E C O N O M I C S B U L L E T I N

A theoretical model of wage discrimination with inspection fines

Juan Prieto-Rodriguez
Universidad de Oviedo, Spain

Juan Gabriel Rodriguez
Universidad Rey Juan Carlos, Madrid, Spain

Rafael Salas
Universidad Complutense de Madrid, Spain

Abstract

In neoclassical models, workers are classified a priori into discrimination groups. We develop a probabilistic model of wage discrimination in which workers need not be classified a priori. Our model is a generalization of the standard framework, whereas Becker's model is an extreme case. A second implication is that the traditional approach to measuring discrimination (the Oaxaca–Blinder approach) must be modified to take into account this probabilistic framework.

This research has benefited from the Spanish Ministry of Science and Technology Projects SEC2003-08397 and MEC-04-SEJ-04065. The usual disclaimer applies.
Citation: Prieto-Rodriguez, Juan, Juan Gabriel Rodriguez, and Rafael Salas, (2008) "A theoretical model of wage discrimination with inspection fines." *Economics Bulletin,* Vol. 10, No. 3 pp. 1-9
Submitted: September 24, 2007. **Accepted:** March 12, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume10/EB-07J70003A.pdf

1. Introduction

We propose a nondeterministic model of wage discrimination in which workers need not be classified a priori into discrimination groups. In this sense, our model extends the neoclassical model of wage discrimination (see Becker, 1957 and Arrow, 1973) in which, typically, all females (or blacks) are assumed to be discriminated against. We show that Becker's model is an extreme case of our model when workers are deterministically classified. A straightforward implication is that the traditional Oaxaca–Blinder approach (Oaxaca, 1973 and Blinder, 1973) to measuring discrimination needs to be revised.

We start with some intuition that we formalize in subsequent sections. Suppose that being female, black, foreign or ugly are disadvantages according to employers' preferences. Assume also that only an unknown critical number (three in our model), rather than all, of those characteristics need to be possessed for workers to be discriminated against. The idea behind this assumption is that the more 'negative' characteristics a worker has, the more likely is such a worker to be discriminated against.

In Figure 1, workers are divided into two groups: nondiscriminated against workers (group 1) and discriminated against workers (group 2). If we adopt, for instance, a gender view, most discriminated against workers are female, although men who are black, foreign or ugly are also discriminated against. If all parties (employers, employees, government, judges, lawyers, researchers, and so on) have perfect information about the relevant characteristics used to classify workers, the Oaxaca–Blinder approach to measuring discrimination is appropriate.

However, this information is often known only by employers, sometimes by employees, but not by third parties, particularly for nonobservable characteristics such as subjective evaluations of beauty, empathy and other personal circumstances. This information problem cannot be solved by improving the quality of the data. Another source of unobservable information for third parties is that available surveys may not include relevant variables such as religion and political affinity. Furthermore, to avoid the penalties imposed by authorities, employers try to hide their discriminatory conduct by modifying their behavior.

Given this asymmetric information problem, deterministic models may be biased. Thus, in Figure 1, neither are all females discriminated against nor are all males not discriminated against. Consequently, an antidiscrimination policy based on only one characteristic, such as gender, results in a misallocation of resources. This is because not all male and female workers are in groups 1 and 2, respectively.

[FIGURE 1 ABOUT HERE]

Even if policymakers account for all observable variables (such as gender, race and nationality), the presence of nonobservable characteristics (such as beauty) reduces the efficacy of policy. This simple example reveals the potential of nondeterministic approaches that do not rely on a priori deterministic classifications of workers.

2. Theoretical Model

We consider two possible cases. In the first case, the authorities do not prosecute against discrimination. Thus, there is no social control over discriminatory behavior and employers are relatively free to set wages and determine contracts. We show that this situation can be represented by a two-stage sequential game. In the second case, we introduce antidiscrimination policy, which gives rise to a three-stage sequential game.

2.1 No prosecutions against discrimination

Assume that there are two groups of workers based on employers' preferences: nondiscriminated against workers (group 1) and discriminated against workers (group 2). These groups comprise workers with mixed characteristics. Some variables are observable (such as race, age and gender), whereas others may only be known by employers. Therefore, groups are not directly observable by third parties. By contrast, in neoclassical models of discrimination, the strong assumption is made that there is perfect information about groups. This implies a deterministic classification of workers.
In the first stage, discriminatory employers determine optimal contracts to maximize their utility. Utility, U, depends on profits (π) and aversion to workers from group 2 (v). Employer utility is represented by the following function:

$$U(l_1, l_2, K) = \pi(l_1, l_2, K) - v(l_2) \tag{1}$$

where K is the amount of capital employed by the firm, and l_1 and l_2 denote the numbers of workers in groups 1 and 2, respectively. Normalizing the price of output to unity (without loss of generality) allows us to define profits as follows:

$$\pi(l_1, l_2, K) = F(l_1, l_2, K) - w_1 l_1 - w_2 l_2 - rK \tag{2}$$

where F is a strictly concave production function, r is the cost of capital, and w_1 and w_2 denote wages in group 1 (for nondiscriminated against workers) and wages in group 2 (for discriminated against workers), respectively. We assume that workers in both groups have the same human capital, and thus the same labor productivity, which is given by $F^{'} = \dfrac{\partial F(l_1, l_2, K)}{\partial l_1} = \dfrac{\partial F(l_1, l_2, K)}{\partial l_2}$. Given this assumption, wages can only differ because of discrimination.
The aversion component depends on the number of workers in group 2 and is represented by the aversion function $v(l_2)$, where $v' = \dfrac{\partial v}{\partial l_2} \geq 0$ and $v'' = \dfrac{\partial^2 v}{\partial l_2{}^2} \geq 0$. The restriction on the second derivative ensures that the function v is convex and, therefore, that utility (given by equation (1)) is increasing and strictly concave.
In the second stage, employees may accept or reject the contracts, depending on the reservation wage.
The first-order conditions for utility maximization are:

$$\frac{\partial U}{\partial l_1} = F' - w_1 = 0$$

$$\frac{\partial U}{\partial l_2} = F' - w_2 - v' = 0 \tag{3}$$

The first-order condition for capital is $\frac{\partial F}{\partial K} - r = 0$. We do not refer to this condition in the rest of the paper, as it is not relevant for our analysis. From the equations in (3), we can determine l_1 and l_2 if an explicit utility function is defined.

It is clear from (3) that employers pay higher salaries to nondiscriminated against workers; i.e., $w_1 > w_2$. The higher the marginal aversion to workers in group 2, the greater the wage discrimination. Discrimination disappears as v' approaches zero. Optimal conditions in (3) imply that the marginal utility of hiring an extra worker from either group at the equilibrium is zero. Furthermore, the marginal utility of exchanging a worker from group 1 for a worker from group 2 is also zero: the marginal benefit (the wage differential between the groups, $w_1 - w_2$) equals the marginal cost (v') at the equilibrium. Now we can analyze gender discrimination.[1]

As explained above, discriminated against workers (from group 2) are not easily identified if wage discrimination also depends on other characteristics (such as age, nationality, religion and appearance); there is an asymmetric information problem relating to these variables. Although this does not mean that workers are unaware of being discriminated against, independent observers can only assign probabilities to this. Hence, we characterize groups 1 and 2 as male and female workers as follows: $l_1 \equiv p_m l_m + (1 - p_f) l_f$ and $l_2 \equiv (1 - p_m) l_m + p_f l_f$, where p_m and p_f denote the probabilities of male and female workers being in groups 1 and 2, respectively, and l_m and l_f denote the numbers of male and female workers in the firm, respectively. Thus, group 2 may include men and group 1 may include women (see example above). Therefore, bias against women can be identified by comparing the estimated probabilities of males and females being discriminated against. (Prieto *et al.* (2006) estimate these probabilities and derive the associated wage equations by using finite mixture models.)

Because both women and men may be discriminated against (belong to group 2), aversion, v, depends on the numbers of women, l_f, and men, l_m. If there is a bias against female workers, then, $v_f' > v_m'$, where $v_f' = \frac{\partial v}{\partial l_f}$ and $v_m' = \frac{\partial v}{\partial l_m}$.

We can rewrite profits (given by equation (2)) in terms of male and female employment as follows:

$$\pi(l_m, l_f, K) = F(l_m, l_f, K) - w_m l_m - w_f l_f - rK \tag{4}$$

where the male average wage is $w_m \equiv [p_m w_1 + (1 - p_m) w_2]$ and the female average wage is $w_f \equiv [(1 - p_f) w_1 + p_f w_2]$.

Assuming that male and female workers have the same labor productivity and that the firm maximizes its utility function with respect to l_m and l_f, the first-order conditions are:

$$\frac{\partial U}{\partial l_m} = F' - w_m - v_m' = 0$$

[1] We focus on gender discrimination but the same analysis applies, *mutatis mutandis*, to race or any other kind of wage discrimination.

$$\frac{\partial U}{\partial l_f} = F' - w_f - v_f' = 0 \qquad (5)$$

Although the equations in (5) yield the same equilibrium as those in (3), they enable us to determine the optimal values of l_m and l_f.
Male and female average wages are:

$$\begin{aligned} w_m &= F' - v_m' \\ w_f &= F' - v_f' \end{aligned} \qquad (6)$$

Therefore, there is discrimination against women in terms of average wages if and only if there is more aversion to women. Furthermore, the wage gap between equally productive men and women, given by $w_m - w_f = v_f' - v_m'$, is less than the wage gap between groups 1 and 2, given by $w_1 - w_2$. This is because w_m and w_f are convex combinations of w_1 and w_2.
From (5), we derive the following probabilities:

$$\begin{aligned} 0 \le p_m &= \frac{F' - w_2 - v_m'}{w_1 - w_2} \le 1 \\ 0 \le p_f &= \frac{w_1 - F' + v_f'}{w_1 - w_2} \le 1 \end{aligned} \qquad (7)$$

Substituting $w_1 = F'$ and $w_1 - w_2 = v'$ (from (3)) into (7) yields $v_m' = v'(1 - p_m)$ and $v_f' = v' p_f$.
Given the above results, the following are equivalent:

$$\begin{aligned} w_m > w_f &\equiv v_m' < v_f' \equiv 1 - p_m < p_f \\ w_m = w_f &\equiv v_m' = v_f' \equiv 1 - p_m = p_f \\ w_m < w_f &\equiv v_m' > v_f' \equiv 1 - p_m > p_f \end{aligned} \qquad (8)$$

The interpretation of this result is intuitive: women are discriminated against in terms of average wages and probabilities if and only if there is greater aversion to women. The same applies to male workers if and only if there is larger aversion to men.
Even when there is no bias against women ($v_f' = v_m'$), there could be discrimination (but not based on gender); i.e., it could be that $w_1 > w_2$ although $w_m = w_f$, $p_m = 1 - p_f$ and $p_f = 1 - p_m$.

A particular case: the deterministic approach
If all relevant variables on which discrimination is based are observable and known, our model without prosecution is equivalent to the so-called deterministic model, which is in turn equivalent to the neoclassical models of Becker (1957) and Arrow (1973). That is, the deterministic model is an extreme case of our probabilistic model.

Suppose that gender is the only relevant variable for discrimination. In this case, group 1 is composed entirely of men and group 2 includes only women. Thus, $l_1 = l_m$, $l_2 = l_f$, $w_1 = w_m$, $w_2 = w_f$ and $p_m = p_f = 1$. Furthermore, the only argument in the aversion function v is the total number of women, l_f. Hence, v'_m is zero and $v'_f = v'$. From (3) and using w_m for w_1, w_f for w_2 and v'_f for v', we obtain $w_m - w_f = v'_f > 0$.

The nondeterministic model's equilibrium, given by (7), implies $p_m = p_f = 1$, and $w_m - w_f = v'_f > 0$, as in the deterministic model.

Therefore, the nondeterministic approach is a more general framework for understanding discrimination because one need not assume that all women are discriminated against.

2.2 Prosecutions against gender discrimination

Now we consider a government that prosecutes against gender discrimination. Employer preferences for female and male workers do not change when there is prosecution against discrimination. However, when maximizing expected utility, employers consider not only the disutility of hiring group 2 workers, but also the expected penalty to be paid. This affects the optimal numbers of female and male workers in both groups and reduces the gender wage gap.

Formally, if the government tries to reduce gender discrimination, the penalties imposed should depend on the bias against female workers, $v'_f - v'_m$. However, because these parameters are not directly observable, they cannot be used by the authorities. The government could adopt a second-best antidiscrimination policy based on the participation of female workers and salary differentials.

If the probability of inspection is q, the expected penalty imposed on employers who discriminate is:

$$E[h] = q \cdot h\left(\delta - \frac{l_f}{L}\right) \tag{9}$$

where $h(\cdot)$ is the penalty function and δ is the female participation rate in the labor market.[2, 3] Moreover, the total labor force in the firm, L, equals the sum of male and female workers, $L = l_m + l_f$. The penalty depends on the participation of female workers in the company relative to δ.[4] Gender discrimination can be practiced not only by paying different salaries, but also by hiring fewer women. Consequently, the government imposes no penalty (sets $h = 0$) if and only if average salaries are equal ($w_m = w_f$) and the firm's female participation rate is at the reference level $\left(\frac{l_f}{L} = \delta\right)$. In

---

[2] The parameter δ is a benchmark and can be defined as any social optimum in terms of female labor participation.

[3] There exists a maximum penalty $\tilde{h}$ that reflects the impossibility of some penalties: either they could be excessive from an ethical point of view or they could be uncollectible.

[4] The penalty should also depend on the wage gap. However, firms are wage takers, so w_1 and w_2 are fixed and the gender wage gap depends only on p_m and p_f.

particular, we assume that $h_l^{'} = \dfrac{\partial h}{\partial\left(\delta - \dfrac{l_f}{L}\right)} > 0$. Note that this derivative guarantees that the penalty increases as the gender wage gap increases.
The employer's expected utility is represented by the following function:[5]

$$EU = \pi(l_m, l_f, K) - v(l_m, l_f) - q \cdot h\left(\delta - \frac{l_f}{l_m + l_f}\right) \tag{10}$$

Under the assumption that labor productivity is the same for male and female workers, the first-order conditions are:

$$\begin{aligned} \frac{\partial EU}{\partial l_m} &= F^{'} - w_m^* - v_m^{'} - q\, h_l^{'} \frac{l_f}{L^2} = 0 \\ \frac{\partial EU}{\partial l_f} &= F^{'} - w_f^* - v_f^{'} + q\, h_l^{'} \frac{l_m}{L^2} = 0 \end{aligned} \tag{11}$$

where w_m^* and w_f^* are the corresponding equilibrium wages when there is prosecution. These wages are:

$$\begin{aligned} w_m^* &= F^{'} - v_m^{'} - q\, h_l^{'} \frac{l_f}{L^2} \\ w_f^* &= F^{'} - v_f^{'} + q\, h_l^{'} \frac{l_m}{L^2} \end{aligned} \tag{12}$$

Prosecution raises the average wage of women and lowers the average wage of men. The wage gap between male and female workers, $w_m^* - w_f^* = \left(v_f^{'} - v_m^{'}\right) - \dfrac{q \cdot h_l^{'}}{L}$, is lower than that when there is no prosecution against discrimination. This gap may even be reversed if expected penalty costs are sufficiently high.
The equations in (11), together with the definitions $w_m \equiv p_m w_1 + (1 - p_m) w_2$ and $w_f \equiv (1 - p_f) w_1 + p_f w_2$, allow us to obtain the following probability parameters at the equilibrium:

$$0 \le p_m^* = \frac{F^{'} - w_2 - v_m^{'}}{w_1 - w_2} - \frac{q \cdot h_l^{'} \dfrac{l_f}{L^2}}{w_1 - w_2} \le 1$$

[5] We ensure that the expected utility function is strictly concave by assuming that the penalty function $h(\cdot)$ is convex; i.e., $h(\cdot)$ has a positive semidefinite Hessian matrix.

$$0 \le p_f^* = \frac{w_1 - F' + v_f'}{w_1 - w_2} - \frac{q \cdot h_l' \frac{l_m}{L^2}}{w_1 - w_2} \le 1 \qquad (13)$$

Comparing (7) and (13) reveals that prosecution against discrimination lowers the probability of a male worker being in group 1, but also lowers the probability that a female worker is in group 2. Furthermore, under prosecution, the following are equivalent:

$$\begin{aligned}
&w_m^* > w_f^* \equiv v_f' - \frac{q \cdot h_l'}{L} > v_m' \equiv 1 - p_m^* < p_f^* \\
&w_m^* = w_f^* \equiv v_f' - \frac{q \cdot h_l'}{L} = v_m' \equiv 1 - p_m^* = p_f^* \\
&w_m^* < w_f^* \equiv v_f' - \frac{q \cdot h_l'}{L} < v_m' \equiv 1 - p_m^* > p_f^* \qquad (14)
\end{aligned}$$

The left-hand sides of the expressions in (14) come from (12). The right-hand sides of (14) derive from (13). The interpretation of (14) is that a necessary and sufficient condition for discrimination against women (over wages and probabilities) is that employers are more averse to hiring women, net of the expected penalty costs, than they are to hiring men.

It is beyond the scope of this paper to determine the government's optimal antidiscrimination policy. However, the optimal policy for a government that prosecutes against wage discrimination would minimize gender discrimination subject to the net resources available for prosecution. That is, the probability of inspection, q, and the penalty function, h, would depend on how averse the government is to gender discrimination and on the net costs of prosecution (information costs, inspection costs and administrative costs net of fine revenue).

3. Conclusions

In neoclassical models, workers are classified a priori into discrimination groups. In this paper, we developed a probabilistic model of wage discrimination in which workers need not be classified a priori. Becker's model is an extreme case of our model. Another implication of our model is that the traditional approach to measuring discrimination (the Oaxaca–Blinder approach) must be modified to take into account this probabilistic framework.

References

Arrow, Kenneth J., “The Theory of Discrimination”. In Orley Ashenfelter and Albert Rees, eds. *Discrimination in Labor Markets.* Princeton, N.J.: Princeton University Press, 1973.

Becker, Gary S., *The Economics of Discrimination*. Chicago: Chicago University Press, 1957.

Blinder, Alan S., “Wage Discrimination: Reduced Form and Structural Estimates”, *Journal of Human Resources*, 1973, 8 (4), pp. 436–455.

Oaxaca, Ronald L., “Male–Female Wage Differentials in Urban Labor Markets”, *International Economic Review*, 1973, 14 (3), pp. 693–709.

Prieto, J., Rodríguez, J.G. and R. Salas, “On the Measurement of Illegal Wage Discrimination: the Michael Jordan Paradox”, ECINEQ Working Paper 2006–38.

Figure 1. Distribution of workers by characteristics.