Why Do I Like People Like Me?*

Manuel F. Bagues **María José Pérez-Villadóniga**

Universidad Carlos III Universidad de Oviedo

December 24, 2009

Abstract

In this paper we extend the standard model of statistical discrimination to a multidimensional framework where the accuracy of evaluators depends on how knowledgeable they are in each dimension. The model yields two main implications. First, candidates who excel in the same dimensions as the evaluator tend to be preferred. Second, if two equally productive groups of workers differ in their distribution of ability across dimensions group discrimination will arise unless (i) evaluators are well informed about the extent of these differences and (ii) evaluators are allowed to take candidates' group belonging into account in their assessments. These results provide support for affirmative action policies and suggest that in some cases group blind antidiscrimination policies may actually be counterproductive.

JEL codes: J71, D82

Keywords: Statistical Discrimination, Affirmative Action.

*We would like to thank Antonio Cabrales, Juan Jose Dolado, Florentino Felgueroso, Walter García-Fontes, Juan F. Jimeno, Günter Fink, Marco Haas, Franck Malherbet, Pedro Marín, Eduardo Melero, Nicola Pavoni and Enrico Pennings for many useful comments on an earlier version of this paper. We are also grateful to participants in seminars at the universities of Bocconi, Verona, Pavia, Las Palmas de Gran Canaria, Oviedo, León, Sant'Anna School of Advanced Studies, FEDEA and conference presentations at the *Jornadas de Economía Industrial*, *Simposio de Análisis Económico* and SOLE-EALE. Any remaining error is our own.

1 Introduction

The fact that individuals might be treated differently according to exogenous characteristics such as gender, age or race has been well documented in the literature. Most of the evidence refers to the labour market, where differences in wages or hiring and promotion that cannot be accounted for by differences in productivity have been observed.[1] Discriminatory behaviours have also been observed in housing decisions (Massey and Denton 1993), lending (Hunter and Walker 1996), car selling (Ayres and Siegelman 1995) or even in the refereeing of academic papers (Blank 1991; Fisher et al. 1994).

In the economics literature, two distinct general sets of explanations that focus on the demand side of the labour market have been proposed to explain the origin and persistence of discrimination. On the one hand, taste models, as in Gary Becker's (1957) seminal work, suggest a preference-based motivation for the existence of discrimination. The difference in wages between two equally productive groups of workers arises because employers, customers or co-workers dislike interacting with employees that belong to certain groups. On the other hand, statistical models of discrimination argue that, in the presence of information asymmetries about the real productivity of workers, the group-belonging of an individual can be considered as a signal that provides additional information. Groups of workers may differ in their expected productivity (Phelps 1972, Lazear and Rosen 1990) or in the reliability of the observable signals (Aigner and Cain 1977, Cornell and Welch 1996). In this context, taking into account an individual's group affiliation may be a rational response to its informational content and the wage gap might persist in the long run.

In this paper we extend the standard model of statistical discrimination

[1]For a survey see, for instance, Altonji and Blank (1999).

presented by Phelps (1972) and we introduce two novel features. First, we allow for the existence of multiple dimensions of ability. These dimensions can be understood either as different tasks that the worker needs to undertake, or as separable skills that are required to perform a single task. Second, while standard models of statistical discrimination assume that the accuracy with which employers assess the productivity of potential employees is exogenous or depends on the group belonging of the employer and the candidate,[2] we assume that the capability of an employer to evaluate quality at a certain dimension increases with her knowledge of that dimension. This assumption is consistent with experimental evidence, where it has often been found that, in many dimensions, individuals who are less competent are also relatively less accurate at evaluating ability.[3]

Combining these features our model yields the following two predictions. First, we show that a similar-to-me-in-skills effect arises in the evaluation. Since individuals can assess knowledge more accurately at those dimensions where they are more knowledgeable, an employer who makes an optimal use of the available information will give relatively more weight to those signals that are observed in dimensions where she is most knowledgeable. As a result, given any two equally productive candidates, the employer will tend to give a higher valuation to the candidate who excels in the same dimensions as she does. This result is consistent with the evidence found by Bagues et al. (2009) who observe that, in public exams in Spain, evaluators take most into account candidates' performance in those dimensions in which they are themselves relatively more knowledgeable. More generally, the fact that evaluators tend

[2]See Phelps 1972, Arrow 1973, Aigner and Cain 1977 or Cornell and Welch 1996.

[3]Knowledgeable people are more accurate in their evaluations in the field of chess (Chi 1978), physics (Chi et al. 1982), grammar (Kruger and Dunning 1999) or academic performance (Everson and Tobias 1998). Similarly, in the context of firms R&D strategies, Cohen and Levinthal (1990) argue that "the ability of a firm to recognize the value of new, external information, is largely a function of the firm's level of prior related knowledge".

to give higher ratings to candidates who are similar to themselves has been widely documented both in psychology and sociology (Byrne 1971) and in organizational processes, such as supervisors' assessments of subordinates or recruitment (Goldberg 2005).

Second, the model shows that, even if members of different groups are equally productive, group discrimination might arise if groups differ in their distribution of ability across dimensions.[4] In particular, group discrimination will arise if (i) employers are not fully aware of the extent of these differences or (ii) employers are perfectly informed but cannot condition their evaluations on candidates' group-belonging. The intuition behind this result is the following. Employers will tend to give more weight to signals that have been observed in those dimensions where they are more knowledgeable. In principle this favours candidates belonging to the same group as the employer, as they are more likely to excel precisely in these dimensions. Still, a well-informed evaluator who was allowed to take into account the group belonging of candidates might use this information in order to adjust her priors appropiately. This would not only be efficient from an informational point of view but, as well, it would yield similar average evaluations across groups of candidates.

Our model overcomes some of the drawbacks of standard models of statistical discrimination (Aigner and Cain 1977, Cornell and Welch 1996). First, these models assume that the screening technology is exogenous or depends on the group belonging of the candidate. However, empirical evidence supporting that the accuracy of evaluations depends on the groups belonging of candidates is scarce.[5] In contrast, in our setup the accuracy of the evaluation

[4]Following Aigner and Cain (1977), we consider group discrimination as the situation where "groups that have the same average ability may receive different average pay" (pp.178). Note that in a multidimensional framework the term *same ability* should be interpreted as meaning *same total ability* rather than *same ability at every dimension*.

[5]Up to our knowledge, and according to Holzer and Neumark (2000), Neumark (1999)

depends on the evaluator's knowledge of each dimension, an assumption that is consistent with an abundant literature which finds a positive relationship between evaluators' knowledge of a field and the quality of their assessment.[6] Second, our model produces testable implications that fit better the available empirical evidence. Standard models of statistical discrimination predict that among highly productive candidates, those belonging to the evaluator's group will tend to be hired but, when all candidates are relatively unproductive, those who do not belong to the employer's group will tend to be preferred, given that the observed (low) signal about their quality is a weaker indicator of their productivity. Still, up to our knowledge there is no empirical evidence supporting the later implication, this is, the reversal of the race and gender gap for low productivity levels. In contrast, in the (multidimensional) model proposed here those candidates akin to the evaluator tend to be preferred for every level of productivity, an implication that is supported by most empirical studies on discrimination.

The paper is structured as follows. In the remaining of this section we offer an example that helps to clarify the intuition underlying our model. The formal model is discussed in Section 2 and in Section 3 we discuss the potencial policy implications of the model.

1.1 Example: The Academic Job Market

Every year, PhD candidates in Economics attend the academic job market. In terms of reseach, a good candidate is understood as someone who will be able to make a relevant scientific contribution. Since the true quality of candidates cannot be immediately disclosed, evaluators typically rely on

provides the only empirical analysis about how the reliability of labor market information varies across groups of applicants.

[6] See footnote 4.

candidates' job market paper in order to infer quality.[7]

Evaluating a paper is in general a multidimensional task. First, a good paper should provide an interesting and novel economic idea. Second, ideally this idea should be presented through an elegant mathematical formalization. Evaluating each of these dimensions is likely to be complex and evaluators usually will be able to assess more accurately the quality of the paper in that dimension where they are themselves more knowledgeable[8]. For instance, while evaluating the novelty of a paper requires knowledge of the previous related literature, in order to appreciate its mathematical quality being skilled in mathematics might be highly convenient.

Following with the example, let us assume for simplicity that the total quality of a paper is equal to the sum of its quality in these two dimensions: its quality in terms of economic novelty or relevance and its quality in terms of mathematical elegance. Let us consider the case of an evaluator who is perfectly knowledgeable in mathematics but who ignores completely any related literature. In this case, given the evaluator's incapability to appreciate accurately the economic novelty of the paper, her optimal evaluation involves taking mostly into account the information that she observes along the mathematical dimension. Therefore, if this evaluator had to evaluate two papers of identical quality, the evaluator would tend to give a higher valuation to the paper which excels relatively more in the mathematical dimension.

Moreover, imagine that there exist gender differences in the distribution of ability across dimensions. Consider, for instance, a scenario where the (total) quality of papers is independent of the author's gender but, where

[7]As Levinovitz and Ringertz (2001) point out, "it usually takes a longer time in economics (and social sciences in general) than in the natural sciences to find out if a new contribution is solid or if it is just a fad. In other words, it is important to wait for scrutiny, criticism and repeated tests of the quality and relevance of a contribution."

[8]The fact that a paper's evaluation may depend heavily on the characteristics of the evaluator is illustrated by the low correlation -only 0.24- that Blank (1991) found in the ratings given by the two referees to papers submitted to the American Economic Review.

male economists tend to have a higher capability for mathematical abstraction, while female economists are better in terms of the economic relevance of their work. If this was the case, the above similar-to-me-in-skills effect could also generate gender discrimination if evaluators are unaware of the existence of these gender differences or if they are not allowed to condition their assessments on candidates' gender. The intuition is the following. Evaluators' optimal evaluation involves giving more weight to information that is observed along the dimension where they excel. This favours candidates belonging to the same gender as the evaluator, as they are more likely to excel in that dimension. For instance, if the evaluator was male, papers produced by male candidates would tend to obtain a higher valuation since the (male) evaluator's optimal evaluation would involve giving a higher consideration to the information observed across the mathematical dimension. This gender discrimination would only disappear if the evaluator was well informed about the existence of such gender differences in the distribution of quality and, most importantly, if he was allowed to adjust his priors at each dimension taking into account candidates' gender.

2 The model

Let us consider the case of an individual whose total quality depends on his abilities or skills in a number D of different dimensions or fields [$q_i = f(x_{i1}, ..., x_{iD})$]. These fields can be understood as different tasks that the worker needs to undertake or as separable skills that are required to perform a single task. Candidates' abilities are exogenously given and independently and normally distributed,

$$\mathbf{x}_i \rightarrow N\left(\mathbf{p}_i, \mathbf{\Sigma}\right)$$

where $\mathbf{x}_i$ represents the Dx1 vector of abilities, $\mathbf{p}_i$ is a Dx1 vector of mean abilities and $\mathbf{\Sigma}$ is the corresponding variance-covariance matrix.

Without loss of generality, we impose two simplifying assumptions on the populational distribution of quality. First, we restrict the variance of quality to be equal across dimensions and normalize it equal to one. With this constraint we want to avoid a more general case where ability may vary systematically more along certain dimensions.

$$Var\left(x_{id}\right) = 1 \qquad \forall d = 1, ..., D$$

Second, we assume that an individual's ability along a certain field is independent of his ability along any other dimension. In other words, the knowledge of an individual's ability in one dimension does not provide any information about his ability in any other dimension.[9]

$$E\left(x_{id}x_{id'}\right) = 0 \qquad \forall d \neq d'$$

These two assumptions restrict $\mathbf{\Sigma}$ to be a diagonal matrix where the on-the-diagonal elements are equal to one.

In this multidimensional framework let us consider the case where individuals' total productivity is not observable but evaluators can observe some noisy and imperfect signal of candidates' ability at each dimension. These signals could be interpreted as the result of some tests or job interviews and their value will be a function of the candidates' true ability at each specific field plus an error term η which is assumed to be independently and normally distributed with zero mean and finite variance.

$$y_{id} = x_{id} + \eta_{id} \qquad \text{where } \eta_{id} \rightarrow N\left(0, \sigma_{\eta_d}\right)$$

[9] As long as there exists some kind of multidimensionality, this is, provided that quality in different dimensions is not perfectly correlated, dimensions could always be appropriately redefined such that this condition is satisfied.

Moreover, let us assume that in each dimension the accuracy of the signal is independent of the quality of the candidate who is being evaluated.

$$E\left(x_{id}\eta_{id}\right) = 0$$

Given the above assumptions, evaluator h will infer the quality of candidate i in dimension d as the weighted sum of the signal observed in this dimension and the distributional prior, where the weight given to the signal will depend on how accurately this signal is perceived by the evaluator:

$$E_h\left(x_{id}/y_{id}\right) = \gamma_d^h y_{id} + \left(1 - \gamma_d^h\right) p_{id} \tag{1}$$

where $p_{id} = E\left(x_{id}\right)$ and $\gamma_d^h = \frac{E_h(x_{id}y_{id})}{E_h(y_{id}y_{id})} = \frac{1}{1+\sigma_{\eta_d^h}}$. If, for simplicity, we assume that a candidate's total productivity is equal to the sum of his quality at each dimension $\left[q_i = f\left(x_{i1}, ..., x_{iD}\right) = \sum\limits_{d \in D} x_{id}\right]$, it follows that:

$$E_h\left(q_i/y_{i1}, ..., y_{iD}\right) = \sum_{d \in D} \left[\gamma_d^h y_{id} + \left(1 - \gamma_d^h\right) p_{id}\right]$$

This is, employer h will take relatively more into account those signals that she observes in fields where she can assess information more accurately.

2.1 Similar-to-me-in-skills effect

Let us define an evaluation as being complex as the situation where an evaluator's relative ability to assess quality is positively related to her own quality. More precisely, in a context where, without loss of generality, D is equal to two, we define an evaluation as being complex if:

$$x_{h1} > x_{h2} \Longrightarrow \sigma_{\eta_1^h} < \sigma_{\eta_2^h} \tag{2}$$

It easily follows that when the evaluation is complex, an evaluator who makes an optimal use of the available information will give a larger weight to those signals that have been observed in that dimension where her own ability is larger. This is, given an evaluator h,

$$x_{h1} > x_{h2} \Longrightarrow \gamma_1^h > \gamma_2^h$$

As a result, faced with two equally productive candidates i and j, evaluator h will tend to give a higher evaluation to the candidate who excels in the same dimension where she herself is best. More precisely,

Proposition 1 *Similar-to-me-in-skills effect*

$$q_i = q_j,\ x_{h1} > x_{h2}\ \&\ x_{i1} > x_{j1} \Rightarrow E_h[q_i] > E_h[q_j]$$

Proof. The evaluator h, who observes at each dimension d a noisy signal of quality, can use y_d as a least-squares predictor of the candidate's true ability in that dimension, x_d, according to the regression-type relation $x_d = \gamma_d^h y_d + \left(1 - \gamma_d^h\right) p_{id} + u_d$ where $E[y_d u_d] = 0$ and coefficient γ_d^h will be determined by the accuracy of the signal $\left[\sigma_{\eta_d^h}\right]$ [as in equation (1)]. The evaluator's expected valuation of candidate i's total productivity will be equal to:

$$E_h[q_i] = E_h[x_{i1} + x_{i2}] = E_h\left[\sum_{d=1,2}\left(\gamma_d^h y_{id} + \left(1-\gamma_d^h\right)p_d\right)\right] = \sum_{d=1,2}\left(\gamma_d^h x_{id} + \left(1-\gamma_d^h\right)p_d\right)$$

Candidate j's productivity can be estimated in a similar way. The difference in the expected observed quality of the two candidates is equal to:

$$E_h[q_i] - E_h[q_j] = \sum_{d=1,2}\left(\gamma_d^h x_{id} + \left(1-\gamma_d^h\right)p_d\right) - \sum_{d=1,2}\left(\gamma_d^h x_{jd} + \left(1-\gamma_d^h\right)p_d\right) = \sum_{d=1,2}\gamma_d^h\left(x_{id} - x_{jd}\right)$$

which is positive since $q_i = q_j \Longrightarrow x_{i1} - x_{j1} = x_{j2} - x_{i2} > 0$ and $x_{h1} > x_{h2} \Longrightarrow \gamma_1^h > \gamma_2^h$. ■

2.2 In-group bias

As shown above, when productivity is multidimensional and the evaluation is complex, evaluators are more likely to give a higher evaluation to candidates alike to them. In this subsection we investigate whether the existence of this similar-to-me-in-skills effect can generate an in-group bias. This is, if individuals belonging to the same group tend to possess knowledge in the same dimensions, will evaluators have a tendency to prefer group mates over equally productive candidates from other groups? As we will see, it depends on whether the evaluator is well informed about the extent of inter-group differences and whether she is allowed to take them into account.

Consider that individuals may belong to different groups defined according to gender, age, or some other easily observable and exogenous characteristic. Let us also consider the case where there are only two groups g_1 and g_2 and where candidates' total productivity is independent of group belonging:

$$E\left[q_i / i \in g_1\right] = \overline{q}^{(g_1)} = \overline{q}^{(g_2)} = E\left[q_j / j \in g_2\right] \tag{3}$$

This assumption does not prevent the possibility that members of the two groups tend to excel in different dimensions. More particularly, let us represent the existence of group-related variations in the distribution of quality in the following way:

$$x_{id} = \sum_{g=g_1,g_2} x_{id}^{(g)} + \mu_{id} \qquad d = 1, 2.$$

where $x_{id}^{(g)} = \left(p_d^{(g)} + \varepsilon_{id}\right) c_i^{(g)}$ and $c_i^{(g)} = 1$ if candidate i belongs to group

g and zero otherwise. Let us also assume that μ_{id} and ε_{id} are normally and independently distributed with zero mean. Therefore, $x_{id}^{(g)}$ measures the differences in dimension d that can be explained by the candidate's belonging to group g, and $p_d^{(g)}$ is the expected ability in dimension d of individuals in group g. Finally, let us, for simplicity, consider the case where the distribution of quality across groups is symmetric so that the following condition is satisfied:

$$p_1^{(g_1)} = p_2^{(g_2)} \ \&\ p_2^{(g_1)} = p_1^{(g_2)} \tag{4}$$

In this set up, evaluators will estimate candidates' quality in a similar way as in (1). Let us define $\lambda_{id} = \mu_{id} + \sum\limits_{g=g_1,g_2} \varepsilon_{id} c_i^{(g)}$ and $z_{id} = \sum\limits_{g=g_1,g_2} p_d^{(g)} c_i^{(g)}$. Therefore $x_{id} = z_{id} + \lambda_{id}$ and, given that $y_{id} = x_{id} + \eta_{id}$, it follows that in each dimension the relationship between quality and signal, net of the group effect, will be equal to $x_{id} - z_{id} = \gamma_d^h (y_{id} - z_{id}) + u_{id}$. Thus, $E_h(x_{id}) = E_h\left[\gamma_d^h y_{id} + \left(1 - \gamma_d^h\right) z_{id}\right]$ where $\gamma_d^h = \frac{Var(\lambda_{id})}{Var(\lambda_{id}) + Var(\eta_{id})} = \frac{\sigma_\lambda}{\sigma_\lambda + \sigma_{\eta_d^h}}$.

In our analysis we will distinguish between two possible situations. First, we present the case where in their evaluation employers may take into account candidates' observable signals of quality but do not condition their evaluation on candidates' group belonging. Second, we study the case where the evaluators condition their evaluation both on the observed signals of quality and candidates' group belonging.

2.2.1 Non-discriminatory practices

Let us define as non-discriminatory practices those situations where evaluators do not condition their evaluation on candidates' group belonging [$\forall i \forall d \qquad E_h(z_{id}) = p_d$]. Several reasons may prevent evaluators from taking into account the group belonging of candidates. Evaluators may not be aware of the existence of differences in quality profiles across groups. As well, even if evaluators are well informed about these differences, they may be restricted not to use this information. This is the case, for instance, in many firms and institutions where the hiring process is subject to a strict equal opportunities policy.

Paradoxically, when members of different groups are, on average, equally productive but differ in their distribution of ability, if the evaluator does not or cannot take into account candidates' group belonging, individuals belonging to her own group will tend to be favoured.

Proposition 2 *Non-discriminatory practices yield discriminatory outcomes*

$$\begin{aligned} \overline{q}^{(g_1)} &= \overline{q}^{(g_2)},\, p_d^{(g_1)} \neq p_d^{(g_2)} \ \&\ E_h\left(z_{id}\right) = E_h\left(z_{jd}\right) = z_d \Longrightarrow \\ &\Longrightarrow \quad E_h\left(q_i\right) > E_h\left(q_j\right) \qquad i, h \in g_1, j \in g_2, d = 1, 2. \end{aligned}$$

Proof. Without loss of generality let us assume that members of group g_1 tend to excel in dimension one $\left[p_1^{(g_1)} > p_2^{(g_1)}\right]$. Let us also for simplicity consider the case where the evaluator h is a typical group g_1 member such that $x_{h1} > x_{h2}$, so that from assumption (2) it follows that $\gamma_1^{(h)} > \gamma_2^{(h)}$. Then,

$$E_h\left(q_i\right) - E_h\left(q_j\right) = E_h\left[\sum_{d=1,2}\left(\gamma_d^h y_{id} + \left(1-\gamma_d^h\right) z_{id}\right)\right] - E_h\left[\sum_{d=1,2}\left(\gamma_d^h y_{jd} + \left(1-\gamma_d^h\right) z_{jd}\right)\right] =$$

$$= \left\{E_h\left(z_{id}\right) = E_h\left(z_{jd}\right) = p_d\right\} = \sum_{d=1,2}\left(\gamma_d^h p_d^{(g_1)} + \left(1-\gamma_d^h\right) p_d\right) - \sum_{d=1,2}\left(\gamma_d^h p_d^{(g_2)} + \left(1-\gamma_d^h\right) p_d\right) =$$

$$= \sum_{d=1,2}\left[\gamma_d^h\left(p_d^{(g_1)} - p_d^{(g_2)}\right)\right] = \{\text{by } (4)\} = (\gamma_1^{(h)} - \gamma_2^{(h)})(\bar{x}_1^{(g)} - \bar{x}_2^{(g)}) > 0$$

■

This is, in a framework where evaluating is complex, if groups differ in their distribution of quality and evaluators do not take into account group-belonging, they will assign a higher valuation to those candidates that excel in the same dimensions as they do and, since the distribution of ability across fields is group dependent, this bias will tend to favour candidates that belong to the same group as the evaluator. This is, not taking into account group priors is not only informationally suboptimal but, moreover, it generates discriminatory outcomes.

2.2.2 Discriminatory Practices

If employers observe that employees belonging to certain groups tend to perform better on certain dimensions, it is likely that these employers will update their beliefs and they will take into account this information in their evaluations, at least, as long as they are allowed to do so. If the evaluator can condition her evaluation both on the observed quality signals and on the group belonging of the candidates, then any two equally productive candidates will tend to obtain the same valuations independently of group belonging.

Proposition 3 *Discriminatory practices yield non-discriminatory outcomes*

$$\begin{aligned}\overline{q}^{(g_1)} &= \overline{q}^{(g_2)},\, p_d^{(g_1)} \neq p_d^{(g_2)},\, E_h\left(z_{id}\right) = p_d^{(g_1)} \,\&\, E_h\left(z_{jd}\right) = p_d^{(g_2)} \Longrightarrow \\ &\Longrightarrow E_h\left(q_i\right) > E_h\left(q_j\right) \qquad i, h \in g_1, j \in g_2, d = 1, 2.\end{aligned}$$

Proof. As in proposition (1), without loss of generality let us assume that members of group g_1 tend to excel in dimension one $\left[p_1^{(g_1)} > p_2^{(g_1)}\right]$ and let us also for simplicity consider the case where the evaluator h is a typical group g_1 member such that $x_{h1} > x_{h2}$, so that from assumption (2) it follows that $\gamma_1^{(h)} > \gamma_2^{(h)}$. Then,

$$E_h\left(q_i\right) - E_h\left(q_j\right) = E_h\left[\sum_{d=1,2}\left(\gamma_d^h y_{id} + \left(1-\gamma_d^h\right) z_{id}\right)\right] - E_h\left[\sum_{d=1,2}\left(\gamma_d^h y_{jd} + \left(1-\gamma_d^h\right) z_{jd}\right)\right] =$$

$$= \sum_{d=1,2}\left(\gamma_d^h p_d^{(g_1)} + \left(1-\gamma_d^h\right) p_d^{(g_1)}\right) - \sum_{d=1,2}\left(\gamma_d^h p_d^{(g_2)} + \left(1-\gamma_d^h\right) p_d^{(g_2)}\right) = \overline{q}^{(g_1)} - \overline{q}^{(g_2)} = 0$$

■

This is, if well-informed employers may condition their evaluation on the group belonging of candidates the outcome of evaluations will be independent of the evaluators' group belonging.

3 Conclusion

In this paper we build on the standard model of statistical discrimination where an employer must select a candidate out of a pool of applicants in a context of imperfect information. Our main departure from the traditional framework is to allow for the existence of multiple dimensions of ability and to make the accuracy of the evaluation at each dimension depend on the evaluators' knowledge of this dimension. The model yields two main results. First, it rationalizes the existence of a similar-to-me-in-skills effect which favours candidates who excel in the same dimensions as the evaluator. Second, the model casts doubts on the capability of blind evaluations to eradicate discrimination. It is shown that, if groups of individuals -according to their gender, race or any other observable and exogenously given characteristic- differ in their distribution of ability across dimensions, group discrimination will arise unless evaluators are well informed about the extent of these differences and, moreover, they can condition their assessments on candidates' group belonging.

There are many cases where this might not be possible. Sometimes employers might not be fully aware of the fact that groups are equally productive but differ in terms of how good they are in each dimension. This may happen when groups have little interaction, when the size of the minority is relatively small[10] or in the presence of a number of cognitive biases such as observational selection bias, availability bias or anchoring that can generate a divergence between individuals' perception of other groups' quality at each dimension and their true quality distribution. As well, even if evaluators were well informed, in some cases evaluators may not be able to take group belonging into account. This may occur if evaluators for some reason do not observe candidates' group belonging or if, even if they observe it, they have been instructed not to take it into account. This may be the case in the presence of group blind antidiscrimination policies according to which evaluators are explicitly ordered not to consider candidates' group belonging or if candidates' identity is kept anonymous (as in Blank 1991 or Goldin and Rouse 2000). Paradoxically, in the framework considered here these policies agravate discrimination. On the contrary, when candidates' quality is not easily observable and productivity is multidimensional, affirmative action may be the only way to promote equal opportunity.

[10] As it would increase the cost of rationality. See for instance Fryer and Jackson (2007).

References

[1] Aigner, D. J. and G. Cain (1977), "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30(2): 175-87.

[2] Altonji, J.G. and R. M. Blank (1999) "Race and Gender in the Labor Market" in: O. Ashenfelter and D.Card. eds., Handbook of Labor Economics, vol. 3, (North-Holland, Amsterdam): 3143-3259.

[3] Arrow, K. J. "The Theory of Discrimination" (1973) in Orley Ashenfelter and Albert Reeds, eds., *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press: 3-33.

[4] Ayres, I. and P. Siegelman (1995) "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, 85(3): 304-321.

[5] Bagues, M, and M. Pérez-Villadóniga (2009) "Do Recruiters Prefer Applicants with Similar Skills? Evidence from a Randomized Natural Experiment," Universidad Carlos III Working Paper 2009/62.

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

[7] Blank, R. M. (1991) "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review," *American Economic Review*, 81(5): 1041-1067.

[8] Byrne, D. (1971), "The Attraction Paradigm," New York: Academic Press.

[9] Chi, M. T. H. (1978). "Knowledge structures and memory development." (In R. Siegler (Ed.), Children's thinking: What develops? (pp. 7396). Hillsdale, NJ: Erlbaum.)

[10] Chi, M. T. H., Glaser, R. and E. Rees (1982). Expertise in problem solving.(In R. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 1776). Hillsdale, NJ: Erlbaum.)

[11] Cohen, W. and D. Levinthal (1990) "Absorptive capacity: A new perspective on learning and innovation," *Administrative Science Quarterly*, 35:128-152.

[12] Cornell, B. and I. Welch (1996) "Culture, Information, and Screening Discrimination," *Journal of Political Economy*, 104(3): 542-71.

[13] Everson, H. T. and S. Tobias (1998) "The Ability to Estimate Knowledge and Performance in College: A Metacognitive Analysis," *Instructional Science*, 26: 65-79.

[14] Fisher, M., Stanford B., Friedman M.D. and B. Strauss (1994) "The Effects of Blinding on Acceptance of Research Papers by Peer Review," *Journal of American Medical Association*, Vol. 272(2):143-146l.

[15] Fryer, R. G. and M. O. Jackson (2007) "A Categorical Model of Cognition and Biased Decision-Making," forthcoming in the *B.E. Press Journal of Theoretical Economics (Contributions)*.

[16] Goldberg, Caren B. (2005), "Relational Demography and Similarity Attraction in Interview Assessments and Subsequent Offer Decisions," *Group and Organization Marketing*, Vol. 30(6): 597-624.

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

[18] Holzer, H. and D. Neumark (2000) "Assessing Affirmative Action," *Journal of Economic Literature*, 38: 483-568.

[19] Hunter, W. C. and M. B. Walker (1996) “The Cultural Affinity Hypothesis and Mortgage Lending Decisions,” *Journal of Real Estate, Finance and Economics* 13: 57-70.

[20] Kruger, J. and D. Dunning (1999) “Unskilled and Unaware if It: How Difficulties in Recognizing One’s Own Incompetence Lead to Inflated Self-Assessments,” *Journal of Personality and Social Psychology,* 6: 1121-34.

[21] Lazear, E. P. and S. Rosen (1990) “Male-Female Wage Differentials in Job Ladders,” *Journal of Labor Economics,* 8(1): S106-23.

[22] Levinovitz, A. W. and N. Ringertz (2001), ”The Nobel Prize: The first 100 years,” Imperial College Press and World Scientific Publishing Co. Pte. Ltd.

[23] Massey, D., and N. Denton (1993) “American Apartheid: Segregation and the Making of the Underclass,” Cambridge, Mass.: Harvard University Press.

[24] Neumark, D. (1999) “Wage differentials by sex and race: The roles of taste discrimination and labor market information,” *Industrial Relations: A Journal of Economy and Society*, 38(3): 414-45.

[25] Phelps, E. (1972) “The Statistical Theory of Racism and Sexism,” *American Economic Review,* 62(4): 659-61.