Working Paper Series

On The Measurement Of Illegal Wage Discrimination: The Michael Jordan Paradox

Juan Prieto Rodríguez
Juan Gabriel Rodríguez
Rafael Salas

ECINEQ WP 2006 – 38

ECINEQ 2006-38
April 2006

www.ecineq.org

On The Measurement Of Illegal Wage Discrimination: The Michael Jordan Paradox*

Juan Prieto Rodríguez†
Universidad de Oviedo
and

Juan Gabriel Rodríguez‡
Universidad Rey Juan Carlos

Rafael Salas§
Universidad Complutense de Madrid

Abstract

Standard wage discrimination models assume that independent observers are able to distinguish a priori which workers are suffering from discrimination. However, this assumption may be inadequate when severe penalties can be imposed on discriminatory employers. Antidiscrimination laws will induce firms to behave in such a way that independent observers (for instance, lawyers, economists) cannot easily detect discriminatory practices. This problem can be solved by estimating the discriminatory wage gap using finite mixture or latent class models because these procedures do not require the a priori classification of workers. In fact, the standard discrimination model can be seen as a particular case of our method when the probabilities of belonging to a group are fixed (to one or zero). We estimate discrimination coefficients for Germany and United Kingdom using the European Community Household Panel (ECHP). We obtain unambiguous higher discrimination in Germany for a wide set of measures.

Keywords: discrimination; wages; latent class model; finite mixture models
JEL Classification: J60; J70

*This research has benefited from the Spanish Ministry of Science and Technology Projects SEC2003-08397 and MEC-04-SEJ-04065. The usual disclaimer applies.
† Universidad de Oviedo, Avda. del Cristo s/n, 33071 Oviedo, Tel: +34 985 103768. jprietor@uniovi.es
‡ Universidad Rey Juan Carlos, Campus de Vicálvaro, 28032 Madrid, Tel: +34 91 4887948. juangabriel.rodriguez@urjc.es
§ Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, Tel: +34 91 3942512. r.salas@ccee.ucm.es

1. Introduction

Was Michael Jordan, the famous basketball African-American player, discriminated against? Standard models answer this question by first assuming that he belonged to a group being discriminated against. But if this is the *a priori* classification, a paradox arises: he is so rich that we think the colour of his skin does not guarantee that he actually belonged to the group being discriminated against. In fact, Michael Jordan received a huge salary during almost all his career.[2] However, he belongs to the group being discriminated against, according to mainstream discrimination economics. In this paper, we offer a different answer by making certain estimations assuming that the data will reveal the probability of his being discriminated against. Under this proposal, the conclusion may be quite different. Jordan would have some probability of belonging to the group that is discriminated against or he might not have belonged to such a group. The possibility of not belonging to the group being discriminated against is eliminated in the standard wage discrimination analysis.

Tracing back to Becker (1957), the topic of discrimination in the labour market is concerned with wages, selection, promotion and occupational differences between distinct groups of people. These differences arise as a result of belonging to a given group and are not related to any productivity characteristics of individuals, otherwise they could not be considered discriminatory. Moreover, when studying sex discrimination, we must take into account possible differences in human capital endowments for each sex, and consider the existence of different behaviours and aims with respect to the labour market that are related to gender and marital status.

[2] He was the sixth best paid athlete in the world in 2005 in spite of he was retired in 2003! See http://www.forbes.com/business/2006/03/22/woods-sharapova-nike_cx_lr_0322athletes_2.html.

The fact that differences in human capital are not the only factor explaining differences in productivity makes discrimination difficult to measure. Moreover, because laws in Western countries prohibit intentional, conscious or patent discrimination against members of a specific group, discrimination is not easily identifiable. In fact, we believe that it is almost impossible to observe wage discrimination directly and that it can be rather imprudent or risky to classify people *a priori* as members of a group that is discriminated against or a group that is favoured. However, the traditional way to measure discrimination relies explicitly on this assumption because it is based on the estimation of a pair of independent wage equations for each group—one for the group that is discriminated against and one for the group that is favoured—and these groups are defined using some observable characteristic of the workers, such as sex or race (Oaxaca, 1973 and Blinder, 1973). Hence, it seems that economists base their wage discrimination research on information that, as employers are well aware, can be very harmful for employers in legal terms, as the importance of litigation proves. Provided that there is an informational problem, it would be better from a methodological point of view not to assume an aprioristic classification of workers, but rather to attempt to estimate the membership of specific groups.

As a possible solution, we propose the estimation of a finite mixture model (see, for example, McLachlan and Peel, 2000), also known as a semiparametric heterogeneity model (Heckman and Singer, 1985) or a latent class model (Aitken and Rubin, 1985 and Greene, 2002).[3] This technique allows the individuals in the sample to be classified into different groups and enables us to evaluate the probability of a specific individual belonging to a particular group, even without enough information about his/her membership, or if these groups are not directly

[3] Millsap and Taylor (1996) used a latent variable rather than finite mixture model to analyze wage discrimination.

observed. That is, if wage discrimination exists in the labour market and observable worker characteristics cannot fully determine whether an individual is a member of the group that is discriminated against because of the legal threat that this implies to the employer, we can at least evaluate the probability of the individual being a member of such a group using finite mixture models. This task will be undertaken without any *a priori* classification, but instead by utilizing data according to probabilistic criteria. Therefore, instead of using the standard Oaxaca–Blinder approach to classify workers, we propose a method that evaluates workers on a case-by-case basis (maximizing the correspondent likelihood function) to reach a conclusion on the existence of wage discrimination and the membership of each worker in the group or class that is discriminated against.

Therefore, the proposed method is a generalization of standard approaches. The standard discrimination model (Oaxaca, 1973 and Blinder, 1973) is deterministic as it assumes fixed membership probabilities (either zero or one). However, the proposed approach is probabilistic and includes the standard model as a particular case, when the estimated probabilities converge for all individuals to either zero or one, depending on an observable characteristic, such as sex or race. Moreover, we can estimate the parameters that evaluate the effect of a certain variable on the wages of each class and the expected effect on a given worker's salary by taking into account the probability of him/her belonging to each group.

The theoretical framework of this paper is based on denominated statistical discrimination models (Phelps, 1972, and Aigner and Cain, 1977). These models view information problems in the labour market as the possible cause of the persistence of discrimination. They allow us to explain the existence of wage differences using the actions of profit-maximizing firms, which do not have monopsonistic power or preferences for any group of workers (Lundahl and Wadensjö, 1984).

If we assume that there is imperfect information in the labour market, workers will be reluctant to reveal any information that may negatively affect their wages or working conditions. In this context, determining the productivity of each worker may be too expensive for an employer, who will prefer to judge each worker using the average characteristics of the group to which the worker belongs. A particular worker may consider her/himself discriminated against if her/his labour productivity is closer to another group's average than to that of her/his own group yet s/he is paid a similar wage to the members of the group to which s/he belongs. Such a case can be called individual discrimination. Moreover, women in the workforce may expect some individual discrimination, which could cause them to behave differently before entering the labour market. Such differences may affect average productivity and labour behaviour. Therefore, the present behavioural differences could help to maintain wage differences between sexes in the long run. Models of statistical discrimination deal with the signalling role of different groups of workers, which demonstrate to the employer the productivity and the type of behaviour that could be expected from the group. In this context, not all women or all men will have the same relative position in the labour market: some of them will be better off (worse off) compared with the rest of their group and with other groups. Hence, we think that any empirical approach should be able to control for this individual heterogeneity without assuming that all the members of a particular group are discriminated against. These two conditions can be achieved using finite mixture models.

In the empirical exercise, we estimate the degree of discrimination in the United Kingdom (UK) and Germany using the European Community Household Panel (ECHP). The discrimination coefficients for the standard approach and for our proposal are obtained and compared. Furthermore, we also apply latent class model methodology to the distributional discrimination approach in Jenkins (1994) and del Rio *et al.* (2006). We detect an overestimation of discrimination under the standard approach, less gender discrimination in the UK than in Germany, and an increase and decrease in gender discrimination in Germany and the UK during the period 1994–2001, respectively. Moreover, these results are obtained from the dominance criterion based on the discrimination curves[4] so the results are valid for all indices of discrimination consistent with this criterion: in this sense they are robust.

The paper is organized as follows. Section 2 provides a brief review of the standard Oaxaca–Blinder model. The finite mixture model is examined in Section 3, and Section 4 presents the main empirical results. The final section includes some concluding remarks.

2. The traditional empirical framework

The traditional approach (see Oaxaca, 1973 and Blinder, 1973) seeks to equate discrimination with all wage differentials between groups that cannot be explained by differences in individuals' productivity characteristics. However, the groups are defined apurioristically, and each worker is assigned to a group before the wage equations are estimated. Hence, when dealing with gender discrimination, we will estimate a different wage equation for each gender:

$$\mathrm{Ln}\, W_{fi} = \beta_f^{'} \mathrm{X}_{\mathrm{fi}} + \varepsilon_{fi} \quad (1)$$

$$\mathrm{Ln}\, W_{mi} = \beta_m^{'} \mathrm{X}_{\mathrm{mi}} + \varepsilon_{mi}\,, \quad (2)$$

[4] The Discrimination Curve is the result of applying the Inverse Generalized Lorenz Curve to the distribution of wage gaps (see del Rio *et al.*, 2006).

where W_i is the ith individual's earnings per hour, X_i is the vector of human capital variables, β_{m} and β_{f} are the vectors of returns to these characteristics, ε_i is the error term and the subscripts m and f represent the worker's gender; male and female, respectively.

The estimated female individual wage gap (in logarithms) is:

$$\mathrm{Ln}\,\hat{W}_{fi}^{cf} - \mathrm{Ln}\,\hat{W}_{fi} = \left(\hat{\beta}_m^{'} - \hat{\beta}_f^{'}\right)\mathrm{X}_{fi} \tag{3}$$

where $\hat{W}_{fi}^{cf}$ is the ith individual's estimated wage if her attributes were remunerated as if she were a male, and $\hat{\beta}_m$ and $\hat{\beta}_f$ are the male and female estimated wage coefficient vectors, respectively. The estimated individual wage gap in (3) reflects the discrimination experienced by a female i.

The aggregated gap between men's and women's wages has traditionally been evaluated in the mean distribution of the characteristics. This aggregated gap can be decomposed into a discriminatory and a nondiscriminatory wage differential, as in the following expression:

$$\overline{\mathrm{Ln}\,\hat{W}_m} - \overline{\mathrm{Ln}\,\hat{W}_f} = \hat{\beta}_m^{'}\ \overline{\mathrm{X}}_{\mathrm{m}} - \hat{\beta}_f^{'}\ \overline{\mathrm{X}}_{\mathrm{f}} = \hat{\beta}_m^{'}\ (\overline{\mathrm{X}}_{\mathrm{m}} - \overline{\mathrm{X}}_{\mathrm{f}}) + (\hat{\beta}_m^{'} - \hat{\beta}_f^{'})\overline{\mathrm{X}}_{\mathrm{f}}\,. \tag{4}$$

In this equation, the term $\hat{\beta}_m^{'}\ (\overline{\mathrm{X}}_{\mathrm{m}} - \overline{\mathrm{X}}_{\mathrm{f}})$ shows the wage differences that would persist even if salary discrimination disappeared. These differences arise from the different average human capital endowments among men and women. The discriminatory differences are evaluated as $\left(\hat{\beta}_m^{'} - \hat{\beta}_f^{'}\right)\overline{X_f}$, which is the mean value of equation (3).

Recently, different econometric techniques have been proposed in the literature to take into consideration distributional aspects of wage discrimination. Three main approaches can be distinguished. In the first approach, the use of quantile regressions has been proposed to increase the number of points in the earnings distribution at which discrimination is evaluated (see, for example, Newell and Reilly, 2001, Albrecht, Björklund and Vroman, 2003 and Gardeazábal and Ugidos, 2005). A second approach is seen in the proposals to estimate counterfactual wage distribution functions to quantify differentials with the observed earnings distribution throughout the whole wage range. See, for example, DiNardo *et al.* (1996) and Machado and Mata (2004). Finally, Jenkins (1994) and del Rio *et al.* (2006) have proposed the use of theoretical results in poverty and deprivation literatures to focus the discrimination analysis on the entire distribution of individual wage gaps.

All these approaches distinguish *a priori* which workers suffer discrimination. However, this *a priori* classification of observations does not reflect the inherent uncertainty of wage discrimination. In the following section, we propose a method to overcome this problem.

3. A finite mixture model to measure wage discrimination

Finite mixture models, or latent class analysis, are a statistical method for finding subtypes of related cases (latent classes) from multivariate categorical data.[5] In our case, they can be used to find different kinds of workers in terms of their wage structure, allowing us to find different labour market segments and classify workers into these segments.

It is worth noting that methods that use predefined groups of workers to estimate wage equations for each group do not use information contained in one class to estimate wage equations for workers who belong to other classes. However, in most empirical applications, this interclass information may be quite important because individuals belonging to different classes often share common features. If this kind of information were not exploited, it is

[5] See Cameron and Trivedi (2005) or Beard *et al.* (1991) for a survey of latent class models. Other recent examples can be found in Clark *et al*. (2005) and Greene (2005).

possible that the estimation method would be inefficient.[6] We use a procedure that combines standard wage equations and latent class models in order to exploit the information contained in the data with greater efficiency. In our model, it is unnecessary to know beforehand which group produced an observation because both the individual's wages and the probability of membership of a particular group are estimated *simultaneously*. Individuals are probabilistically separated into several classes, and a wage equation is estimated for each class. As each observation may have a nonzero probability of belonging to any class, all the observations in the sample are used to estimate all the wage equations.[7] Moreover, the proposed methodology allows the sample to be split into groups even when sample-separating information is not available. In this case, the finite mixture model uses the goodness of fit of each estimated function as additional information to identify groups of individuals.

In latent class models, individuals are assumed to belong to one of J classes $j = 1, ..., J$ where class membership is unknown. Assuming that the wage structures for each class follow standard Mincer's equations, the standardized normal density function for individual i assuming that s/he belongs to class j can be written as:

$$f_{ij}(\beta_j) = \phi\left(\frac{\ln W_i - \beta_j' X_i}{\sigma_j}\right), \tag{5}$$

where X_i is the vector of human capital variables, including educational level, tenure and labour experience (see Table A1 for a definition of these variables) and σ_j is the standard deviation of the corresponding group j. As sex is not included in this vector, wages vary

[6] Analysis cluster classifies observations according to *a priori* sample separation information. Therefore, it does not use information contained in one class to estimate the other class's wage equation.
[7] In the standard procedure, we are *implicitly* restricting the cross-class probabilities to zero and the own probabilities to one. This precludes using observations from other classes to estimate a particular class's probability.

within a specific group because of human capital differences, but expected wages for men and women will be equal once controlled by human capital given group membership.

The unconditional response probabilities become finite mixtures. The unconditional likelihood for individual *i* is obtained as the weighted sum of her/his *j*-class likelihood functions, where the weights are the probabilities of class membership. That is,

$$f_i(\beta,\delta) = \sum_{j=1}^{J} f_{ij}(\beta_j)\cdot P_{ij}(\delta_j) \quad , \quad 0 \le P_{ij} \le 1 \quad , \quad \Sigma_j P_{ij} = 1 , \tag{6}$$

where $\beta = (\beta_1, \ldots, \beta_J)$, $\delta = (\delta_1, \ldots, \delta_J)$ and $P_{ij}(\delta_j)$ represent the probability of individual *i* being a member of class *j*. These class probabilities are parameterized as a multinomial logit model:

$$P_{ij}(\delta_j) = \frac{\exp(\delta_j' q_i)}{\sum_{j=1}^{J} \exp(\delta_j' q_i)} \quad , \quad j = 1,\ldots,J \quad , \quad \delta_J = 0 , \tag{7}$$

where q_i is a vector of individual-specific variables. In order to estimate sex discrimination, we have included workers' gender in this vector. Therefore, discriminatory wage differences between men and women arise from the differences in their probabilities of being members of each class of workers, not because men and women are paid differently within each class. That is, wage discrimination will arise from the higher probability that women will be members of those groups with lower human capital returns, but these groups will include men as well. Thus, wages will reflect the uncertainty that we have about the true partitioning of the sample, and we believe that this uncertainty is central to the problem that we are analysing.

The overall log-likelihood function resulting from (6) is a continuous function of the vectors of parameters β and δ, and can be written as:

$$\log f(\beta,\delta)=\sum_{i=1}^{N}\log f_i(\beta,\delta)=\sum_{i=1}^{N}\log\left\{\sum_{j=1}^{J} f_{ij}(\beta_j)\cdot P_{ij}(\delta_j)\right\}. \tag{8}$$

Under the maintained assumptions, maximum likelihood techniques will give asymptotically efficient estimates of all the parameters. The number of classes, *J*, is taken as given. In our case, there will be two classes, *J*=2, the discriminated and the nondiscriminated classes.

The estimated parameters can be used to compute the conditional posterior class probabilities as follows:

$$P(j\,|\,i)=\frac{f_{ij}(\hat{\beta}_j)\cdot P_{ij}(\hat{\delta}_j)}{\sum_{j=1}^{J} f_{ij}(\hat{\beta}_j)\cdot P_{ij}(\hat{\delta}_j)}. \tag{9}$$

This expression shows that the posterior class probabilities depend not only on the estimated δ parameters, but also on the vector β, i.e., the parameters from the wage equations and the observed wages.

Once we have efficiently estimated the returns parameters according to the maximum likelihood function and the prior and posterior class probabilities, we need only know individual's returns to evaluate wage discrimination. There are two ways to measure individual's returns. First, we can evaluate the marginal effects as:

$$\hat{\beta}_i=\sum_{j=1}^{J}\hat{\beta}_j\cdot P_{ij}(\hat{\delta}_j), \tag{10}$$

where P_{ij} is individual *i*'s prior probability of belonging to class *j*, defined in (7), and $\hat{\beta}_j$ is the estimated marginal effect using the wage equation for class *j*. Note that here we take into account all the possible returns related to the different classes—$\hat{\beta}_j$—weighted by their prior class probabilities—P_{ij}.

Alternatively, we can examine the posterior class probabilities for each individual following equation (9) and use them as weights to evaluate individuals' coefficients:

$$\hat{\beta}_i = \sum_{j=1}^{J} \hat{\beta}_j \cdot P(j \mid i) . \tag{11}$$

Obviously, the results obtained by both methods are different.

To evaluate the individual female wage discrimination coefficient (DC_{fi}), we need a counterfactual wage for each female worker to calculate the individual's discrimination gap. This counterfactual wage depends on the human capital endowments and the probabilities of belonging to each class if the worker were man rather than woman. In order to do this, we estimate counterfactual posterior probabilities, $P_{j|i}^{cf}$, using equation (9) by changing the value of the dummy variable for sex included in vector q when computing the prior probabilities following equation (7). Then, we compute the following expression for all the women in the sample:

$$\begin{aligned} DC_{fi} &= \left(\hat{\beta}_{fi}^{cf\,'} - \hat{\beta}_{fi}^{'}\right) X_{fi} \\ &= \sum_{j=1}^{J} P_{j|i}^{cf} \cdot \left(\hat{\beta}_j{}' X_{fi}\right) - \sum_{j=1}^{J} P(j \mid i) \cdot \left(\hat{\beta}_j{}' X_{fi}\right) \\ &= \sum_{j=1}^{J} \left(P_{j|i}^{cf} - P(j \mid i)\right) \cdot \left(\hat{\beta}_j{}' X_{fi}\right). \end{aligned} \tag{12}$$

For each woman, this expression evaluates the difference between her actual expected wage, $\sum_{j=1}^{J} P(j \mid i) \cdot \left(\hat{\beta}_j{}' X_{fi}\right)$ and the wage she would receive if discrimination did not exist, that is, the wage she would receive if she had the same membership probabilities as a man, $\sum_{j=1}^{J} P_{j|i}^{cf} \cdot \left(\hat{\beta}_j{}' X_{fi}\right)$.

In the gender discrimination problem, there are two possible classes: not discriminated against workers (group 1) and discriminated against workers (group 2); so the wage discrimination coefficient can be written as:

$$\begin{aligned} DC_{fi} &= \left(P_{1|i}^{cf} - P(1\,|\,i)\right)\cdot\left(\hat{\beta}_1{}' X_{fi}\right) + \left(P_{2|i}^{cf} - P(2\,|\,i)\right)\cdot\left(\hat{\beta}_2{}' X_{fi}\right) \\ &= \left(P_{1|i}^{cf} - P(1\,|\,i)\right)\cdot\left(\hat{\beta}_1{}' X_{fi}\right) + \left(\left(1 - P_{1|i}^{cf}\right) - \left(1 - P(1\,|\,i)\right)\right)\cdot\left(\hat{\beta}_2{}' X_{fi}\right) \\ &= \left(P_{1|i}^{cf} - P(1\,|\,i)\right)\cdot\left(\hat{\beta}_1{}' - \hat{\beta}_2{}'\right) X_{fi}. \end{aligned} \tag{13}$$

Note the similarities between this expression and equation (3). In fact, if we consider (in line with the standard literature on gender discrimination) that men are the group that suffers no discrimination (J =1) and women are the group discriminated against, then we will have $P_{1|i} = 0$ and $P_{1|i}^{cf} = 1$ for women. Moreover, the female individual wage discrimination coefficient DC_{fi} becomes the female individual wage gap in (3) and the standard Oaxaca–Blinder approach is obtained as a particular case of our proposal.

The female wage discrimination coefficient is $DC_f = \frac{1}{N}\sum_{i=1}^{N} DC_{fi}$, that is, the women's mean discrimination gap. This expression evaluated for the women's subsample is:

$$\begin{aligned} DC_f &= \frac{1}{N}\sum_{i=1}^{N}\left[\sum_{j=1}^{J} P_{j|i}^{cf}\cdot\left(\hat{\beta}_j{}' X_{fi}\right) - \sum_{j=1}^{J} P(j\,|\,i)\cdot\left(\hat{\beta}_j{}' X_{fi}\right)\right] \\ &= \frac{1}{N}\sum_{i=1}^{N}\left[\sum_{j=1}^{J} P_{j|i}^{cf}\cdot\hat{\beta}_j' - \sum_{j=1}^{J} P(j\,|\,i)\hat{\beta}_j'\right] X_{fi} \\ &= \left(\overline{\hat{\beta}_m}' - \overline{\hat{\beta}_f}'\right)\overline{X_f}. \end{aligned} \tag{14}$$

where $\overline{\hat{\beta}_m} = \sum_{j=1}^{J} P_{j|i}^{cf} \hat{\beta}_j$ and $\overline{\hat{\beta}_f} = \sum_{j=1}^{J} P(j \mid i) \hat{\beta}_j$ denote the average vector of coefficients for men and women, respectively, weighted by the posterior probabilities. We used this definition of the wage discrimination coefficient to enable comparisons with the standard Oaxaca–Blinder approach. This expression is equivalent to the discrimination coefficient defined in Section 2 (the second term on the right-hand side of equation 4). The only difference is that membership in a specific group has a random component, as we assume that groups of workers who are discriminated against cannot be defined *a priori*.

Two final comments must be done before presenting the empirical results. First, the above methodology can be similarly employed to any kind of wage discrimination. We have focused the analysis on gender discrimination, but other kinds of discrimination could be considered: race, religion, nationality, etc.

Second, we have applied the latent class methodology to estimate wage equations and we have complemented it with the distributional approach in Jenkins (1994) and del Rio *et al.* (2006).

Let us define $h_i(\mathrm{DC}_{fi}) = \max\,\{\mathrm{DC}_{fi}, 0\}\ \forall\ i = 1,\ldots,\mathrm{N}$ and h as the vector of h_i ranked from a higher to a lower level of discrimination. The discrimination curve is the sum of the first p per cent ($0 \leq p \leq 100$) of h values divided by the total number of females (see del Rio *et al.*, 2006). Comparison among different discrimination curves is a dominance criterion, which is incomplete and it can be characterized by a set of axioms: continuity, focus, monotonicity, symmetry, progressive transfers and replication invariance. However, complete indices of discrimination consistent with the above axioms can be used if the estimated discrimination curves cross. We will see below that the analysis of the discrimination indices consistent with the above axioms is not necessary in our case as the estimated discrimination curves for UK and Germany do not cross.

4. Results

In this section, we estimate the finite mixture model and the standard Oaxaca–Blinder (1973) model to evaluate sex discrimination in Germany and the United Kingdom. Moreover, we apply the proposed methodology to the distributional approach in del Rio *et al.* (2006). The database used in the estimations is the European Community Household Panel (ECHP). This database contains data on individuals in 12 European countries, with data for eight years available (1994–2001). The information is homogenous across the countries because the elaboration process of the survey was coordinated by EUROSTAT, although the size of samples varies across countries and years. Given the nature of our study, we have used the eight-year incomplete panel for both countries. Moreover, we have singled out information on personal characteristics (age, work experience, length of time in the firm, and education), wages and labour status for individuals in the samples.

As mentioned above, we have estimated Mincer's wage equations. Hence, the natural logarithm of the hourly wage in constant terms, **Ln(W)**, is the dependent variable of the wage equations. The independent variables are those factors that economic theory suggests as possible sources of compensating wage differences for differences in human capital: labour experience, tenure and education. Moreover, yearly dummy variables were included to control for cyclical effects.

Education is included by levels using two dummies (**EDUC1** and **EDUC2**). Potential experience (**POTEXP**) and tenure (**TENURE**) are included in the model to account for *on-*

the-job-training. A quadratic relationship between these variables and wages is expected. To represent these relationships, the wage equations include the potential experience and tenure squares (**POTEXP2** and **TENURE2**). Table 1 contains information on the average wage rates and human capital variables by country and sex. As expected, women have lower wage rates in both countries.

[Table 1]

Table 2 shows the results of the two-class finite mixture model. For both countries, being a woman increases the probability of belonging to the second class, that is, of being rewarded on the basis of the wage equation with the lowest intercept. In addition, for the UK, being a woman implies the lowest human capital returns. For Germany, human capital returns are higher for the second-class members. However, the important differences in both class intercepts compensate first-class members well for these lower human capital returns and they are better rewarded than second-class members. Thus, as expected, women have a higher probability of being in the least-rewarded labour market group, since they have a lower prior probability of being in the first group related to their negative estimated coefficient.

[Table 2]

POTEXP and **POTEXP 2** present the quadratic expected relationship with wages. For both groups in both countries, wages increase until approximately 24–25 years of experience have been gained. After these maximums, there are diminishing returns to experience. In addition, a quadratic relationship is detected for **TENURE** and **TENURE2,** but the maximum returns differ by country and class.

In order to estimate the Oaxaca–Blinder discrimination coefficients, we have estimated men's and women's wage equations for Germany and the UK. Results are displayed in Table 3.

[Table 3]

The estimated equations show a higher intercept for men than for women in both countries, with a greater difference between the two intercepts in the German case. Furthermore, similarly to the finite mixture estimations, educational returns are lower for German men than for German women. However, adding up the intercept and the education level estimated effects on wages, German men are better off than women, regardless of their educational attainments. On the other hand, in the UK, educational returns are higher for men than for women. It is remarkable that among the human capital variables, potential labour experience accounts for the main wage differences in both countries. All the estimates in table 2 and 3 are statistically significant at the 5% level of significance. As mentioned above, following Oaxaca (1973) and Blinder (1973), we have used the coefficients from the men's equation to predict the earnings that women would have earned if they had been treated as men. We then decompose the wage gap into the unexplained differential and the component that can be explained by characteristics.

Discrimination coefficients were calculated using these wage equations and the finite mixture models estimations. The results are presented in Table 4, which includes in the first row the total wage gap between men and women. The second and third rows of Table 4 present discrimination coefficients estimated using both approaches.

[Table 4]

It can be observed that the estimated coefficients are lower when workers are not classified *a priori*. It may be an overestimation problem if it is assumed that all women are discriminated against from the beginning. It should be pointed out, however, that individual values of the estimated discrimination coefficient from the finite mixture model are non-negative for all women, although some women have higher posterior probabilities of belonging to the privilege group than to the less paid group.

The discriminating wage difference is the most important component of the observed wage difference between men and women. However, its estimated magnitude varies depending on the approach: using the finite mixture models, it is around 60%, whereas using the Oaxaca–Blinder approach, it exceeds 90%. It is also apparent from Table 4 that the wage discrimination degree in Germany is larger than in the UK. This is confirmed by the graphical analysis below. It can be observed that British total wage gap is smaller than the Oaxaca-Blinder estimated discrimination coefficient that means that women should earn more than men if discrimination would be eliminated; i.e., characteristic included in the wage equations takes in average higher values for women than for men although these differences are quite small.

As mentioned above, we have applied the distributional approach in del Rio *et al.* (2006) to the finite mixture model results. We have estimated the discrimination curves for the UK and Germany in 1994 and 2001. Figures 1.A and 1.B show the discrimination curves for Germany and the UK in 2001 using the wage gaps according to the Oaxaca–Blinder model and the individual discrimination coefficients according to the posterior probabilities. We observe that the discrimination curve evaluated for the posterior probabilities dominates the Oaxaca–Blinder-based curve, due to the overestimation of discrimination with the latter model.

[Figures 1.A and 1.B]

Figure 2 shows the dominance of UK over Germany according to the discrimination curves, indicating greater discrimination in Germany according to any discrimination index consistent with these curves.

[Figure 2]

Finally, in Figure 3.A, despite having greater discrimination, it is shown that Germany showed an additional increase in discrimination in the period 1994–2001. This result contradicts the evolution given by the classical Oaxaca-Blinder model. In Figure 3.B it is shown that the UK reduces the discrimination in this period.

[Figures 3.A and 3.B]

5. Discussion and conclusions

If wage discrimination exists in the labour market, observable worker characteristics cannot determine whether an individual is a member of the group that is discriminated against because of the legal threat that this implies to the employer. However, the traditional Oaxaca–Blinder (1973) approach assumes that people can be classified *a priori* as members of a group that is discriminated against or a group that is favoured. We believe that this assumption can bias results of studies on wage discrimination.

To solve this problem we propose the use of finite mixture models. These allow us to classify the workers into different groups and estimate, simultaneously, the wage structure for each group. Moreover, they also permit evaluation of the probability of a specific individual belonging to a particular segment of the labour market and estimate her/his expected wage.

These models were estimated for Germany and the United Kingdom using the data of the European Community Household Panel. From the results of the estimated finite mixture models, we can assert that being a woman increases the probability of belonging to the labour group with the lowest wages. This result is the main factor explaining wage discrimination because wage differences arise from differences in human capital endowments, rather than from discrimination, within each group. Moreover, human capital variables (education, labour experience and tenure) have been shown to be very important in explaining wage differences between individuals.

Furthermore, the discriminating wage difference is the most important component of the observed wage difference between men and women. However, the estimated magnitude of this varies from one approach to another. As expected, Oaxaca–Blinder's (1973) approach gives us higher wage discrimination coefficients for both countries, that is, it seems to overestimate wage discrimination due to the aprioristic classification of workers. In any case, although the discrimination coefficient proposed in this research and based on finite mixture models reduces the evaluated differences in wages due to discrimination, this is a problem that remains far from solved since the estimated individual discrimination coefficient for all women, regardless of country, is positive.

Finally, inverse discrimination curves are deduced and we obtain unambiguous higher discrimination in Germany for a wide set of measures consistent with the inverse discrimination curves.

Beyond the scope of this paper, the proposed methodology allows us to analyze a relevant issue of political economy. As countries follow different active wage discrimination policies, the probabilities of women being discriminated against, differs. The larger is the active prosecution of wage discrimination, the lower is the probability of discrimination occurring. Thus, the greater the enforcement of antidiscrimination policies, the further away from one is the probability of women being discriminated against, and, therefore, the larger is the bias achieved by estimating the discrimination with the standard Oaxaca–Blinder methodology in comparison with the proposed procedure. In this context, the Oaxaca–Blinder approach is more appropriate in those countries where discrimination is more "permissive" (or even legal). This is an open line to be explored in future research. Another hypothesis to be tested is to what extent this bias is an indicator of the degree of active enforcement of antidiscrimination policies (in a general sense, the combination of laws, resources, etc.).

References

Aigner, D. J. and Cain, G. G. (1977) “Statistical Theories of Discrimination in Labor Markets”, *Industrial and Labor Relations Review*, vol 30, n 2, pp175-187.

Aitken, M. and D.B. Rubin (1985): “Estimation and hypothesis testing in finite mixture models”, *Journal of the Royal Statistical Society*, B, 47, 67-75.

Albrecht, J., A. Björklund and S. Vroman (2003): “Is there a glass ceiling in Sweden?”, *Journal of Labor Economics*, 21, 145-177.

Beard, T., S. Caudill and D. Grooper (1991): “Finite Mixture Estimation of Multiproduct Cost Functions.” *Review of Economics and Statistics,* vol. 73, 654-664.

Becker, G. (1957): *The Economics of Discrimination*. Chicago. Chicago University Press.

Blinder, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, vol. 8, nº 4, 436-455.

Cameron, A.C. and P.K. Trivedi (2005): Microeconometrics. Cambridge University Press, New York, USA.

Clark, Andrew; Etilé, Fabrice; Postel-Vinay, Fabien; Senik, Claudia; Van der Straeten, Karine (2005) ”Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries”, *Economic Journal*, Vol. 115, n 502, pp 118-132.

DiNardo, J., N.M. Fortin and T. Lemieux (1996): “Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach”, *Econometrica*, 64, 5, 1001-1044.

Del Rio, C., C. Gradín and O.Cantó (2006): “The measurement of gender wage discrimination: the distributional approach revisited”, ECINEQ working paper 2006-25.

Gardeazábal, P. and Ugidos, A. (2005): “Gender wage discrimination at quantiles”, *Journal of Population Economics*, 18, 165-179.

Greene, W. (2002): “Alternative Panel Data Estimators for Stochastic Frontier Models”, *Working Paper,* Department of Economics, Stern School of Business, NYU.

Greene, W. (2005): “Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model”, *Journal of Econometrics,* vol 126, pp 269-303.

Heckman, J.J. and B. Singer (1984): “A method for minimizing the impact of distributional assumptions in econometric models of duration data”, *Econometrica*, 52, 271-320.

Jenkins, S.P. (1994): “Earnings discrimination measurement: a distributional approach”, *Journal of Econometrics*, 61, 81-102.

Lundahl, M. and Wadensjö, E. (1984): *Unequal Treatment. A Study in the Neo-Classical Theory of Discrimination*, New York University Press, New York.

Machado, J.A.F. and J. Mata (2005): “Counterfactual decomposition of changes in wage distributions using quantile regression”, *Journal of Applied Econometrics*, 20, 4, 445-577.

Millsap, R. E. and Taylor, R. (1996) **“**Latent Variable Models in the Investigation of Salary Discrimination: Theory and Practice”, *Journal of Management*, vol 22, n 4, 653-673.

McLachlan, G. and Peel, D. (2000): *Finite mixture models*, New York, John Wiley & Son Ltd.

Newell, A. and Reilly, B. (2001): “The gender pay gap in the transition from communism: some empirical evidence”, *Economic Systems*, 25, 287-304.

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

Phelps, E. S. (1972): "The Statistical Theory of Racism and Sexism", *American Economic Review*, vol. 62, 659-661.

Table A1. Variable definitions

Name	*Definition*
Ln(W)	natural logarithm of the woman's hourly real wage
EDUC1	=1 if the individual has university studies; =0 otherwise
EDUC2	=1 if the individual has secondary school studies; =0 otherwise
POTEXP	potential experience (present age-age when started work)
POTEXP2	square of potential experience
TENURE	years of experience at the current firm
TENURE2	square of tenure

Table 1. Summary statistics

	GERMANY		UNITED KINGDOM	
	Mean	Std. Dev.	Mean	Std. Dev.
		Male's		
Ln(W/H)	2.0190	0.4820	2.0468	0.4920
EDUC1	0.2408	0.4276	0.5007	0.5000
EDUC2	0.5859	0.4926	0.1420	0.3490
TENURE	8.7976	6.7262	5.1892	5.1043
POTEXP	20.6286	11.3937	18.6785	12.8569
N observations	22880		15721	
N individuals	4621		3590	
		Women's		
Ln(W/H)	1.7536	0.4707	1.8667	0.4704
EDUC1	0.2111	0.4081	0.4314	0.4953
EDUC2	0.5936	0.4912	0.1425	0.3496
TENURE	7.2637	6.1056	4.6491	4.4976
POTEXP	19.8731	11.2277	19.4273	13.0059
N observations	17369		15200	
N individuals	3932		3614	

Table 2. Two latent class model of wages

	GERMANY		UNITED KINGDOM	
	Estimated coefficient	Standard error	Estimated coefficient	Standard error
Wage equation for latent class 1				
CONSTANT	1.53436	0.00887	1.57134	0.01087
EDUC1	0.40728	0.00329	0.28463	0.00373
EDUC2	0.10063	0.00307	0.14508	0.00595
TENURE	0.01950	0.00078	0.00952	0.00130
TENURE2	-0.00018	0.00003	-0.00045	0.00007
POTEXP	0.02860	0.00041	0.03106	0.00050
POTEXP2	-0.00058	0.00001	-0.00061	0.00001
σ	0.28186	0.00027	0.29830	0.00080
Wage equation for latent class 2				
CONSTANT	0.76528	0.01309	1.10080	0.00906
EDUC1	0.43196	0.00664	0.18208	0.00356
EDUC2	0.23043	0.00573	0.14438	0.00559
TENURE	0.03765	0.00143	0.02017	0.00119
TENURE2	-0.00093	0.00006	-0.00074	0.00006
POTEXP	0.03530	0.00066	0.02784	0.00045
POTEXP2	-0.00072	0.00002	-0.00058	0.00001
σ	0.38322	0.00075	0.30189	0.00081
Estimated prior probabilities				
CONSTANT	0.80300	0.03655	0.04640	0.04072
WOMAN	-1.29953	0.05330	-0.89363	0.05871
N observations	40249		30921	
N individuals	8553		7204	
Log-likelihood	-14477.92		-9850.344	

Table 3. Men and women wage equations

	GERMANY		UNITED KINGDOM	
	Estimated coefficient	Standard error	Estimated coefficient	Standard error
	Male's Wage equation			
CONSTANT	1.24468	0.01408	1.34001	0.01212
EDUC1	0.23249	0.01136	0.14568	0.00842
EDUC2	0.14348	0.00881	0.08646	0.00916
TENURE	0.01408	0.00134	0.00790	0.00172
TENURE2	-0.00022	0.00006	-0.00031	0.00010
POTEXP	0.04475	0.00127	0.03974	0.00115
POTEXP2	-0.00082	0.00003	-0.00076	0.00003
$\sigma[e_{it}]$	0.20442		0.21354	
$\sigma[u_i]$	0.41090		0.33097	
N observations	22880		15721	
N individuals	4621		3590	
Log-likelihood	6528.308		4010.758	
	Women's wage equation			
CONSTANT	1.10519	0.01537	1.27302	0.01212
EDUC1	0.33734	0.01432	0.14019	0.00810
EDUC2	0.20267	0.00997	0.09257	0.00842
TENURE	0.01837	0.00162	0.00880	0.00183
TENURE2	-0.00040	0.00008	-0.00037	0.00011
POTEXP	0.02977	0.00141	0.02903	0.00113
POTEXP2	-0.00062	0.00003	-0.00063	0.00003
$\sigma[e_{it}]$	0.21889		0.21144	
$\sigma[u_i]$	0.40486		0.31644	
N observations	17369		15200	
N individuals	3932		3614	
Log-likelihood	4043.424		4121.671	

Table 4. Wage differences decomposition

Model	GERMANY	UNITED KINGDOM
Total Wage Gap	26.542	18.009
Finite mixture model: $\left[\sum_{j=1}^{J} P_{j\mid i}^{cf}\hat{\beta}_j' - \sum_{j=1}^{J} P(j\mid i)\hat{\beta}_j'\right]\overline{X_f}$	11.066	6.439
Oaxaca-Blinder approach: $\left(\hat{\beta}_m' - \hat{\beta}_f'\right)\overline{X_f}$	24.423	18.151

Figure 1.A Wage discrimination in Germany (2001)

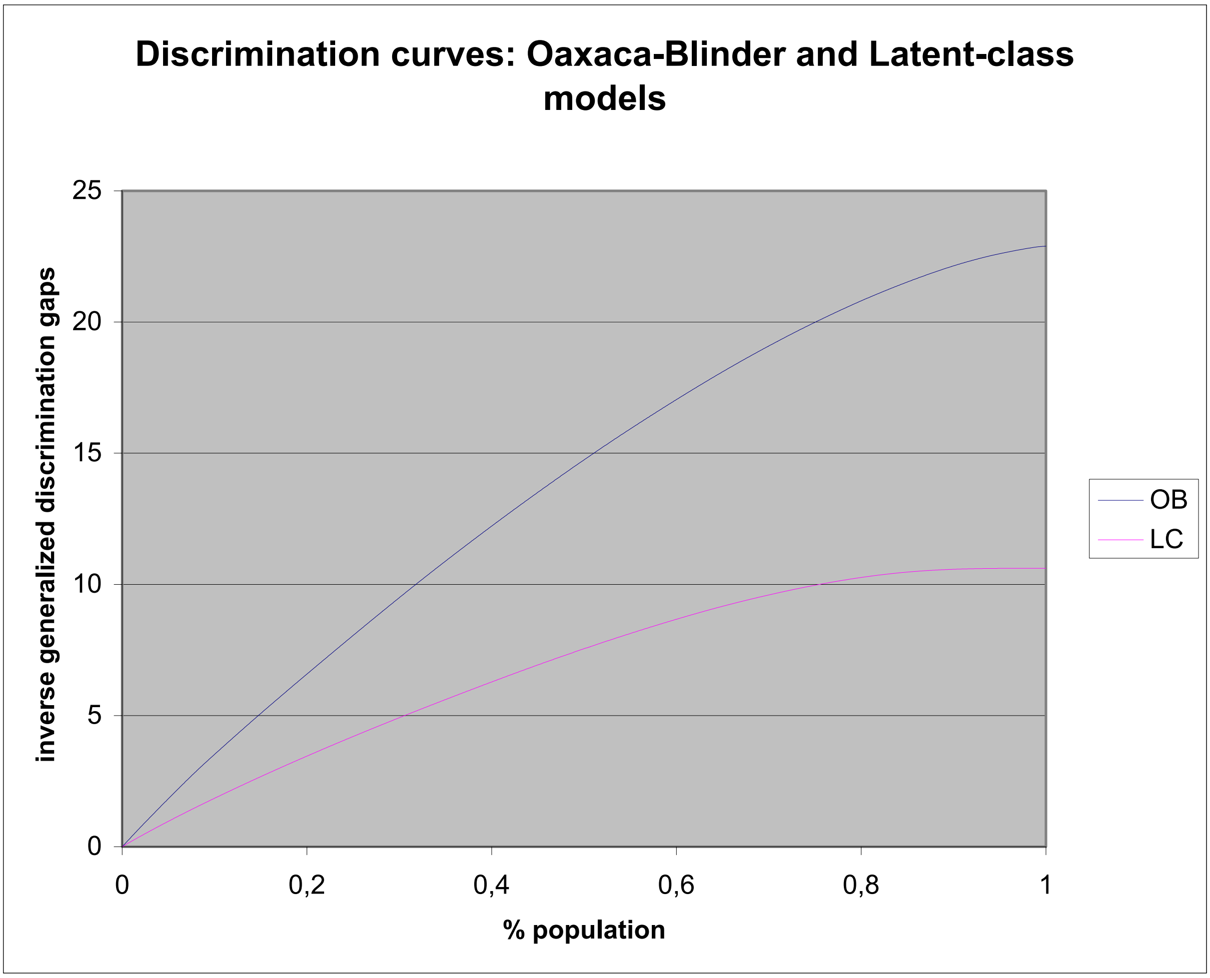

Figure 1.B Wage discrimination in the United Kingdom (2001)

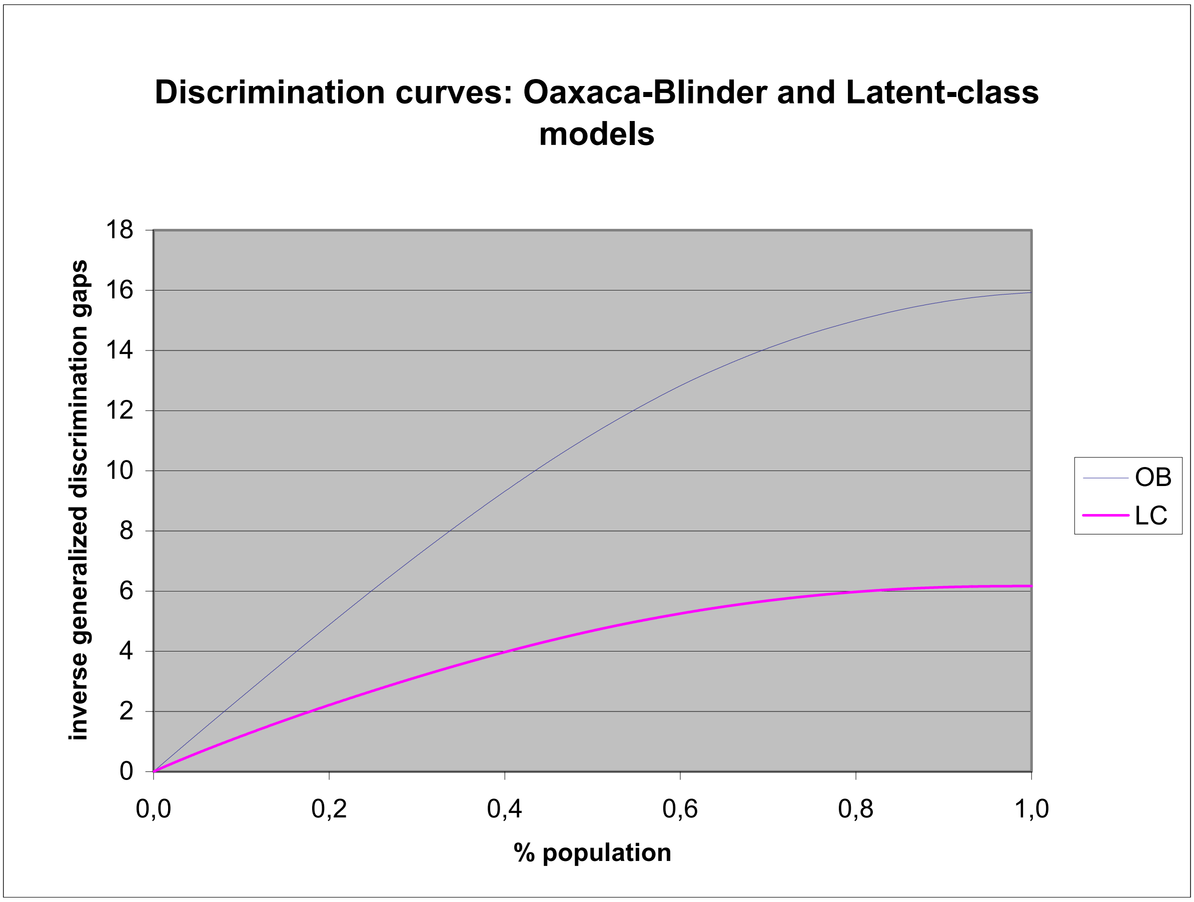

Figure 2. Wage discrimination in the UK and Germany (2001)

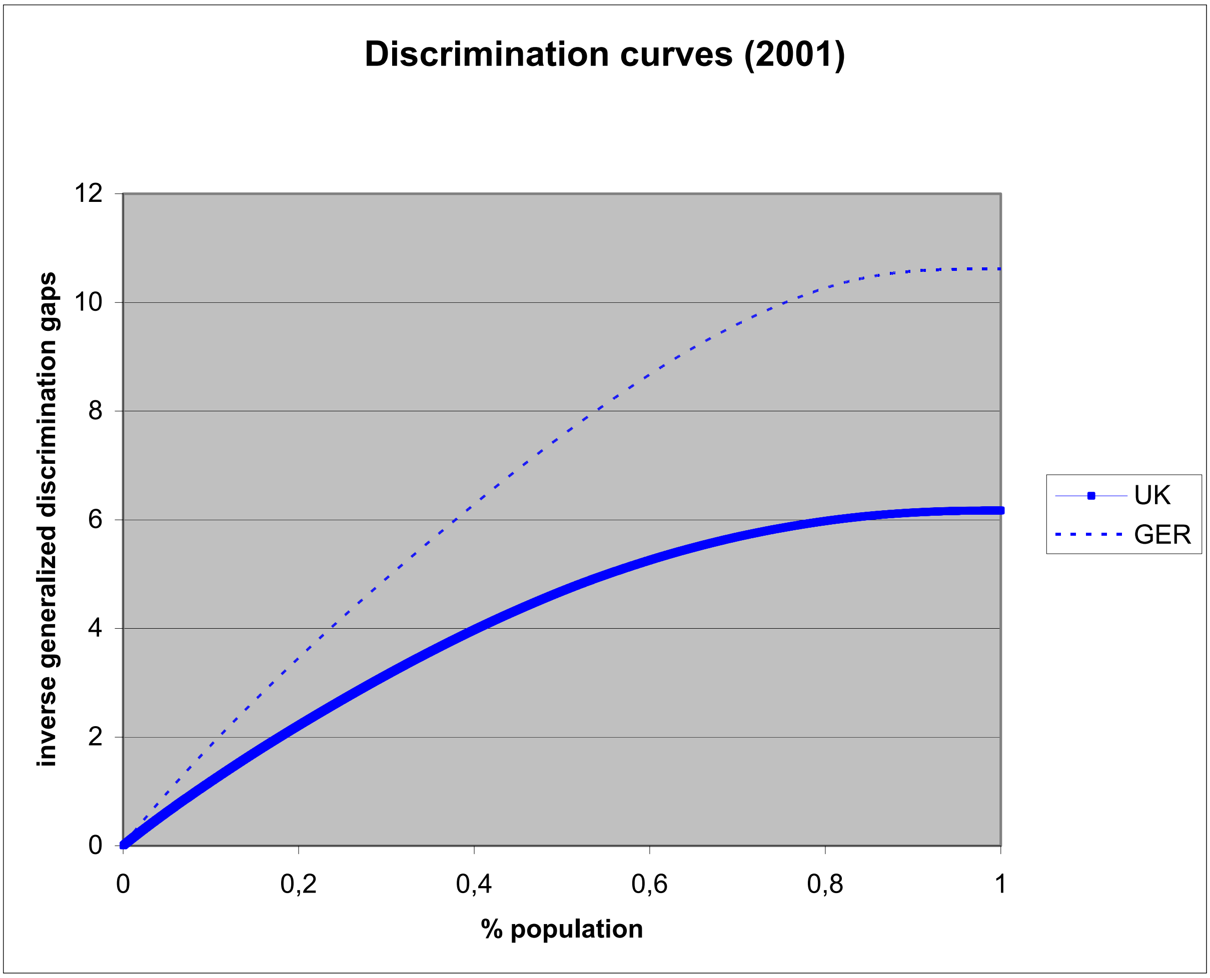

Figure 3.A Discrimination change in Germany (1994-2001)

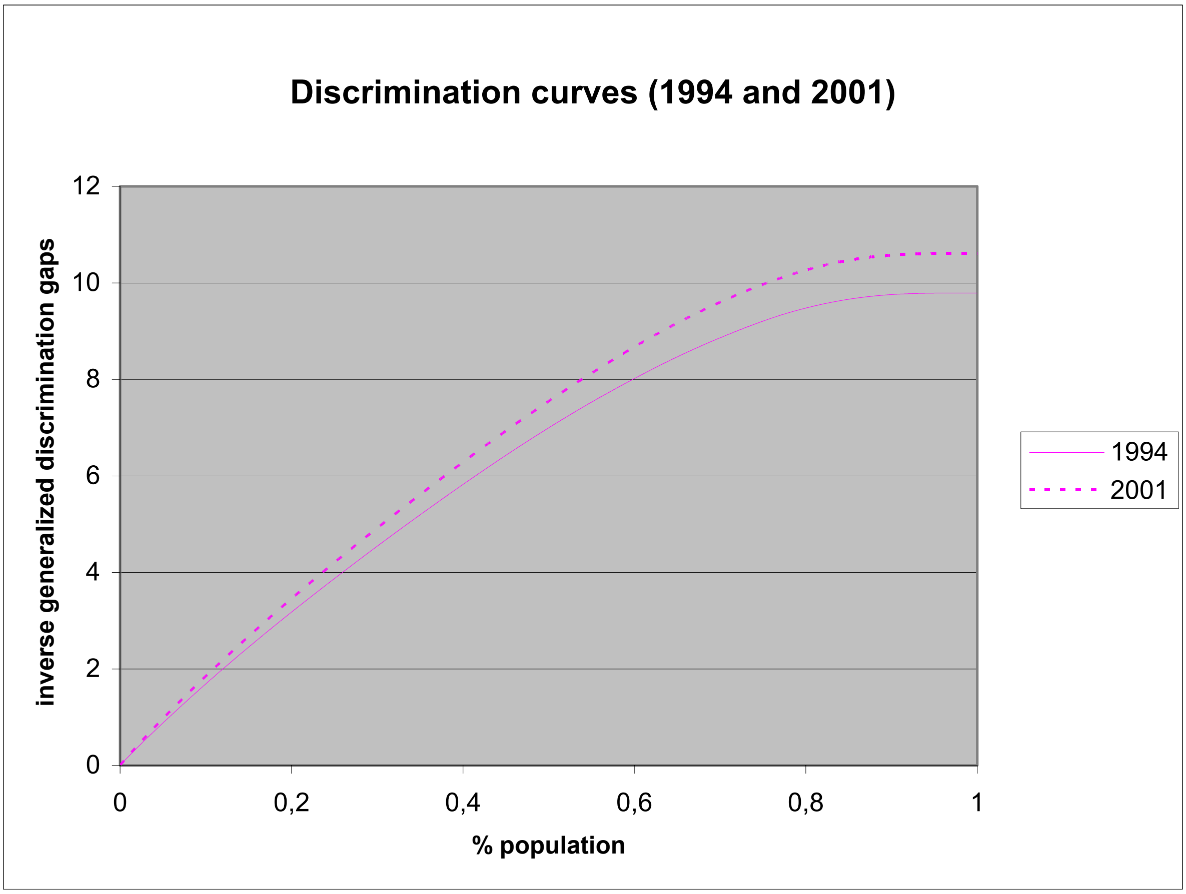

Figure 3.B Discrimination change in the United Kingdom (1994-2001)