identification using the more frequently used Lognormal and the quite popular Normal and Gamma distributions. These restrictions derive from the following identification strategy, formally proved in the Appendix of Flabbi 2005. The observed earnings distribution of women is a mixture of accepted jobs at prejudiced and unprejudiced employers. The proportion in which the two distributions are combined in this mixture is exactly the proportion of prejudiced employers p. Under the assumption of $G(x; \mu, \sigma)$ being characterized by a location and scale parameters, the two distributions composing the mixture have different location parameters but share the same scale parameters. If $G(x; \mu, \sigma)$ allows for the identification of its finite mixtures, we can then directly identify four pieces of information: the proportion of the two distributions in the mixture, the common scale parameter, and the two different location parameters. From the scale parameter and one of the two location parameters we can identify (μ, σ) with the same procedure described equations (25)-(27). As mentioned, the proportion in the mixture directly identifies p. Finally, thanks to the linearity of the mapping between match-specific productivities and the accepted wages, equation (5), the difference between the two location parameters is exactly equal to the discrimination intensity d, thus guaranteeing its identification.

The identification of the *completely heterogenous case with over-time heterogeneity*, which is the specification actually estimated in the paper, is obtained by repeating the previous identification strategy on each year and by assuming that each year is the realization of a (potentially different) steady state.

4 Data

The minimum data requirements necessary to identify and estimate the model are earnings, unemployment durations, employment status and gender of individuals over a sufficiently long period of time. Data from the *Annual Social and Economic Supplement* (ASES or March supplement) of the Current Population Survey (CPS) contains this information, providing a representative sample of the U.S. labor market. The construction of the estimation sample should also address two additional issues: the ex-ante homogeneity conditioning on gender assumed by the model and the steady state equilibrium assumption necessary to specify the likelihood function.

The homogeneity issue is addressed by extracting a sample homogenous with respect to observable characteristics.[9] In particular, the estimation sample is constituted by individuals who are white, 30 to 55 years old, with a College degree or more, and that are not self-employed. Some specifications will also control for out of labor market characteristics, such as marital status and presence of young children.

The steady state assumption issue is addressed by appropriately selecting the years on which the estimation is implemented. The ASES survey is available from 1979 to 2005. However, 1979 and 1980 have an unusually high percentage of top-coded values on earnings variables. Moreover, the proportion of top-coding is very different by gender and significantly affects the shape of the accepted earnings distribution. I therefore limited the sample to 1981-2005. Out of this sample I have selected three years - 1985, 1995 and 2005 - because they satisfy the following criteria: (i) they are equally spaced over-time and enough far away to describe different steady-states; (ii) they are not boom or recession years and therefore they seem more appropriate to describe a steady state; (iii) 1985 is in the middle of the convergence period, 1995 is in the middle of the slowdown and 2005 is the last year available.

Descriptive statistics on the estimation sample are reported in Table 1 after some trimming is performed on the raw data.[10] The first three rows report mean and standard deviation of hourly earnings for the overall sample of employed individuals, for employed women and for employed men. Hourly earnings are obtained either by using the value directly reported in the CPS survey or by computing the value dividing *weekly earnings* by the *usual hours worked per week*.[11] Earnings are deflated using the *Consumer Price Index*[12] and they are expressed in 2005 US dollars. The data show a differential with a trend similar to the evidence presented in the Introduction: a decrease from 1985 to 1995 and then some variability from 1995 to 2005. The differential is always higher than 20%.

The second group of variables reported in Table 1 are unemployment du-

[9] It is an assumption fairly common in the literature, see for example Flinn 2005 and Bowlus and Eckstein 2002.

[10] Validation data and the observation of extremely low hourly earnings suggest measurement errors asymmetric by gender in the low tail (Bollinger 1998). The solution used in this paper is to trim the bottom 5% of the sample (Bowlus 1997). Changing of the bottom trimming point has a direct impact on the reservation wages but the other parameters are reasonably stable, as already found in Flabbi 2005. The difference here is that the range over which the estimates are stable is smaller (1% to 5% as compared to 1% to 10% in Flabbi 2005) due to sensitivity of the estimates for the year with the lower numerosity (1985). Top-coding on income and earnings variables is also present in the CPS data. I have run the estimates both including and excluding top-coded values and the main results do not change. The estimation sample presented in the paper includes top-coded values.

[11] There is a relative high proportion of missing values in the *usual hours worked* variable. To reduce the impact of this problem, for those individuals that are not paid by the hour, do not report hourly earnings, and have missing value on the *usual hours worked* variable I compute hourly earnings by dividing weekly earnings by 45 if they work full-time and by 25 if they work part-time. Hourly earnings obtained in this way are less than 2% of the sample. I have also estimated the model excluding individuals with imputed hours obtaining very similar estimates to those reported in Table 2.

[12] Consumer Price Index for All Urban Consumers; Series Id: CUUR0000SA0.

rations. They are measured in months and they are obtained by rescaling the original weekly unemployment durations. As found in other works, unemployment durations are lower for women, even if the differential is shrinking over the twenty years period under consideration. Finally, the bottom of Table 1 reports unemployment rates and participation rates.[13] Interestingly unemployment rate ranking reverses: from an higher unemployment rate for women in 1985 and 1995 to a lower unemployment rate in 2005. Participation rates are always lower for women by at least ten percentage points. It is worth noticing that this "homogenous" group of individuals has generally a lower unemployment rate and a higher participation rate than the population not selected with respect to the homogeneity controls. The participation rate differential is also lower than the one computed on more aggregate samples, as for example the one reported in Figure 2.

5 Results

Maximum Likelihood estimates are implemented on the sample obtained by pooling the three years together. Beyond some efficiency gains, the joint estimation allows for the introduction of restrictions across years to test different specifications or for the isolation of specific equilibrium effects. The preferred specification - reported in Table 2 - estimates a reparametrization of the original model where in place of directly estimating the disutility d, the following parameter k is estimated:

$$k \equiv \frac{d}{E\left(x_i | i \in M\right)} = \frac{d}{\exp\left[\mu_M + 0.5\sigma_M^2\right]} \tag{29}$$

It defines the *relative* disutility of hiring a woman, that is the ratio between the absolute value of the discrimination intensity and an indicator of the average productivity in the economy. Since disutility is measured on the same scale of productivity and since productivity is changing over time, it is more instructive to use such relative measure instead of the absolute value d. The interpretation is that k represents the proportion of an average match-specific productivity that a prejudice employer is willing to sacrifice to hire a men instead of a women generating the same x.

The reparameterization is also useful as a scale normalization for computational purposes and as a meaningful way to impose parameters restrictions across years. In particular, I will restrict this ratio to be the same for all the three years. Notice that this does not mean to impose the same disutility but that the ratio between the disutility and a synthetic measure of productivity in a given year should be the same. The behavioral assumption is that preferences

[13] The labor market status (employment, unemployment and nonparticipation) is obtained by a set of questions organized by the CPS team in the *monthly labor force recode* variable which directly assings each individual in the sample to employment, unemployment or not-in-the labor force status. Excluded from the universe are kids and individuals in the armed forces.

are very slow to adjust and therefore the average *relative* disutility represented by the parameter k is not very likely to change over time. A Likelihood Ratio (LR) test performed on the joint sample does *not* reject this restriction and it is reported at the bottom of Table 2.[14] This restriction also helps obtaining a more precise estimate of the disutility parameter for 2005 and makes the comparison across time easier: by estimating a common *relative* disutility the proportion of prejudiced employers becomes the parameter that describes the evolution of prejudice. This is also the parameter that should be affected by over-time changes in the original Becker's model.

5.1 Estimation results

Estimation results are presented in Table 2: the specification is - using the terminology introduced in Section 3 - a *completely heterogenous case with over-time heterogeneity.* Estimates are precise and the restriction on k is not rejected (P-value: 0.799). The main result is that the proportion of prejudiced employers is decreasing over time: the proportion decreases at a faster rate in the last decade than in the previous decade, moving from 69.2% in 1985, to about 57.1% in 1995, to about 32.2% in 2005.[15] The relative disutility is estimated to be about 32%, a value comparable with estimates obtained by Bowlus and Eckstein 2002 on a sample of black and white workers from NLSY, and by Flabbi 2005 on CPS data for 1995 on men and women.

Arrival rates are estimated to be always higher for women while termination rates are higher for women in 1985 and 1995 but not in 2005. These results are similar to Bowlus 1997 that estimates higher arrival and termination rates for women on a sample extracted from NLSY 1979-1991. Hazard rates out of unemployment are not exogenous because they are the product of the exogenous arrival rate and the endogenous probability to accept the match. The estimates imply higher hazard rates for women but a strong convergence over time: the positive differential for women decreases from 95.8% in 1985, to 68.5% in 1995, to 16.1% in 2005. The participation coefficient implies an average value of nonparticipation higher for women in all years: \$ 4.7 per hour for women and \$ 2.8 dollars per hour for men in 1985, to 5.1 dollars for women and 3.2 for men in 2005.

Productivity trends are better evaluated by looking at predicted values rather than at the primitive μ and σ. The first two rows in Table 3 report the predicted average productivity for the two groups. Men's productivity is steadily increasing while women's productivity is increasing from 1985 to 1995 and decreasing from 1995 to 2005. This last result is surprising and may be due to the limited controls for sample selection in the model. However, participa-

[14] Other specification tests have been performed: in particular a "full prejudice" model that imposes a proportion of prejudiced employers equal to one for all years is rejected with respect to a specification without restrictions (P-value = 0.047).

[15] The result of a decreasing proportion in p is robust to the removal of the cross-restriction on k. However, differences are generally not significant: for example, the restriction $p_{85} = p_{95} = p_{05}$ is not rejected at conventional significance levels (P-value=0.398).

tion rates for this homogeneous group of individuals do not change much over the period and they actually decrease both for men and women from 1995 to 2005. Hence, to rationalize the results, a sample selection story should propose a mapping from out-of-labor market characteristics to labor market productivity that could potentially change over time.[16] To partially address the issue, I have performed estimates on samples homogenous with respect to observables usually correlated with sample selection. Results are reported in Table 4 and will be discussed below.[17]

The second group of variables presented in Table 3 concerns average offered and accepted earnings by gender. The gender differential in average accepted earnings within this relatively homogenous group of workers is comparable with the conditional differential presented in the Introduction and it is conceptually similar to results obtained in the wage regressions literature. In addition, the model allows for an estimation of the differential in *offered* earnings: this is in many respect a more relevant variable to compare groups or returns, as pointed out among others by Eckstein and Wolpin 1995. Offered earnings are closer to be a primitive of the model[18] since accepted earnings are an endogenous equilibrium outcome resulting from the optimal decision rule. Since these equilibrium effects are stronger for women, the estimates suggest that simply looking at accepted wages conditioning on an homogenous sample may underestimate wage discrimination. In other words, women are receiving offers that are worse, with respect to men, than the ones that are accepted but they are able to partially overcome these worse offers by reacting optimally. In 2005, for example, we observe a negative wage differential of 23.1% but the underlying differential in wage offers was higher and equal to 26.5%.

Table 3 is also useful to evaluate the fit of the model by comparing predicted moments with the sample moments reported in Table 1. The fit on first moments is always very good for female accepted earnings while average male earnings are slightly overestimated in 1995 (28.24$/h instead of a sample value of 27.89$/h). Unemployment rates are interesting because they are an overidentifying restrictions implied by the equilibrium of the model and therefore matching them is a test for the model. The bottom of Table 3 shows a perfect fit implying that the model is not rejected on this ground.

Table 4 presents selected parameters estimates from the same specification

[16] Blau and Kahn 2004 presents some evidence that women moved from being positively selected in the labor market in the 1980s (i.e. more productive women were the ones actually supplying labor) to being slightly negatively selected in the 1990s.

[17] Another potentially relevant selectivity issue over time concerns selection into education: if the proportion or the composition of individuals holding a college degree changes over time, we are effectively estimating on different groups, thus affecting the comparison across years. To address the issue I have computed the proportion of individuals holding a college degree by cohorts and estimated the model over different cohorts. The results are that the proportion of individuals with College degree increases significantly from 1985 to 1995 but not from 1995 and 2005 and that if the point estimates change slightly, the main results are confirmed.

[18] The primitive in the model is the productivity distribution while also offered wages are endogenous because they are the outcome of the bargaining process. However, the offered wage distribution is computed before the optimal decisons on accepting or rejecting the match and in this sense is *closer* to be a primitive of the model than the accepted wages distribution.

reported in Table 2 but on different samples. The estimation samples are selected to be homogenous with respect to some out-of-labor market variables that may have gender-specific impacts. To present results in a succinct way, Table 4 does not report all the estimated structural parameters but only the prejudice parameters (p, k) and an indicator of productivity differentials: the ratio of average female/male ex-ante productivity. The first set of results is obtained by imposing no controls and simply reports the estimates from Table 2. The second set of results refers to a sample of individuals married at least once at the time of the interview. The third to a sample of individuals living with a family member under age 18 and the fourth imposes both requirements. Results confirm the general pattern of a decreasing proportion of prejudiced employers over time. However, when considering only individuals with young family members, there is no appreciable difference between the proportion of prejudiced employers in 1995 and 2005. The general trend found in Table 2 in terms of productivity differentials is also confirmed: productivity differentials increase between 1995 and 2005. Finally, the comparison of results for a given year shows quite robust point estimates for 2005, a larger variation in 1995 and an intermediate value for 1985.

5.2 Counterfactual experiments

Different explanations have been proposed to account for the slowdown in the convergence of gender differentials in the 1990s.[19] One contribution of this paper is to evaluate whether changes in prejudice over time may have played a role. To better assess this impact it is interesting to perform counterfactual experiments that isolate the impact of prejudice on labor market variables. The proposed experiments set all the structural parameters for each group to the estimated male values for 2005 while changing the parameters that describe prejudice: the proportion of prejudiced employers p and the intensity of discrimination relative to productivity k. For each of these environments, a new equilibrium is computed and the corresponding labor market variables for each gender are obtained allowing for an evaluation of the impact of prejudice once equilibrium effects are taken into account. The specification used in the experiments corresponds to the *heterogenous case* defined in Section 3.

Table 5 reports the results of the first experiment: the proportion of prejudiced employers is set at the estimated values for 1985, 1995 and 2005 and then is set at two additional and arbitrary low values in columns 4 and 5. The bottom part of the table reports the gender differentials generated in these counterfactual environments on statistics of the relevant labor market variables. For example, the second row in the first column reports that women would experience a negative differential of 27.5% in average offered earnings in an environment where they are ex-ante equal to men and about 69% of the employers are prejudiced. The following rows report differentials computed on accepted earnings, unemployment durations, unemployment and participation rates. The

[19] For example Fortin and Lemieux 2000 relates this trend to changes in male wage inequality.

last row reports the differential on the value of unemployment U: this may be considered a welfare measure since it is the value to the worker of participating in the specific labor market under consideration.[20]

Results confirm the expected convergence between men and women as the proportion of prejudiced employers decreases; the convergence is achieved on all the labor market variables. Looking at the value of unemployment as a welfare measure to summarize them all, the differential would move from more than 50% in 1985 to about half of this value in 2005 to about 3% in the hypothetical $p = 5\%$ environment. Additionally, what is interesting in this exercise is the speed of the convergence as p decreases, once the equilibrium effects are taken into account. These equilibrium effects are important because they have the potential to magnify the impact of a reduction in p. For example a decrease in p affects the accepted earnings differentials by the direct impact on the wage schedule offered by the prejudiced employers, by the equilibrium impact on the reservation wage and by the equilibrium impact on the outside option in bargaining with any type of employer including the unprejudiced. The result is that the impact on labor market outcomes can be highly non-linear as the proportion of prejudiced employers decreases. As a result we observe that for high values of p, as in 1985 and in 1995, female workers are able to partially reduce the impact of prejudice: the accepted earnings differential - about 22% in 1985 and 17% in 1995 - is much smaller than the offered earnings differential - about 27% in 1985 and 23% in 1995. These relatively high proportions of prejudiced employers are also the only ones able to generate earnings differentials comparable to the ones observed in the data. For a proportion of prejudiced employers of about 30%, as in 2005, the observed earnings differential would be comparable to the lowest in the world, such as values typically experienced by Scandinavian countries. For a proportion of less than 5%, convergence in labor market variables is basically achieved. This implies that, at the current rate of reduction in p, it will require slightly more than ten years to achieve full convergence. However, nothing guarantees that this rate will remain constant: for example at the rate of the previous decade, from 1985 to 1995, it will require more than twenty years to achieve convergence.

Table 6 reports the results of the second experiment: as before all the structural parameters are fixed at the 2005 male values but now also the proportion of prejudiced employers is fixed at the estimated 2005 value while it is the relative discrimination intensity that is changing. Since only one value of k is estimated for the three time periods, it is necessary to generate a meaningful variation of k. The change is obtained as follows. A summary measure of the amount of prejudice in the economy is the average discrimination intensity which, under the model parametrization is simply:

$$E(d) = p \times d$$

The year-specific $k's$ are then obtained as those values that generate the estimated average discrimination intensity once all the parameters excluding k

[20]See Flinn 2002 for a discussion on welfare measures in a search environment.

(but including p) are fixed at the 2005 value.[21] The logic of the experiment is as follow: given the same average discrimination intensity, does it matter that this amount is generated by many prejudiced employers with low intensity or by few prejudiced employers with high intensity? The first row of Table 6 reports the estimated average discrimination intensity and the second row the counterfactual $k's$. As before, new equilibria are generated using these $k's$ and for each equilibrium the distribution on the relevant labor market variables are obtained. The bottom panel reports the implied differentials on these labor market variables and on the welfare measure U.

As expected, we observe convergence between men and women as time passes. Comparing Table 5 and 6 we obtain an interesting result: the reduction in discrimination intensity improves women's condition more than the reduction in p (at given $E(d)$). For example, in 1995 the welfare/value of unemployment differential is 43.5% after the reduction in p from 1985 but 35.9% after the reduction in k. By looking at labor market variables, we see that this is a result of the optimal acceptance rule: the differential on the accepted wages is significantly lower in Table 6 than it is in Table 5. Also the other labor market variables register a relatively better outcome for women but it is on the accepted wages that most of the action is taking place. These results are also comforting in terms of the separate identification of p and k because they emphasize that the proportion of prejudiced employers and the intensity of discrimination, albeit having a degree of substitutability, have different impacts on observed labor market outcomes.

6 Conclusion

This paper was set to address two issues: first, to assess whether changes in prejudice over time could explain the slowdown in the convergence of earnings between men and women in the 1990s[22]; second, to assess the validity of Becker's model of taste discrimination in explaining groups differentials by providing an estimate for the proportion of prejudiced employers over time .

The paper estimates that the proportion of prejudiced employers has been decreasing in the last twenty years. This reduction has happened at an increasing rate both in absolute and relative terms: the proportion has gone from 69% in 1985, to 57% in 1995, to 32% in 2005. The estimates are obtained by developing a search model with matching and bargaining and by allowing men and women to differ in terms of behavior (participation rate, arrival rate of job offers, termination rate of the employment relation) and match-specific productivity.

With respect to the first objective, the estimation of a decreasing proportion of prejudiced employers implies that employers' prejudice does not explain

[21] Therefore, by construction, the third column of Table 6 should be equal to the third column of Table 5 because they have the same underlining parameters even if one is obtained by keeping k fixed and moving p and the other by keeping p fixed and moving k.

[22] See for example Blau and Kahn 2004, Fortin and Lemieux 2000 and Eckstein and Nagypal 2004.

the slowdown. Moreover the decrease is faster over the 1995-2005 period than over the 1985-1995 period reinforcing the argument that a change in prejudiced behavior cannot be responsible for the slowdown in the converge of the labor market position of men and women. Estimation results attribute most of the slowdown to the faster increase in average male productivity with respect to female productivity: the ratio of average female productivity over average male productivity has decreased from about 95% in 1995 to about 82% in 2005. One possible explanation for this increase in productivity differentials is the change in the sample selection of women participating in the labor market. Blau and Kahn 2004, for example, tentatively suggests that women may have moved from being positively selected in the labor market to be negatively selected in the labor market. However, if such a process is taking place, it is really endogenous with respect to the labor market environment, including its proportion of prejudiced employers. A better understanding of the issue therefore requires a complete general equilibrium treatment of some of the channels developed and estimated in this paper.

With respect to the second objective, the estimated decreasing trend in the proportion of prejudiced employers is broadly consistent with the predictions of the Becker's taste discrimination model. However, the implied adjustment process is estimated to be very slow, covering at least three decades and contradicting formulations of the model developed in perfectly competitive labor markets.

References

[1] Arrow, K. (1973), "The Theory of Discrimination" in O. Ashenfelter and A. Rees (eds.) *Discrimination in Labor Markets*, Princeton, N.J.: Princeton University Press, 3-33.

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

[3] Blau, F.D. and L.M. Kahn (2004) "The US Gender Gap in the 1990s: Slowing Convergence", *NBER Working Paper* 10853.

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

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

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

[7] Eckstein, Z. and E. Nagypal (2004) "The Evolution of U.S. Earnings Inequality: 1961-2002" *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(2): 10-29.

[8] 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.

[9] Flabbi, L. (2005), “Gender Discrimination Estimation in a Search Model with Matching and Bargaining", *IZA Discussion Paper* 1764.

[10] Flinn, C. (2006), “Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining and Endogenous Contact Rates”, *Econometrica*, 73: 1013-1062.

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

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

[13] Fortin, N.M. and T. Lemieux (2000) "Are Women’s Wage Gains Men’s Losses? A Distributional Test" *American Economic Review Papers & Proceedings*, 90(2): 457-460.

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

[15] O’Neill (2003) "The Gender Gap in Wages, circa 2000" *American Economic Review Papers & Proceedings*, 93(2): 309-314.

Table 1: Descriptive Statistics on Estimation Sample

	1985		1995		2005	
	Mean	SD	Mean	SD	Mean	SD
$w_i \vert i \in E$	24.257	10.253	25.286	11.756	27.831	14.236
$w_i \vert i \in E, i \in W$	20.090	9.413	22.133	10.452	24.378	12.380
$w_i \vert i \in E, i \in M$	27.202	9.794	27.891	12.134	31.342	15.119
Differential (%)	-26.14		-20.64		-22.22	
$t_i \vert i \in U$	4.846	6.448	5.394	5.735	4.846	5.539
$t_i \vert i \in U, i \in W$	3.209	3.798	4.220	4.759	4.432	5.175
$t_i \vert i \in U, i \in M$	6.279	7.900	7.106	6.656	5.146	5.825
Differential (%)	-48.90		-40.61		-13.87	
u	0.023		0.026		0.030	
u_W	0.026		0.034		0.025	
u_M	0.021		0.019		0.034	
Differential (%)	23.18		73.99		-28.13	
l	0.902		0.913		0.884	
l_W	0.804		0.852		0.813	
l_M	0.977		0.968		0.961	
Differential (%)	-17.71		-11.98		-15.40	
$N_U + N_E$	1919		2283		2741	

Note: Definitions: w_i =hourly earnings in 2005 dollars, t_i =unemployment duration in months, u =unemployment rate, l =participation rate. The participation rate is computed before trimming the sample.

Table 2: Maximum Likelihood Estimates

	1985	1995	2005
λ_M	0.1598	0.1413	0.1962
	(0.0326)	(0.0288)	(0.0422)
λ_W	0.3291	0.2467	0.2446
	(0.0721)	(0.0418)	(0.0422)
η_M	0.0035	0.0028	0.0069
	(0.0010)	(0.0008)	(0.0014)
η_W	0.0084	0.0082	0.0057
	(0.0026)	(0.0020)	(0.0014)
μ_M	3.6701	3.6966	3.7757
	(0.0149)	(0.0166)	(0.0176)
σ_M	0.4880	0.5685	0.6151
	(0.0111)	(0.0124)	(0.0135)
μ_W	3.6425	3.7217	3.6282
	(0.0884)	(0.0622)	(0.0428)
σ_W	0.3858	0.4175	0.5286
	(0.0305)	(0.0263)	(0.0235)
p	0.6917	0.5714	0.3224
	(0.1902)	(0.1498)	(0.1021)
k		0.3254	
		(0.0414)	
ρU_M	10.4765	8.9373	10.2587
ρU_W	7.6121	6.8224	8.5550
$_M$	0.3601	0.3851	0.3162
$_W$	0.2141	0.2800	0.1960
$Loglik$		-26523.13	
LR Test (P-value):			
$k_{85} = k_{95} = k_{05}$		0.799	

Note: Asymptotic standard errors in parentheses. Joint estimation on all years.

Table 3: Predicted Values

	1985	1995	2005
$E(x_i \mid i \in M)$	44.22	47.38	52.71
$E(x_i \mid i \in W)$	41.14	45.10	43.29
$E(w_i \mid i \in M)$	27.35	28.16	31.49
$E(w_i \mid i \in E, i \in M)$	27.41	28.24	31.69
$E(w_i \mid i \in W)$	19.40	21.56	23.16
$E(w_i \mid i \in E, i \in W)$	20.04	22.11	24.37
$E(t_i \mid i \in U, i \in M)$	6.280	7.108	5.145
$E(t_i \mid i \in U, i \in W)$	3.208	4.219	4.432
u_M	0.021	0.019	0.034
u_W	0.026	0.034	0.025
l_M	0.977	0.968	0.961
l_W	0.804	0.852	0.813

Note: Predicted values obtained from estimates reported in Table 2.

Table 4: Maximum Likelihood Estimates by different Out of Labor Market Characteristics

	1985	1995	2005
No controls:			
k		0.3254	
		(0.0414)	
p	0.6917	0.5714	0.3224
	(0.1902)	(0.1498)	(0.1021)
$\frac{E(x_i \mid i \in W)}{E(x_i \mid i \in M)}$	0.9303	0.9519	0.8212
Married at least once:			
k		0.3132	
		(0.0564)	
p	0.8697	0.5531	0.3791
	(0.2866)	(0.1670)	(0.1187)
$\frac{E(x_i \mid i \in W)}{E(x_i \mid i \in M)}$	0.9806	0.9168	0.8048
With family members under age 18:			
k		0.3507	
		(0.1181)	
p	0.7661	0.3552	0.3312
	(0.2762)	(0.1964)	(0.1375)
$\frac{E(x_i \mid i \in W)}{E(x_i \mid i \in M)}$	0.9456	0.8636	0.7965
Married at least once and with family members under age 18:			
k		0.3494	
		(0.1255)	
p	0.7862	0.3516	0.3406
	(0.2780)	(0.1970)	(0.1438)
$\frac{E(x_i \mid i \in W)}{E(x_i \mid i \in M)}$	0.9512	0.8622	0.7918

Note: Asymptotic standard errors in parentheses. Joint estimation on all years. The specification is the same as in Table 2. The only difference between the estimates are the homogeneity controls imposed on the estimation sample with respect to the reported out of labor market characteristics.

Table 5: Counterfactual Experiment 1: Isolate the Impact of the Proportion of Prejudiced Employers

	1985	1995	2005	y'	y"
p	0.692	0.571	0.322	0.100	0.050
Differential (%):					
$E(w_i)$	-27.500	-22.659	-12.708	-3.918	-1.956
$E(w_i \mid i \in E)$	-21.973	-17.594	-9.292	-2.718	-1.342
$E(t_i \mid i \in U)$	9.046	8.252	5.424	1.832	0.928
u	8.709	7.947	5.229	1.768	0.896
l	-18.706	-12.615	-4.813	-1.090	-0.510
U	-53.150	-43.551	-24.103	-7.330	-3.648

Note: Results obtained by fixing all the remaining parameters except p to the men's values in 2005. For each value of p in this environment a new equilibrium is computed and the corresponding values are reported in the table.

Table 6: Counterfactual Experiment 2: Isolate the Impact of the Intensity of Discrimination

	1985	1995	2005
$\widehat{E}(d)$	9.953	8.809	5.530
k	0.586	0.518	0.325
Differential (%):			
$E(w_i)$	-22.234	-19.836	-12.708
$E(w_i \mid i \in E)$	-13.556	-12.652	-9.292
$E(t_i \mid i \in U)$	13.627	11.391	5.424
u	13.099	10.957	5.229
l	-10.543	-8.945	-4.813
U	-39.461	-35.889	-24.103

Note: Results obtained by fixing all the remaining parameters except k to the men's values in 2005. Different values of k for each year are obtained by matching the estimated average discrimination intensity $\widehat{E}(d)$ while keeping p fixed at the 2005 value. For each value of k in this environment a new equilibrium is computed and the corresponding values are reported in the table.

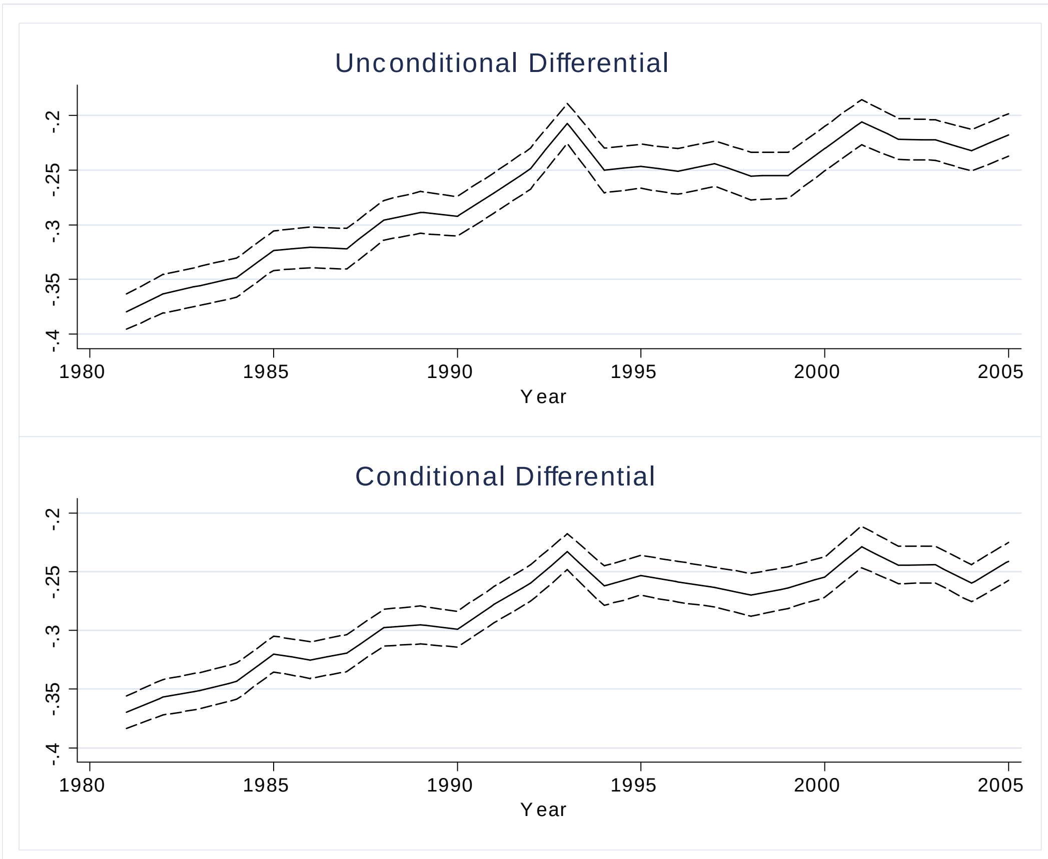

Figure 1: **Gender Earnings Differential over Time: Point Estimates and 95% Confidence Interval.**

Note: Data are from the Annual Social and Economic Supplement (March Supplement) of the Current Population Survey (CPS). Results report the estimated coefficient and the 95% confidence interval of a dummy =1 if the individual is a woman in a regression of log hourly earnings on: (i) a constant and the dummy woman for the unconditional differential case (top panel); (ii) a constant, the dummy woman, three educational dummies, age linear and squared, a dummy for marital status and a dummy for presence of children younger than 18 for the conditional differential case (bottom panel). Complete results available upon request.

Figure 2: **Gender Differential in Labor Market Participation: Participation Rate and Employment Rate conditioning on Participation.**

Note: Data from the Annual Social and Economic Supplement (March Supplement) of the Current Population Survey (CPS). Results report the ratio $\frac{rate_W - rate_M}{rate_M}$ where $rate$ means participation rate (solid line) and employment rate (dotted line); and W denotes women and M men.