Two Concepts of Discrimination:
Inequality of Opportunity *versus* Unequal Treatment of Equals

Bart Capéau
SBS-EM, ECARES, Université Libre de Bruxelles

Lieve Eeman
Katholieke Universiteit Leuven

Steven Groenez
Katholieke Universiteit Leuven

Miet Lamberts
Katholieke Universiteit Leuven

ECARES working paper 2012-021

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Two concepts of discrimination: inequality of opportunity *versus* unequal treatment of equals*

Bart Capéau†, Lieve Eeman, Steven Groenez, Miet Lamberts‡

Revised, July 2012

* This is a revised and abbreviated version of a paper that was first circulating under the title "A proposal to measure multidimensional discrimination". The longer version, in which we discuss some unfruitful alternative specifications of the regressions we tried out, is still available upon request.

† ECARES – ULB, Avenue F.D. Roosevelt, 50 CP114, B-1050 Bruxelles, Belgium.

‡ Research Institute for Work and Society, K.U.Leuven, Parkstraat 47, B-3000 Leuven, Belgium.

Correspondence to: bcapeau@ulb.ac.be or bart.capeau@yahoo.com.

We thank seminar participants at HIVA, Michaël Bennetsen, Patrick Charlier, Hans Bruyninckx, Edouard Delruelle, Carla Rijmenams, Michel Vanderkam, Michiel van de Voorde and Hildegard Van Hove for their remarks during this research. The financial support of the CGKR and the HIVA–hefboomfonds is gratefully acknowledged. The content of this paper is the sole responsibility of the authors and does not reflect the views of any institution. All remaining errors in the text are ours.

Abstract

We propose a novel way to cope with the measurement of discrimination in a multidimensional context. We adopt from the pairwise approach the idea that discrimination occurs when a (systematic) distinction is made between two persons on the basis of personal traits (sex, age, national origin or physical state) that are normatively considered as unjustifiable reasons to treat persons differently.

We extend this approach to a situation in which the types differ with respect to more than one of the dimensions that are potential sources of discrimination. We define and compare two concepts of discrimination: inequality of opportunity *versus* unequal treatment of equals. The first looks at overall differences in output scores (*e.g.* the probability to be selected). The second aggregates the occurrences of differences in treatments between persons that should be treated equally in the absence of discrimination. We show how both concepts can be quantified by regression coefficients of discrete choice models.

Our proposal is implemented empirically on a database of 1708 fake resumes sent out during 2010-11 on the Belgian vacancy market. The applicants differed with respect to sex, age, national origin or a particular physical trait or state (a physical limitation or being pregnant). The reaction of the selection managers in terms of an invitation for the candidate to the next step in the hiring procedure is the output variable that was analysed. We argue in favour of the second concept (unequal treatment of equals) as the first concept of discrimination (inequality of opportunity) fails to detect discrimination in some cases in which there is, presumably, consensus about its occurrence. Reversely, in some cases this concept leads to the conclusion that there is discrimination, while many would agree that the differences measured are the consequence of other factors than discrimination.
JEL codes: C93, J71

Keywords: Labour market discrimination, Field experiments.

Two concepts of discrimination: inequality of opportunity *versus* unequal treatment of equals

1 Introduction

There is a growing consensus in the literature that discrimination can only be measured by field experiments (see a.o. Bovenkerk, 1992, Fix and Struyk, 1993 and the overviews by Riach and Rich, 2002, 2004a). In the beginning, these experiments were organised by sending out real persons that resembled as close as possible in all respects, except for the variable of interest (sex, age, national origin, etc.), to apply *e.g.* for the same jobs. In many cases, these persons (auditors) were trained to behave similarly and to reveal similar information during the job interview. When the salient difference between the auditors influences the probability to be hired, it is concluded that there is discrimination. However, some pervasive critiques on this approach remain (Heckman and Siegelman, 1993, Heckman, 1998). Heckman and Siegelman (1993) argue that differences in the variance of characteristics not observed and beyond the control of the experimenter, can affect the ratio of, *in an experimental setting observed*, probabilities of both groups to be selected, even though the true population probabilities are the same.[1] Moreover, the direction of this bias depends upon the level on which people are equalised in observable characteristics. Suppose one investigates differences in hiring probabilities between a minority and majority group, due to discrimination.[2] As was mentioned, one should select the auditors such that they are as similar as possible in all respects, except for the characteristic of interest. One thus has to decide upon the level of the other characteristics, such as their degree of education, they (pretend to) possess. Suppose it is decided they reveal a low score on those observable characteristics that facilitate largely getting the job. Assume that the unobservable characteristics influencing the probability to get the job, are independent from the observables, and symmetrically and identically distributed for both, the minority and majority group, except for their variance, which is higher for the minor-

[1] One can think of small differences in the attitude of both applicants during the interview, or particular physical traits, such as tallness, weight, or how old they look.

[2] Minority and majority are just names for the groups here.

ity group than for the majority group. In order to be selected, given the low score on the observable characteristics, a high score on the unobservables should be attained. Given the distributional assumptions on the unobservables, the minority group will surpass this selection threshold more often, and therefore will be observed to be hired more, even though there is no discrimination on the basis of the characteristic which we try to measure.

Some argue (Lahey and Beasley, 2009, Oreopoulos, 2011) that sending written resumes, with all remaining differences sufficiently randomised, can cure the major part of the objections made by Heckman and Siegelman (1993) and repeated by Heckman (1998). But Heckman (1998) insists that his critique holds for resume or correspondence studies too. Neumark (2011) recently made a statistical proposal to meet some of the issues raised by Heckman (1998), in the case of resume studies.

According to our opinion, the basic problem is not in the first place a statistical one. Behind the statistics, there is often a hidden disagreement about how to measure discrimination from the outcomes of these experiments. Two positions can be discerned:

(1) *Discrimination between two groups occurs when the group distinguishing characteristics of their members give rise to inequality in their respective opportunities.* Discrimination is then measured by investigating differences in success between two or more groups, e.g. hiring probabilities for job seekers (see a.o. Bertrand and Mullainathan, 2004, Lahey, 2008, Oreopoulos, 2009, Moreno *et al.*, 2012)[3];

(2) *Discrimination is difference in treatment of persons because of their difference in some personal characteristics not deemed to be a justifiable basis for selection.* This is the idea behind the *net degree of discrimination* (see e.g. Riach and Rich, 2002, 2006, 2010).[4] This measure is defined as the proportion of the difference in the number of cases in which the person belonging to the majority group is favoured (has success, while the minority group person has not) and the number of cases in which the minority group is favoured, in the total number of cases in which at least one person has success. The

[3] To be fair, it should be mentioned that Bertrand and Mullainathan (2004) also report measures which are more in line with the second concept.

[4] This measure is also favoured by the International Labour Office (Bovenkerk, 1992).

conditional sign test proposed by Heckman and Siegelman (1993), is a measure for the degree to which the differences in cases of asymmetric success (one has success while the other has not) is statistically significant.

This paper has as objective to extend these concepts to a multidimensional context. A multidimensional context occurs when people differ in more than one aspect (e.g. sex, age, national origin or physical state), as is usually the case in the real world. This poses as an additional problem to fix a method to determine the contribution of each of these dimensions separately, to the unfavourable treatment a person may undergo. Which part *e.g.* is due to age, and which to sex when persons differ in only those two characteristics? We propose a measure embodying the second concept of discrimination mentioned above, based on an interpretation of the regression coefficients of an ordered trinomial discrete choice model. We ran a correspondence experiment in order to implement this measure. The results are compared with those of a discrimination measure based on a binary discrete choice model, which is in line with the first concept of discrimination.

2 A field experiment

In order to implement our proposal empirically, we conducted a field experiment. From the end of July 2010 to the beginning of June 2011 we selected 854 publicly announced vacancies on the Belgian labour market, evenly spread across the three regions of the country (Flanders, Wallonia and Brussels), four economic sectors (industry and manufacturing; commerce, transport and catering; communication, administration and financial services; public sector, health care, non-profit and other services) and four professions (skilled and unskilled blue collars, routine and executive white collars). Since we wanted to study the behaviour of employers, we did not take into account vacancies which are treated by interim or selection agencies. Because of the quantitatively and economically poor importance of the primary sector for the Belgian labour market, vacancies in this sector were excluded too. To each vacancy, two resumes were sent. So, in total, there where 1708 applications. The type of one of the applicants was kept fixed across vacancies: a 35 years old male of Belgian origin. It will henceforth be labeled as the reference type. The other applicant differed from this reference type in at least one of the following dimensions:

(1) sex: some of the applicants were female;

(2) age: in addition to applicants of 35 years, some of the job candidates were of age 23, 47,or 53;

(3) national origin or physical limitation: next to persons from Belgian origin, we also had candidates that indicate (by mentioning birthplace in the CV in combination with a foreign name) to be of Italian, Congolese, Turkish or Moroccan origin. For the last group we distinguished between candidates with Belgian nationality and persons with Moroccan nationality. Other candidates not born in Belgium all had Belgian nationality (mentioned in the CV). Some applicants, born in Belgium and with Belgian sounding name, indicated in the CV that they suffered from a physical limitation (not specified), that did however not hamper their independence.

Combining these characteristics, we create 55 types of applicants, in addition to the reference type (2 sexes times 4 age groups times 6 possible national origins (=48) *plus* 2 sexes times 4 age groups of persons signalling a physical limitation (=8)). The following two types were added:

(4) pregnancy: some female applicants (born in Belgium and with Belgian name) indicated in the letter of application that they were immediately available, but pregnant since three months. These types of applicants belonged to the youngest two age groups (23 and 35 years old).

So, we have in total 58 types of applicants. For each of the vacancies, the second applicant was selected at random from the list of 57 non reference types.

We always matched the applicants' skill level with the requirements in the vacancy. All applicants had a complete job history, and a faultless school career (all obtained at least a certificate of the secondary school degree). Irrespective of birthplace or nationality, all applicants obtained their education in Belgium.

Small differences in the CV, such as template, font, lay-out, the education institutions at which the (required) certificates were obtained and the job history and address of the applicants, were randomly alternated between the reference type and the other applicant. Distance between residence and future workplace were kept more or less equal for both candidates, and, if necessary, differences were alternated.

We registered whether the applicants were selected for the next step in the hiring procedure (usually an invitation for an oral interview). After a reaction of the employer (either a positive reaction as just defined or a request for further information, or a negative reaction) or when no answer was obtained within two months after the application, the application was closed. So, all applications were formally closed. In this way we limit the burden of this involvement without consent of the selection responsibles into experimental research to a minimum and meet the ethical standards usually put on this approach (see Riach and Rich, 2004b). The distribution of the vacancies and the (positive) response rate by region, sector and profession can be found in Appendix I. In Appendix I the number of applications and the relative selection frequencies by sex, age and national origin/physical state is also given. In the next section we explain how we analysed the hiring decision in order to derive measures of discrimination.

3 Modelling the selection behaviour of employers

We test whether the selection behaviour of an employer thus registered, can be approached by a latent variable model. It is assumed that the selection manager implicitly gives a score to each candidate. The score reflects the potential of a candidate to be selected, also called the *selection propensity* in the sequel. This score on the potential to be selected can be influenced by some observable characteristics of the candidate, such as sex, age, national origin and particular physical states. The value of these characteristics for a particular job candidate i is denoted by $\mathbf{x}_i$. The impact of the characteristics on the score is denoted by $\boldsymbol{\beta}$. The deterministic part of the sore equals therefore $\boldsymbol{\beta}'\mathbf{x}_i$. Random deviations from the deterministic part of the score attributed to candidate i by selection responsible j are denoted by ε_{ij}. So, the unobserved score of candidate i when applying for a job with employer j, say

y_{ij}^* equals:

$$y_{ij}^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_{ij}. \tag{1}$$

There are two possibilities to convert this unobservable score into observable behaviour, which coincide with the two rivalling views on measuring discrimination mentioned in the introduction. The first behavioural rule says that candidates surpassing a score equal to zero, will be selected. Formally, the candidate i applying for job j will be selected (say $y_{ij} = 1$) if and only if $y_{ij}^* \geq 0$, and she will not be selected ($y_{ij} = 0$) iff $y_{ij}^* < 0$:

$$\begin{aligned} y_{ij} = 1 \quad &\Longleftrightarrow \quad y_{ij}^* \geq 0, \\ y_{ij} = 0 \quad &\Longleftrightarrow \quad y_{ij}^* < 0. \end{aligned} \tag{2}$$

Hence, the probability that a candidate i will be selected ($P\left(y_{ij} = 1\right)$) is equal to:

$$P\left(y_{ij} = 1\right) = P\left(y_{ij}^* \geq 0\right) = 1 - P\left(\varepsilon_{ij} < -\boldsymbol{\beta}'\mathbf{x}_i\right). \tag{3}$$

The second approach interprets the latent variable as the job candidate's capacity to be preferred for selection by the employer above other candidates, also called *discrimination susceptibility* in the sequel. The higher the score of the candidate on this variable, the higher the probability that this candidate will be selected *at cost* of the reference type. We will call this the discriminative capability of the candidate. Reversely, discrimination vulnerability is the liability of this particular candidate to be set aside in advantage of the reference type. Discrimination vulnerability and discriminative capability are *individual* characteristics. If a candidate differing from the reference type is selected while the reference type is not selected, the reference type is considered to have a high degree of discrimination vulnerability and the other candidate has a high discriminative capability. There are assumed to be two thresholds for the score. When a candidate's score is below the lower threshold, say $\underline{y}^*$, she will not be selected, while the reference candidate would be. Reversely, if the score of the candidate surpasses the upper bound $\bar{y}^*$ (with $\bar{y}^* > \underline{y}^*$), she will be selected while the reference candidate will not be selected. Hence, in that case, the reference candidate will have a score on the discriminative capability measure below the lower threshold $\underline{y}^*$. Let us denote for each candidate i for job j, differing from a reference type r, three possible events:

– the candidate is selected while the reference type is not selected, denoted as $z_{ij} = 1$;

– both, the candidate and the reference type, are selected, or none of them is selected, denoted as $z_{ij} = 2$;

– the candidate is not selected, while the reference type is selected, denoted as $z_{ij} = 3$.

We call these events the occurrence of a discrimination advantage ($z_{ij} = 1$), a discrimination disadvantage ($z_{ij} = 3$), and the discrimination neutral state or equal treatment case ($z_{ij} = 2$). Consequently, for the reference type we have:

– If $z_{ij} = 1$ then $z_{rj} = 3$;

– If $z_{ij} = 2$ then $z_{rj} = 2$;

– If $z_{ij} = 3$ then $z_{rj} = 1$.

The scores on the latent variable associated with these events are as follows:

$$\begin{aligned} z_{ij} = 1 \text{ and } z_{rj} = 3 \quad &\Longleftrightarrow \quad y^*_{ij} > \bar{y}^* > \underline{y}^* > y^*_{rj}, \\ z_{ij} = 2 \text{ and } z_{rj} = 2 \quad &\Longleftrightarrow \quad \bar{y}^* \geq y^*_{ij}, y^*_{rj} \geq \underline{y}^*, \\ z_{ij} = 3 \text{ and } z_{rj} = 1 \quad &\Longleftrightarrow \quad y^*_{rj} > \bar{y}^* > \underline{y}^* > y^*_{ij}. \end{aligned} \tag{4}$$

We stick however to the interpretation of the latent variable as a purely individual characteristic. It means that possessing characteristics which increase the discrimination capacity, will increase the probability that this person will be selected from a randomly drawn pool of other candidates. Of course, if some of the rivals also possess (possibly other) characteristics that make her a salient candidate for selection, maybe both candidates would be selected in that case. The more frequently it is observed that a candidate with particular characteristics is not selected alone, the less this characteristic is considered to increase the probability to obtain a discrimination advantage. If such individual characteristics fostering discrimination advantage or disadvantage would not exist, the model would empirically be rejected.

Hence, the probabilities to observe one of the three possible events for a candidate i for vacancy j, equal:

$$\begin{aligned}
P(z_{ij}=1) &= 1-P\left(\varepsilon_{ij} \leq \bar{y}^* - \boldsymbol{\beta}'\mathbf{x}_i\right), \\
P(z_{ij}=2) &= P\left(\bar{y}^* - \boldsymbol{\beta}'\mathbf{x}_i > \varepsilon_{ij} > \underline{y}^* - \boldsymbol{\beta}'\mathbf{x}_i\right), \\
P(z_{ij}=3) &= P\left(\varepsilon_{ij} \leq \underline{y}^* - \boldsymbol{\beta}'\mathbf{x}_i\right).
\end{aligned} \tag{5}$$

Depending on the assumptions on the distribution of ε_{ij} (logistic or standard normal) both models can be estimated by a logit or probit model. We will use the probit approach in the sequel: the ε_{ij}'s are assumed to be standard normally distributed. For the second approach (interpreting the latent variable as discriminative susceptibility), the appropriate model is an ordered (or cumulative) one.

Let $\Phi(z)$ be the standard normal distribution function, and $\varphi(z)$ the associated density function. Then, for the probit model, the binary model (3) reduces to:

$$\begin{aligned}
P(y_{ij}=1) &= 1-P\left(\varepsilon_{ij} < -\boldsymbol{\beta}'\mathbf{x}_i\right) &&= 1-\Phi\left(-\boldsymbol{\beta}'\mathbf{x}_i\right) = \Phi\left(\boldsymbol{\beta}'\mathbf{x}_i\right), \\
P(y_{ij}=0) &= P\left(\varepsilon_{ij} < -\boldsymbol{\beta}'\mathbf{x}_i\right) &&= \Phi\left(-\boldsymbol{\beta}'\mathbf{x}_i\right),
\end{aligned} \tag{6}$$

where, for the last step in the first equality, use has been made of the symmetry around zero of the standard normal distribution function ($\varphi(z) = \varphi(-z)$ for all $z \in \mathbb{R}$, implying that $\Phi(z) = 1-\Phi(-z)$, for all $z \in \mathbb{R}$).

To formulate the probit version of the ordered model (5), an identification issue has to be settled first. Usually the vector of covariates $\mathbf{x}_i$ includes a constant term equal to 1 with associated coefficient β_1 say. In model (5), this coefficient always occurs in an linear combination with $\underline{y}^*$ and $\bar{y}^*$. Therefore, β_1 will be normalised to zero, and only coefficients $\widetilde{\boldsymbol{\beta}}$ for the non-constant covariates, denoted by $\widetilde{\mathbf{x}}_i$, will be estimated in this model. The thus normalised boundaries are denoted as $\underline{\mu}^*$ and $\bar{\mu}^*$. Given this normalisation, we can now formulate the probit version of model (5) as:

$$\begin{aligned}
P(z_{ij}=1) &= 1-P\left(\varepsilon_{ij} \leq \bar{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right) &&= 1-\Phi\left(\bar{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right), \\
& &&= \Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i - \bar{\mu}^*\right) \\
P(z_{ij}=2) &= P\left(\bar{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i > \varepsilon_{ij} > \underline{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right) &&= \Phi\left(\bar{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right) - \Phi\left(\underline{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right) \\
P(z_{ij}=3) &= P\left(\varepsilon_{ij} \leq \underline{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right) &&= \Phi\left(\underline{\mu}^* - \widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right),
\end{aligned} \tag{7}$$

where again symmetry of Φ has been used in the first set of equations.

The likelihood function for the binary probit model is equal to:

$$\mathcal{L}_{bin}\left(\boldsymbol{\beta}\,|\mathbf{y},\mathbf{X}\right)=\prod_{ij\in\mathcal{N}}\Phi\left(\boldsymbol{\beta}'\mathbf{x}_i\right)^{y_{ij}}\cdot\Phi\left(-\boldsymbol{\beta}'\mathbf{x}_i\right)^{(1-y_{ij})}, \tag{8}$$

where $\mathbf{y}$ is the vector of binary values y_{ij}, and $\mathbf{X}$ is the matrix of covariates, with typical row $\mathbf{x}_i'$, and $\mathcal{N}$ is the set of indices ij for which we have observations in our data set.

Let $\chi\,(\texttt{condition})$ be the indicator function which is equal to one if the condition stated in its argument is true, and zero otherwise. So, *e.g.* $\chi\,(z_{ij}=k)=1$ if $z_{ij}=k$, and $\chi\,(z_{ij}=k)=0$ otherwise, for $k=1,2,3$. The likelihood function for the ordered probit model is then:

$$\begin{aligned}&\mathcal{L}_{ord}\left(\widetilde{\boldsymbol{\beta}},\bar{\mu}^*,\underline{\mu}^*\,|\mathbf{z},\mathbf{X}\right)=\\&\prod_{ij\in\mathcal{N}}\Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i-\bar{\mu}^*\right)^{\chi\left(z_{ij}=1\right)}\cdot\left(\Phi\left(\bar{\mu}^*-\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right)-\Phi\left(\underline{\mu}^*-\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right)\right)^{\chi\left(z_{ij}=2\right)}\cdot\left(\Phi\left(\underline{\mu}^*-\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}_i\right)\right)^{\chi\left(z_{ij}=3\right)}.\end{aligned} \tag{9}$$

We estimated a basic version of this model using as covariates dummies for sex, age, national origin, signalling a physical limitation and mentioning pregnancy.[5] The reference *type* also served as the reference *category* in our estimations. On the basis of these estimates, different measures of discrimination can be constructed. Most straightforward candidates, for this purpose, are differences, or ratio's or odds ratio's of, or marginal effects on the estimated probabilities. One could, for example, calculate the differences in estimated probabilities to be selected between males and females, or evaluate the ratio of the male's to female's odds between falling into the category obtaining a discrimination advantage and not falling into that category. However, the difficulty with all these measures is that their magnitude depends on the complete vector of characteristics (the values for sex, age, national origin, signalling a physical limitation and mentioning pregnancy of the type for which the measure is evaluated). For example, (the impact on) the probability to be selected of being female (as compared to male) depends on the age category, the national origin, having a physical malfunctioning and being pregnant or not. Alternatively, (the impact on) on the probability to obtain a discrimination advantage, of belonging to a certain age category depends on the type's sex, the national origin, exhibiting a physical malfunctioning and being pregnant or not. This can be seen formally as follows. We consider the comparison of the impact of sex and age on the probability to fall into the category with discrimination advantage. This impact comparison is a difference–in–difference measure: the difference in probability between males and females is compared with the difference in probability for persons of age y_1 *versus* persons of age o_1.

[5] We come back to extensions of this basic version of the model at the end of Section 5.

Denote the coefficient of the dummy for females as β_f (male is the reference category for the sex dimension), and the coefficients for the dummies of the age categories y_1 and o_1, are, respectively, β_{y_1} and β_{o_1} (suppose that another category, y_o say, is the reference category for age). Furthermore, the notation $\boldsymbol{\beta}_{-l}$ denotes t he vector of coefficients after removing the effect of covariate l, β_l. Similarly, $\mathbf{x}_{i,-l}$ is the vector of type i's covariate values (dummy values) after deleting the l–th coordinate. With this notation, the comparison between difference in sex and difference in the selected age categories is equal to:

$$\begin{array}{rcl}\Delta\left(\beta_f;\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}-\bar{\mu}^*\right) & = & \Phi\left(\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}+\beta_f-\bar{\mu}^*\right)-\Phi\left(\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}-\bar{\mu}^*\right) \\ \Delta\left(\beta_{y_1},\beta_{o_1};\widetilde{\boldsymbol{\beta}}'_{-y_1,-o_1}\widetilde{\mathbf{x}}_{i,-y_1,-o_1}-\bar{\mu}^*\right) & = & \\ \multicolumn{3}{r}{\Phi\left(\widetilde{\boldsymbol{\beta}}'_{-y_1,-o_1}\widetilde{\mathbf{x}}_{i,-y_1,-o_1}+\beta_{o_1}-\bar{\mu}^*\right)-\Phi\left(\widetilde{\boldsymbol{\beta}}'_{-y_1,-o_1}\widetilde{\mathbf{x}}_{i,-y_1,-o_1}+\beta_{y_1}-\bar{\mu}^*\right).}\end{array} \tag{10}$$

In general, the sign of the difference between both differences, $\Delta\left(\beta_f;\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}-\bar{\mu}^*\right)-\Delta\left(\beta_{y_1},\beta_{o_1};\widetilde{\boldsymbol{\beta}}'_{-y_1,-o_1}\widetilde{\mathbf{x}}_{i,-y_1,-o_1}-\bar{\mu}^*\right)$, can depend on the particular value of the dummies for i on sex and the age categories y_1 and o_1. In words, the effect of sex for a young persons, say belonging to category y_1 on the probability to obtain a discrimination advantage, is different from that of sex on relatively older persons, say belonging to category o_1, so that the difference with the effect calculated for age, can change. However limiting comparisons to any given value of $\widetilde{\mathbf{x}}$, *the sign of the difference* between $\Delta\left(\beta_f;\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}-\bar{\mu}^*\right)=\Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}+\beta_f-\bar{\mu}^*\right)-\Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}-\bar{\mu}^*\right)$ and $\Delta\left(\beta_{o_1};\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}-\bar{\mu}^*\right)=\Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}+\beta_{o_1}-\bar{\mu}^*\right)-\Phi\left(\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}-\bar{\mu}^*\right)$ is independent of $\widetilde{\boldsymbol{\beta}}'\widetilde{\mathbf{x}}-\bar{\mu}^*$. It depends solely on the sign of the difference between between β_f and β_{o_1}. *E.g.*, the effect of a switch in sex on the probability to obtain a discrimination advantage for a male belonging to age category y_1 can unambiguously be signed as compared to the effect, for that same person of switching to age category o_1, and this independent of the national origin or physical state of the type. And, for such a comparison, looking at the difference between the coefficients suffices. Therefore the following claim can be endorsed.

Claim 1. *The magnitude of the regression coefficients in the discrete choice models are a measure for the relative intensity of discrimination with respect to particular characteristics across (possibly different) dimensions.*

The proposal in Claim 1 is intended to serve as a measure for comparing the intensity of discrimination for particular sets of values a type can obtain in possibly different dimensions, such as there are, in our application, sex, age, and national origin or special physical features of the applicant. For example, it is possible to compare the degree of discrimination between the Turkish and Belgian national origin of an applicant, with that between sexes. But one could also compare the intensity of discrimination between different sets of values within one particular dimension, for example, between the age of 53 *vis-à-vis* 35 (measured by β_{53}), with that of 23 *vis-à-vis* 47 (measured by $\beta_{23} - \beta_{47}$).

The measure is about the aggregate selection behaviour of employers, or, in particular, their selection responsibles. It can, for example, be investigated whether there is evidence that this selection behaviour is, on the average, influenced by one of the following factors as a whole: sex, age, national origin, or each of the two special physical conditions (pregnancy and physical limitation). To that purpose, we tested, for each of these five factors in turn, the joint significance of the coefficient(s) of all the possible values referring to that factor (*e.g.* the three coefficients for age). For this reason also, we explicitly opted for not taking up possible cross effects into the regression models. We want to test whether the employer takes into account each of these factors *per se*, and not whether there are specific constellations of individual characteristics which make a person vulnerable to discriminative behaviour. If there is no clear evidence that one or more of these factors play a separate role, we reject the hypothesis that, on average, Belgian employers take this factor into account for the first selection of job applicants.

A negative impact of a type's particular characteristic l, on the probability to obtain a discrimination advantage (negative β_l), implies a positive influence of that same characteristic on the probability to undergo a discrimination disadvantage, as should be clear from inspecting the first and third equality in the set of equations (5) or (7). The direction of the effect on the probability to fall in the neutral category (both selected or none of both selected) is not clear however from the sign of the coefficient. As can be seen from the middle equality in the same set of equations, the effect of, for example, being female, on the probability to fall in the middle category, equals:

$$\begin{aligned}\Delta_2\left(\beta_f;\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f},\bar{\mu}^*,\underline{\mu}^*\right) \quad = \quad & \Phi\left(\bar{\mu}^*-\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}-\beta_f\right)-\Phi\left(\underline{\mu}^*-\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}-\beta_f\right)- \\ & \left(\Phi\left(\bar{\mu}^*-\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}\right)-\Phi\left(\underline{\mu}^*-\widetilde{\boldsymbol{\beta}}'_{-f}\widetilde{\mathbf{x}}_{i,-f}\right)\right).\end{aligned} \tag{11}$$

This is again a difference–in–difference measure, and the expression cannot be signed solely on the basis of the sign of the coefficient β_f. As a measure of absence of discrimination, the effect of particular characteristics on the probability to fall in the neutral category, can therefore not be derived from the sign (or magnitude) of its coefficient. We therefore do not explore this measure further in the present contribution. It suffices to say that in our application, it turned out to be the case that the higher the values of the coefficients were, the bigger the impact of the corresponding characteristic on the probability to fall into the neutral category.

The direction of a characteristic's effect on the cumulatives $P(z_{ij} = 1) + P(z_{ij} = 2)$ and $P(z_{ij} = 3) + P(z_{ij} = 2)$, can be derived from the coefficients' sign. Indeed, from the set of equations (7), it can be seen that a positive (negative) coefficient implies an increasing (decreasing) effect of the associated covariate on $P(z_{ij} = 1) + P(z_{ij} = 2)$, and a decreasing (increasing) effect on $P(z_{ij} = 3) + P(z_{ij} = 2)$.

4 Empirical results

In line with the two proposals to measure discrimination in the previous section, we report in Table 1 the results of our basic regression analyses. The left hand panel gives the results for the binary probit model. A negative (positive) coefficient is to be interpreted as a decreasing (increasing) effect of the variable on the probability to be selected. The right panel gives the result for the ordered probit. A negative coefficient means a decreasing impact on the probability to obtain a discrimination advantage and a positive influence on the probability to experience a discrimination disadvantage.

In line with the discussion at the end of the previous section (Claim 1) we also add in the second column of each panel the rank of the order of magnitude of the coefficients from smallest (most discriminated) to highest (least discriminated or, if positive, most favoured).

In terms of inequality of opportunities to be selected (the binary model), discrimination is strongest towards the oldest age category considered, followed, on almost equal footing, by a Moroccan origin and nationality and the second oldest age category. The next most discriminating characteristics are a Turkish origin and sex (against females). The effect of the other factors is almost nil, and there is poor statistical evidence that these factors did influence the probability to be selected.

The picture changes however when measuring discrimination in terms of unequal treatment of equals (one candidate being selected, the other not, captured by the ordered probit model). Discrimination in these terms is strongest against candidates from Italian and Congolese origin, followed by the two older age categories. Next in ranking is a Moroccan origin (irrespective of nationality). Still substantially negative is the effect of being pregnant, a Turkish origin and exhibiting a physical limitation (in this order), though the statistical evidence is poorer for rejecting the null hypothesis of no discrimination for these groups. The effect of sex almost vanishes according to this measure.

Finally, according to both measures, the difference between 35–years old and 23–years old is negligible.

Empirically, this change in results could be understood from the split-up of the relative frequencies of selected candidates in two categories: those who are invited alone, and those invited together with the other (see Table VII in Appendix II). Persons from Turkish and Moroccan origin are selected less often than those from Italian and Congolese origin. But, when they are selected, it occurs relatively more frequently that the other person (the reference type) was not selected. On the other hand, candidates from Italian and Congolese origin are relatively more frequently invited together with the reference type.

Table 1: Probit models

	Probability to be selected					Probability to obtain discrimination advantage or undergo disadvantage				
Parameter	Estimate	Rank	Std Err	χ^2-value	Pr > χ^2	Estimate	Rank	Std Err	χ^2-value	Pr > χ^2
c $(-\bar{\mu}^*)$	-0.6274	N.A.[a]	0.0448	196.2837	<.0001	-1.0700	N.A.	0.0475	507.1787	<.0001
c_2 $(-\underline{\mu}^*)$	N.A.	N.A.	N.A.	N.A.	N.A.	1.5382	N.A.	0.0560	754.5412	<.0001
sex-f	-0.1103	5	0.1021	1.1672	0.2800	-0.0718	10	0.0886	0.6562	0.4179
age-23	-0.0161	10	0.1189	0.0184	0.8920	-0.0434	11	0.1086	0.1601	0.6891
age-47	-0.2337	3	0.1346	3.0138	0.0826	-0.3369	4	0.1176	8.2125	0.0042
age-53	-0.3960	1	0.1400	8.0035	0.0047	-0.3547	3	0.1169	9.1995	0.0024
nat-morobelg	-0.0856	6	0.1638	0.2729	0.6014	-0.2635	6	0.1446	3.3190	0.0685
nat-moromoro	-0.2393	2	0.1683	2.0217	0.1551	-0.2824	5	0.1426	3.9227	0.0476
nat-cong	-0.0357	8	0.1620	0.0487	0.8253	-0.4104	2	0.1444	8.0773	0.0045
nat-turk	-0.1629	4	0.1694	0.9254	0.3361	-0.2173	8	0.1459	2.2189	0.1363
nat-ital	-0.0318	9	0.1613	0.0388	0.8438	-0.4145	1	0.1440	8.2821	0.0040
phys. lim.	0.0131	12	0.1580	0.0069	0.9338	-0.2015	9	0.1427	1.9933	0.1580
pregnant	-0.0675	7	0.1818	0.1379	0.7104	-0.2521	7	0.1635	2.3787	0.1230
	Value of Objective Function = -895.895					Value of Objective Function = -1083.97				
Model tests	LR 28.75		DF 11	p-value 0.0025		LR 71.53		DF 11	p-value <.0001	
	Wald 27.17		DF 11	p-value 0.0043		Wald 72.36		DF 11	p-value <.0001	

[a] Not applicable

Also, in terms of precision of the coefficients, the second model performs better: on average the standard errors of the model decrease by 12%. Similarly, in terms of the global model fit, the likelihood ratio and Wald–test statistics of the second model outperform those of the first (see the last two lines of Table 1).[6]

In order to test for discrimination along the aggregate dimensions sex, age, national origin, physical limitation and pregnancy, we performed also LR– and Wald–tests for each of those dimensions (see Table 2). In terms of the first model we could only reject absence of discrimination in terms of age. According to the second model, there is strong evidence for both, discrimination according to age and national origin and weak evidence for discrimination against persons signalling pregnancy or a physical limitation. According to the first model, there is only weak, and in the second model almost no evidence for discrimination against females.

Both measures give different answers to the question of discrimination. Though this can empirically be understood from an inspection of the raw data, the need remains to judge the relative merits of both approaches for measuring discrimination. This is the subject of the next section.

5 Critiques on the inequality of opportunity approach

The binary discrete choice model (3) conceives discrimination as a matter of inequality of opportunities. In our application of this model, these opportunities are measured by probabilities to be selected, and unequal selection probabilities between equally valid candidates is then considered as evidence for the occurrence of discrimination. We claim that this conception of

[6] The Likelihood Ratio (LR)-test is equal to two times the difference of the value of the log of the likelihood function of the full model and a model with only the constant term(s). *E.g.* for the binary model (with β_1 the coefficient associated with respect to the constant term): $2 \cdot \left(\ln\left(\mathcal{L}\left(\widehat{\boldsymbol{\beta}};\mathbf{y},\mathbf{X}\right)\right) - \ln\left(\mathcal{L}\left(\widehat{\beta}_1;\mathbf{y}\right)\right)\right)$. The Wald–test statistic for a subset of coefficients is defined as $\widehat{\boldsymbol{\beta}}_s' V^{-1}\left(\widehat{\boldsymbol{\beta}}_s\right)\widehat{\boldsymbol{\beta}}_s$, where $V\left(\widehat{\boldsymbol{\beta}}_s'\right)$ is the variance-covariance matrix of the subset of estimated coefficients $\widehat{\boldsymbol{\beta}}_s$. Some care should be taken comparing these figures. Though the explanatory variables and functional form of the latent variable equation behind both models are equal, the assumptions about the nature of the information provided by the disturbance terms differs, so that these models are not really nested.

Table 2: Model tests

	Probability to be selected			Probability to obtain advantage or to undergo disadvantage		
Test	Sex					
	χ^2-value	DF	p-value	χ^2-value	DF	p-value
LR	1.168	1	0.2798	0.655	1	0.4183
Wald	1.167	1	0.2800	0.656	1	0.4179
Test	Age					
	χ^2-value	DF	p-value	χ^2-value	DF	p-value
LR	10.114	3	0.0176	14.254	3	0.0026
Wald	9.969	3	0.0188	14.178	3	0.0027
Test	National origin					
	χ^2-value	DF	p-value	χ^2-value	DF	p-value
LR	2.637	5	0.7557	13.582	5	0.0185
Wald	2.591	5	0.7627	13.480	5	0.0193
Test	Physical limitation					
	χ^2-value	DF	p-value	χ^2-value	DF	p-value
LR	0.007	1	0.9333	1.959	1	0.1616
Wald	0.007	1	0.9338	1.993	1	0.1580
Test	Pregnancy					
	χ^2-value	DF	p-value	χ^2-value	DF	p-value
LR	0.138	1	0.7103	2.411	1	0.1205
Wald	0.1379	1	0.7104	2.379	1	0.1230

Table 3: Equality of selection probabilities does not imply absence of discrimination

Types	Number of cases			
	Only A selected	Only other (B or C) selected	Both selected	None of both selected
A & B	5	0	20	55
A & C	0	5	15	60
	Number of selections			
Type	Discr. adv. selected alone	Discr. disadv. other sel., cand. not	Equal Treatment both sel. or none of both	Selection ratio
A	5	5	150	40/160
B	0	5	75	20/80
C	5	0	75	20/80

discrimination is flawed.

Claim 2. *Equality of selection probabilities between types is neither a sufficient nor a necessary condition for absence of discrimination.*

The sufficiency part of the claim (equality of selection probabilities does not imply absence of discrimination) can be sustained by the fictitious example in Table 3. There are 160 vacancies and three candidates, differing in one, and only one, respect. Candidate A applies to all vacancies, candidate B only to 80 of the 160 vacancies and candidate C applies to the other 80 vacancies. In 5 of the 80 cases in which A and B apply, A is selected, while B is not, in another 20 cases both are selected, and in the last 55 cases, none of both is selected. In the 80 cases in which A and C apply, C is selected 5 times, without A being selected. In another 15 cases, both are selected. In the remaining 60 cases, none of them is invited. Assume that we are informed that the asymmetric selection cases, are undoubtedly the result of discriminatory motivated behaviour. Clearly, the aggregate selection probability measure by type could not capture this event. Indeed, all three persons are invited in 25% of the cases (see the last column of the lower panel in Table 3).

The necessity part of the claim (equal invitation probabilities are not necessary for absence of discrimination) is a consequence of possible differences in selection probabilities across vacancies. Indeed, if, for instance, candidates are more likely to be selected for certain vacancies, *irrespective of their characteristics*, and less likely for others, then we could possibly agree that there is no discrimination in the selection behaviour of the employers. The urgency to fill in the vacancy can *e.g.* be higher in one firm than in another, and therefore they select more candidates. But, if the distribution of applicant types across these vacancies is unequal, they will end up with different aggregate invitation probabilities. One could argue that in an experimental set-up, one can control the distribution of types across vacancies to be equally. However, vacancy specific characteristics cannot be controlled for in an experiment. This is an instance of pairwise heterogeneity which could bias estimated results in such a set-up, as Heckman and Siegelman (1993) warned. We took this issue up by running a specification with vacancy specific explanatory variables in the binary model, following a suggestion by Yinger in his comments on Heckman and Siegelman (1993), and by estimating a bivariate probit model with vacancy specific correlation. The results of these regressions are available on request. They did not change the conclusions and were less reliable than those of the simple model.

The problem is however not only a statistical one. When discrimination is considered to be a distinction made between persons which should be treated equally, then the cases in Table 3 where one, and only one, of two equally valid applicants were selected, are potential indicators of the occurrence of discrimination in this sense, while cases in which both are selected (or none of both) are not. In our proposal to analyse experimental data on discrimination by means of an ordered discrete choice model, this last group of cases is considered as evidence for absence of discrimination. This might be questioned. But in any way, selecting both equally valid candidates should not be treated on an equal footing as selecting only one of these candidates, in a measure of discrimination, as they tend to be by the inequality of opportunity approach.

6 Conclusions

Two notions of discrimination can be extracted from the results of field experiments. The first defines the absence of discrimination between two groups of persons as equality of opportunities for persons belonging to both of these groups. The second one identifies discrimination as the unequal treatment of persons belonging to different groups.

We propose a method to measure both notions, when people differ in more than one dimension in which discrimination can occur, e.g. sex, age and national origin or physical state. The additional difficulty of a multidimensional approach, is how to ascribe observed differences in opportunities or treatment, to each of these dimensions. This should be done in a way, such that statements can be made about whether discrimination, if any, according to,say, age is larger or smaller than discrimination according to, say, national origin.

We argue that the magnitude of the coefficients in a discrete choice model can serve that purpose. We consider discrimination by sex to be larger than discrimination by age if the coefficient of the sex dummy is larger (or smaller, depending on the reference category of the model, success or failure) than the coefficient of the relevant age group. Moreover, we have shown how the first concept of discrimination could be measured by the coefficients of a binary discrete choice model. The appropriate model for the second concept in a multidimensional context is an ordered trivariate discrete choice model: an agent obtains a discrimination advantage if she has success, while the other candidate, differing only in one or more of the distinguished dimensions, fails. Equal treatment occurs if both have success or fail. She undergoes a discrimination disadvantage if the other has success, while she has not.

We argue that the first concept of discrimination fails to detect discrimination in some cases in which there is, presumably, consensus about its occurrence. Reversely, in some cases this concept leads to the conclusion that there is discrimination, while many would agree that the differences measured are the consequence of other factors than discrimination.

We set up an experiment in which we sent out 1708 resumes to vacancies publicly announced on the Belgian labour market between July 2010 and June 2011. Selected vacancies were equally distributed across the three regions of the country, four large sector classes of economic activities of the firms that opened the job place, and four groups of professions. The candidates were equal in all respects, except for sex (male and female), age (23–, 35–, 47– and 53–years old), national origin (Belgian, Italian, Congolese, Moroccan and Turkish) or physical state (signalling a physical limitation or pregnancy). Their educational qualities and skills were matched to the requirements posed by the vacancy. Unavoidable small remaining differences between resumes (typeface, lay-out, ...), necessary to make the fake candidacies credible, were randomized across types.

Both measures agreed in the conclusion that age is by far out the most important factor of discrimination (in favour of the younger groups, and against older groups). However, as far as the contribution of national origin is concerned, both measures give a substantially different picture. In terms of selection propensities, Moroccans (without Belgian nationality) and persons from Turkish origin are substantially worse off. But an Italian or Congolese background does hardly make any difference in this respect (compared to persons from Belgian origin). Globally, the contribution of national origin to differences in selection probabilities is not significant. In terms of unequal treatment of equals, however, the susceptibility of persons of Italian and Congolese origin to undergo a discrimination disadvantage, is substantially higher, and even surpasses the effect of older age slightly. Also, the small and insignificant negative effect of pregnancy according to the first measure, becomes more substantial and more reliable. The slightly positive (and insignificant) effect of signalling a physical limitation in terms of selection probabilities, turns into a negative (and more reliable) effect on discrimination vulnerability. On the contrary, the negative (but rather unreliable) effect of being female on selection probabilities almost vanishes when looking at its contribution to asymmetric treatment.

It should be stressed that in our opinion, discrimination (in the selection of job candidates) is a trait of the behaviour of employers (or their selection responsibles). Therefore, this study should not be read as an attempt to measure the real opportunities of the different groups we analysed, in society. We only aimed at proposing a measure for differences in treatment due to the discriminative behaviour of employers (or their selection responsibles). As far as the proposed measure is accepted, it follows from our study, that the lower scores of older cohorts on labour market participation, as far as it is age related, cannot be removed by activation policies on the labour supply side alone. There is room for intervention at the labour demand side too. For the purpose of illustration: the probability to undergo a discrimination disadvantage due to an increase in age from 35–years to 47– or 53– years old increases by *70* *to 75* percent, according to our study.

References

[1] Bertrand M. & S. Mullainathan (2004), 'Are Emily and Greg More Employable than Lakisha an Jamal? A Field Experiment on Labor Market Discrimination,' *American Economic Review* Vol.94 No.4, 991-1013.

[2] Bovenkerk F. (1992), *Testing Discrimination in Natural Experiments: a manual for international comparative research on discrimination on the grounds of 'race' and ethnic origin.* International Labour Office, Geneva.

[3] Fix M. & R.J. Struyk (eds.) (1993), *Clear and Convincing Evidence. Measurement of discrimination in America*, The Urban Institute Press, Washington D.C.

[4] Heckman J.J. (1998), 'Detecting Discrimination,' *The Journal of Economic Perspectives* 12(2), 101-116.

[5] Heckman J.J. & P. Siegelman (1993), 'The Urban Institute Audit Studies: Their Methods and Findings,' in: M. Fix & R.J. Struyk (eds.), *Clear and Convincing Evidence. Measurement of discrimination in America*, The Urban Institute Press, Washington D.C., 187-275 (followed by discussion with J. Yinger).

[6] Lahey J.N. (2008), 'Age, Woman, and Hiring,' *Journal of Human Resources* 43 (1), 30-56.

[7] Lahey J.N. & R.A. Beasley (2008), 'Computerizing Audit Studies,' *Journal of Economic Behavior & Organization* 70, 508-514.

[8] Moreno M., H. Ñopo, J. Saavedra & M. Torero (2012), 'Detecting Gender and Racial Discrimination in Hiring Through Monitoring Intermediation Services: The Case of Selected Occupations in Metropolitan Lima, Peru,' *World Development* Vol.40(2), 315-332.

[9] Neumark D. (2011), 'Detecting Discrimination in Audit and Correspondence Studies,' IZA Discussion Paper No. 5263, Bonn.

[10] Oreopoulos P. (2011), 'Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Six Thousand Résumés,' *American Economic Journal: Economic Policy* Vol.4(3), 148-178.

[11] Riach P.A. & J. Rich (2002), 'Field Experiments of Discrimination in the Market Place?' *The Economic Journal* Vol. 112 No. 483, F480-F518.

[12] Riach P.A. & J. Rich (2004a), 'Fishing for Discrimination,' *Review of Social Economy* Vol.42(4), 465-486.

[13] Riach P.A. & J. Rich (2004b), 'Deceptive Field Experiments of Discrimination: Are they Ethical?' *Kyklos* Vol.57 Issue 3, 457-470.

[14] Riach P.A. & J. Rich (2006), 'An Experimental Investigation of Sexual Discrimination in Hiring in the English Labor Market,' *The B.E. Journal of Economic Analysis & Policy*, Vol. 6: Iss. 2 (Advances), Article 1.

[15] Riach P.A. & J. Rich (2010), 'An Experimental Investigation of Age Discrimination in the English Labor Market,' *Annals of Economics and Statistics/Annales d'Économie et de Statistique* No. 99/100, 169-185.

APPENDIX

In this appendix we give a brief description of our dataset. Appendix I contains the geographical, sectoral and professional distributions of the vacancies. We also report the distribution of the applicants across the values for the three dimensions (sex, age and national origin or physical state) in which we studied the presence of discrimination in response rates. Empirical response frequencies for these dimensions are reported in Appendix I, and the rate of occurrence of discrimination advantage, disadvantage and equal treatment in Appendix II. These figures can in no way be interpreted as the contribution of sex, age and national origin or physical state to the degree of discrimination in these dimensions. The group of females is differently composed in terms of age distribution, national origin and physical state, be it only because our option to send out a reference type to each vacancy. Moreover, the figures for the different types of discrimination susceptibility per definition do not reflect the probability to fall into the different categories, advantage, disadvantage and neutrality, for the factors that characterise the reference type (male, 35 years old and of Belgian origin). E.g. the occurrence of 13% of 35 years old persons that fall into the category with discrimination advantage, cannot be solely ascribed to the age of those persons, since, possibly, some of the non selected persons might have the same age as that person (and be of different national origin or sex). These frequencies could be purified from this, by e.g. only considering cases in which the other candidate is not of the same age. Still then, the unequal composition of a group in one dimension, with respect to the other dimensions, impedes ascribing the empirical frequencies to the different factors. Correcting for this, by reweighting, presupposes that *all* differences in the data can be ascribed to some of the the 14 values in the three dimensions that we have studied. We did not buy this assumption, and therefore propose a regression analysis in the main text, and give the raw frequencies only for descriptive purposes.

Appendix I Data

Table A.I Number of applicants (N) and relative selection frequency (P) by sex

Sex	Male	Female	All
N	1227	481	1708
P	24.37%	17.67%	22.48%

Table A.II Number of applicants (N) and relative selection frequency (P) by age

Age	23	35	47	53	All
N	240	1084	191	193	1708
P	21.67%	25.55%	16.23%	12.44%	22.48%

Table A.III Number of applicants (N) and relative selection freq. (P) by origin or physical state

Nat. Origin or Phys. State	Belg.	Phys. Lim.	Pregn.	Ital.	Cong.	Turk	Moroc. Belg.	Moroc.	All
N	948	112	101	109	109	108	109	112	1708
P	25.84%	20.54%	20.79%	19.27%	19.27%	15.74%	18.35%	14.29%	22.48%

Table A.IV Number of vacancies (V) and relative selection freq. (P) by region, sector and profession

Profession / Sector	Blue collar unskilled	Blue collar skilled	White collar	Executive	All
Region=Brussels					
1. Industry and manufacturing					
V	12	16	10	11	49
P	8.33%	25.00%	25.00%	50.00%	26.53%
2. Commerce,transp., catering					
V	18	15	14	13	60
P	11.11%	20.00%	3.57%	23.08%	14.17%
3. ICT, admin., financial serv.					
V	8	7	24	22	61
P	18.75%	21.43%	27.08%	25.00%	24.59%
4. Public, Health Care, Non profit					
V	14	15	20	26	75
P	10.71%	16.67%	27.50%	40.38%	26.67%
Total					
V	52	53	68	72	245
P	11.54%	20.75%	22.06%	34.03%	23.06%
Region=Flanders					
1. Industry and manufacturing					
V	20	26	15	16	77
P	32.50%	32.69%	30.00%	37.50%	33.12%
2. Commerce,transp., catering					
V	31	17	17	16	81
P	14.52%	14.71%	14.71%	43.75%	20.37%
3. ICT, admin., financial serv.					
V	20	18	20	22	80
P	10.00%	13.89%	40.00%	25.00%	22.50%
4. Public, Health Care, Non profit					
V	13	17	25	13	68
P	19.53%	11.76%	34.00%	30.77%	25.00%
Total					
V	84	78	72	67	306
P	18.45%	19.87%	30.52%	33.58%	25.16%
Region=Wallonia					
1. Industry and manufacturing					
V	25	35	26	16	99
P	16.00%	14.29%	10.87%	37.50%	17.68%
2. Commerce,transp., catering					
V	31	19	16	14	80
P	8.06%	13.16%	18.75%	7.14%	11.25%
3. ICT, admin., financial serv.					
V	13	10	15	18	56
P	7.96%	30.00%	23.33%	25.00%	21.43%
4. Public, Health Care, Non profit					
V	17	16	21	14	68
P	23.53%	9.38%	40.48%	42.86%	29.41%
Total					
V	86	80	75	62	303
P	13.37%	15.00%	23.33%	28.23%	19.31%

Belgium					
1. Industry and manufacturing					
V	57	77	48	43	225
P	20.18%	22.73%	19.79%	40.70%	24.89%
2. Commerce,transp., catering					
V	80	51	47	43	221
P	11.25%	15.69%	12.77%	25.38%	15.38%
3. ICT, admin., financial serv.					
V	41	35	59	62	197
P	10.98%	20.00%	30.51%	25.00%	22.84%
4. Public, Health Care, Non profit					
V	44	48	66	53	211
P	18.18%	12.50%	34.09%	38.68%	27.01%
Total					
V	222	211	220	201	854
P	14.86%	18.25%	25.45%	32.09%	22.48%

Appendix II Empirical frequencies of discrimination susceptibility categories

Table A.V Empirical absolute (N) and relative (R) frequencies of discrimination susceptibility by sex

Sex	Discrimination advantage Candidate selected, other not	Absence of discrimination Both selected	None of both selected	Discrimination disadvantage Other selected, candidate not
Male				
N	149	150	819	109
R	12.14%	12.22%	66.75%	8.88%
Female				
N	29	56	327	69
R	6.03%	11.64%	67.98%	14.35%
Total				
N	178	206	1146	178
R	10.42%	12.06%	67.10%	10.42%

Table A.VI Empirical absolute (N) and relative (R) frequencies of discrimination susceptibility by age

Age	Discrimination advantage Candidate selected, other not	Absence of discrimination Both selected	 None of both selected	Discrimination disadvantage Other selected, candidate not
23				
N	21	31	158	30
R	8.75%	12.92%	65.93%	12.5%
35				
N	145	132	724	83
R	13.38%	12.18%	66.79%	7.66%
47				
N	5	26	129	31
R	2.62%	13.61%	67.54%	16.23%
53				
N	7	17	135	34
R	3.63%	8.81%	69.95%	17.62%
Total				
N	178	206	1146	178
R	10.42%	12.06%	67.10%	10.42%

Table A.VII Empirical absolute (N) and relative (R) frequencies of discrimination susceptibility by national origin/physical state

	Discrimination advantage Candidate selected, other not	Absence of discrimination Both selected	 None of both selected	Discrimination disadvantage Other selected, candidate not
National Origin				
	Belgian			
N	132	113	639	64
R	13.92%	11.92%	67.41%	6.75%
	Congolese			
N	5	16	68	20
R	4.59%	14.68%	62.39%	18.35%
	Italian			
N	4	17	69	19
R	3.67%	15.6%	63.3%	17.43%
	Moroccan with Belgian nationality			
N	7	13	73	16
R	6.42%	11.93%	66.97	14.68%
	Moroccan with Moroccan nationality			
N	5	11	81	15
R	4.46%	9.82%	72.32%	13.39%
	Turkish			
N	9	8	74	17
R	8.33%	7.41%	68.52%	15.74%
Physical state				
	Physical limitation			
N	5	18	77	12
R	4.46%	16.07%	68.75%	10.71%
	Pregnancy			
N	11	10	65	15
R	10.89%	9.9%	64.36%	14.85%