

What is the Right Profile for Getting a Job? A Stated Choice Experiment of the Recruitment Process*

by

Stefan Eriksson[†], Per Johansson[‡] and Sophie Langenskiöld[§]

April 21, 2013

Abstract

We study the recruitment behavior of Swedish employers using data from a stated choice experiment. In the experiment, the employers are first asked to describe an employee who recently and voluntarily left the firm, and then to choose between two hypothetical applicants to invite to a job interview or to hire as a replacement for their previous employee. The two applicants differ with respect to characteristics such as gender, age, education, experience, ethnicity, religious beliefs, family situation, weight, and health, but are identical to the previous employee in all other dimensions. Our results show that employers prefer not to recruit applicants who are old, non-European, Muslim, Jewish, obese, have several children, or have a history of sickness absence. The magnitude of these effects is substantial: to eliminate the differences in hiring, wages would have to be reduced by up to 50 percent. Moreover, increasing the firms' cost of uncertainty in hiring – through more firm co-payment in the sickness benefit system – may reduce hiring, but does not seem to affect the employers' choice of whom to hire. Also, there are only small differences in recruitment behavior between different types of recruiters. Overall, our results suggest that statistical discrimination is important.

Keywords: Stated choice experiment, Discrimination, Gender, Age, Ethnicity, Obesity, Sickness absence

JEL-codes: C93, J23, J71

* We are grateful for helpful comments from Ali Ahmed, Per-Anders Edin, Erik Grönqvist, Lena Hensvik, Oskar Nordström Skans, Gerard van den Berg, and seminar participants at the Royal Economic Society 2013 Annual Conference at Royal Holloway University of London, the Conference on Field Experiments in Policy Evaluation in Nuremberg, the Nordic Summer Institute in Labour Economics in Bergen, and Uppsala University. Financial support from the Swedish Ministry of Social Affairs is gratefully acknowledged.

[†] Department of Economics, Uppsala University, PO Box 513, SE-751 20, Uppsala, Sweden, stefan.eriksson@nek.uu.se

[‡] IFAU, Department of Economics, Uppsala University and IZA, PO Box 513, SE-751 20, Uppsala, Sweden, per.johansson@ifau.uu.se

[§] IFAU, PO Box 513, SE-751 20, Uppsala, Sweden, sophie@langenskiold.eu

1. Introduction

Labor market discrimination is a major issue in the policy debate. In most countries, women and immigrants have lower employment rates and wages than men and natives. Moreover, it is often claimed that workers who are old, have several children, are overweight, or have a history of health problems are avoided by employers. In response to such concerns, most countries have introduced anti-discrimination laws and other policy measures to prevent discrimination, but it has proven to be difficult to find effective measures. One interpretation of these policy failures is that policymakers do not have enough information about the degree and nature of the existing discrimination. Hence, there is a need for more studies.

To explain the observed differences in labor market outcomes between groups, it is crucial to understand employers' hiring decisions. However, the existing empirical literature on firm-level hiring choices is rather limited (cf. Oyer & Schaefer, 2011). Profit maximizing (or cost minimizing) employers may have incentives to use easily observed characteristics, such as job applicants' gender, age, ethnicity or weight, as sorting criteria if they believe that these factors are correlated with productivity. Such behavior is particularly likely to occur in Europe, where wage setting typically is not flexible enough to allow wages to fully reflect real or perceived differences in productivity⁵, and where strict employment protection laws make it difficult to fire low productivity workers. In addition, the employers' preferences may directly affect their hiring decisions.

Empirically, it is very difficult to show the existence of discrimination. The traditional approach is to use administrative/survey data to analyze differences between groups. However, using this approach, it is virtually impossible to distinguish between the effects of worker characteristics observed by the firms but not included in the datasets and discrimination. Therefore, many researchers have turned to field experiments in the form of audit and correspondence studies.⁶ In audit studies, fictitious job applicants (auditors) are sent to employers to apply for

⁵ In Europe, wages are often set in collective agreements and the wage distributions are usually compressed. Therefore, discrimination is more likely to occur in hiring rather than in wage setting.

⁶ See Riach & Rich (2002) for a survey.

jobs, but these studies have been severely criticized since it is very difficult to ensure that employers perceive the auditors as identical in all dimensions except for the characteristic of interest (e.g. gender or ethnicity).⁷ Therefore, most recent studies use the correspondence testing methodology, where fictitious job applications which are identical in all dimensions except for the characteristic of interest are sent to employers, the callback rate to a job interview for the two groups are compared, and any difference is interpreted as discrimination. The main advantage of this approach is that the researcher has full control over the applicant information observed by the firms, and hence it is possible to estimate causal effects. However, this type of experiment also has some limitations. First, it is only possible to study discrimination in the initial stage of the hiring process, and hence these studies cannot tell us how discrimination affects the hiring and wage outcome. Second, it is often difficult to separate between preference-based and statistical discrimination since the characteristics of the recruiters are seldom known.⁸ For policymakers, this distinction is crucial since the appropriate policy response is likely to vary depending on the nature of the discrimination. Finally, it may be argued that it is ethically questionable to subject employers to fictitious job search (cf. Riach & Rich, 2004).⁹ Clearly, correspondence studies have increased our knowledge of discrimination. However, given the limitations of this approach, there is still a need for additional methods to analyze discrimination.

The main purpose of this study is to investigate if employers (recruiters) use information about job applicants' gender, age, ethnicity, religious beliefs, number of children, weight, or history of sickness absence in their recruitment decisions;

⁷ The auditors are often aware of the purpose of the experiment, and this may affect their behavior. Moreover, it is not obvious what employers consider as important for productivity. If some important attribute is not averaged out over the groups of auditors, this may affect the estimates of the degree of discrimination (cf. Heckman, 1998, and Heckman & Siegelman, 1993).

⁸ See Becker (1957), Phelps (1972), Arrow (1973) and Aigner & Cain (1977) for the original theoretical models, and Altonji & Blank (1999) and Bertrand (2011) for reviews.

⁹ Another problem is that the estimates of the degree of discrimination may be affected by the choice of the other (fixed) attributes in the applications. Both the choice of attributes and their levels may affect the estimates of the degree of discrimination (cf. Heckman & Siegelman, 1993). Some recent studies try to handle these issues; cf. Lahey (2009) and Neumark (2012).

i.e. if they discriminate.¹⁰ Moreover, we investigate if the degree of discrimination is affected by the firms' cost of uncertainty in hiring, measured by their co-payment in the sickness benefit system, and if the degree of discrimination differs between different types of recruiters (e.g. gender and ethnicity) and firms (e.g. sector and size). With the help of these results, we try to distinguish between preference-based and statistical discrimination.

To this end, we conduct a stated choice experiment. This approach is often used in transport, tourism and environmental economics, but our study is the first which uses it to study employers' hiring behavior.¹¹ In the experiment, which was conducted in the Swedish labor market in 2007, the recruiters are first asked to describe an employee who recently and voluntarily left the firm and then to choose between two hypothetical applicants to invite to a job interview or to hire as a replacement for their previous employee. The two applicants always have the same characteristics as the previous employee except for four attributes which are varied. In the first stage (interview), two of the attributes gender, age, education and experience are varied. In the second stage (hiring), two of the attributes gender, ethnicity, religious beliefs, number of children, weight, and history of sickness absence are varied. In addition, two other attributes are always varied. To measure the marginal value of each attribute (i.e. in monetary terms), the wage is varied (we use three alternatives based on the previous employee's wage). To measure the effect of the cost of uncertainty in hiring, the type of firm co-payment in the sickness benefit system is varied (we use three well-known alternatives). To keep the size of the experiment feasible, we use a fractional factorial design; i.e. a design which allows us to estimate the main effects of each characteristic as well as a limited number of interaction effects (cf. Section 3.2).

¹⁰ Throughout the paper, we label different treatment of all these groups as discrimination. However, it should be noted that in most countries it is not illegal to use factors such as the applicants' weight or history of sickness absence as sorting criteria (cf. Section 2.2).

¹¹ Van Beek *et al.* (1997) use a related approach – called conjoint analysis – to study how employers rank applications. Each employer was asked to rank 18 profiles, and state how many of the profiles they considered acceptable for hiring. They find that employers hardly care about education and experience. Instead, employers rank the profiles based on age, gender, health and

Compared to the other approaches, the stated choice approach has several important advantages. First, we have complete control over the information available to the recruiting firms, and hence we can estimate the causal effect of each of the applicants' characteristics. Also, sorting – i.e. that job applicants avoid employers they believe will discriminate – is not an issue. This is similar to correspondence studies. Second, we can study gender, age, ethnic and other kinds of discrimination simultaneously, which facilitates a comparison of the degree of different kinds of discrimination. This is different from traditional correspondence studies, but is similar to some recent studies (i.e. Rooth, 2011). Third, in sharp contrast to correspondence studies, we can study applicant characteristics which may be relevant in any of the stages of the hiring process; i.e. both the invitation to a job interview and the hiring decision. The fact that we can analyze factors relevant in the hiring decision is similar to audit studies, but we can do this without encountering the problems characterizing such studies. Fourth, the fact that we include the wage as one of the attributes which are varied makes it possible to calculate the marginal value of each characteristic. This tells us how much the wage must be reduced in order to make employers indifferent between applicants with and without a particular attribute (all else equal); i.e. how much 'wage flexibility' which would be needed to eliminate the differences in hiring. Importantly, we can do this even if there is no flexibility in actual wage setting in these dimensions. This is a new way of quantifying the degree of discrimination which may have policy-relevant implications for wage setting. Fifth, the fact that we vary the cost of uncertainty in hiring by considering several firm co-payment schemes in the sickness benefit system, and that we know the exact identity of the recruiters, gives us two new ways of, at least to some extent, disentangle preference-based and statistical discrimination. Sixth, the fact that the employers can only choose between the two hypothetical applicants means that our results are easier to interpret than the results in correspondence studies where employers

origin. However, the employers were not asked to choose a particular applicant to invite to a job interview or to hire. Moreover, the degree of discrimination was not quantified in wage terms.

also can choose other (real) applicants.¹² Finally, the fact that all recruiters know that they participate in an experiment, and are allowed to opt-out, means that no ethical concerns arise.

A limitation of the stated choice approach is that it is based on stated rather than observed behavior. A concern is, therefore, that the employers may not give answers which are consistent with their actual hiring behavior; i.e. a strategic or hypothetical bias may arise. However, in the design of the experiment, we use several methods to minimize this problem; e.g. the recruiters must assess several attributes simultaneously, which makes it very difficult for them to behave strategically. In addition, since the ‘politically correct’ behavior is to not discriminate (all employers we asked in the initial interviews stated that they never discriminated), we expect any remaining bias to reduce the estimates of discrimination. Another limitation is the risk of a low response rate since the study is conducted in the form of a survey.

Our results show that employers prefer not to recruit applicants who are old, non-European, Muslim, Jewish, obese, have several children, or have a history of sickness absence. The magnitude of the discrimination is substantial: to eliminate the differences in hiring, wages would have to be reduced by up to 50 percent. Our estimates of discrimination are similar in magnitude to the relative effects found in correspondence studies of the effects of e.g. age, ethnicity and obesity. Moreover, more firm co-payment in the sickness benefit system may reduce hiring, but does not seem to affect the degree of discrimination. Also, we find only small differences in hiring behavior between different types of recruiters and firms. Overall, our results are consistent with statistical discrimination.

The rest of the paper is organized as follows. Section 2 gives a background to the empirical analysis. The experimental design is described in Section 3, and Section 4 presents the data. In Section 5, the empirical framework and estimation strategy is discussed. Section 6 contains the results and Section 7 concludes.

¹² In correspondence studies, it is difficult to interpret cases when neither of the fictitious applicants is invited to an interview. Since there are other (real) job applicants, an employer may choose one of them and this may affect the magnitude of the estimated degree of discrimination.

2. Background

To provide the requisite background for our empirical analysis, we describe the state of discrimination in the Swedish labor market and discuss relevant institutional details.

2.1 Discrimination in the Swedish labor market

There are substantial differences in labor market outcomes between e.g. men and women and native Swedes and immigrants.

The employment rate for women is almost six percentage points lower than for men, and the gender wage gap is around 14 and 20 percent for blue- and white-collar workers, respectively.¹³ Differences in observed human capital variables explain some of these differences, but especially for white-collar workers a substantial wage gap remains unexplained (le Grand, 1997, and Albrecht *et al.*, 2003). However, there is substantial occupational segregation based on gender, and if detailed controls for occupations are included much of the remaining wage gap is eliminated (Petersen *et al.*, 2007). There are some recent studies of gender discrimination in hiring. Carlsson (2011), using data from a correspondence study, and Eriksson & Lagerström (2012), using data from an online CV database, find that women in general are not discriminated in hiring.¹⁴

The employment rate for immigrants is almost 18 percentage points lower than for native Swedes, and the median disposable income for immigrants is around 75 percent of the corresponding income for native Swedes.¹⁵ Edin & Åslund (2001) show that the average earnings of non-OECD immigrants are around 45 percent of the earnings of similar (on observed characteristics) native Swedes. There are many studies which present evidence that a substantial part of these differences reflects ethnic discrimination. Carlsson & Rooth (2007), using data from a

¹³ According to Statistics Sweden, the employment rate for men and women aged 20-64 in 2011 was 82.8 and 77.2 percent, respectively.

¹⁴ Other studies of gender discrimination using the correspondence testing methodology include Riach & Rich (1987, 2006), Neumark (1996), and Weichselbaumer (2004).

¹⁵ According to Statistics Sweden, the employment rate for native Swedes and immigrants aged 20-64 in 2011 was 83.3 and 65.5 percent, respectively.

correspondence study, show that applicants with Swedish names receive 50 percent more callbacks than applicants with Arabic names. Similar evidence is found in Eriksson & Lagerström (2012).¹⁶

For other types of discrimination, there are fewer studies. For age, Ahmed *et al.* (2012), using data from a correspondence study, show that the callback rate is much lower for older applicants.¹⁷ For weight, Rooth (2009), using data from a correspondence study, show that obese applicants have a lower callback rate.¹⁸

2.2 Swedish institutions

Some institutions which are relevant for the experiment are anti-discrimination legislation, employment protection, parental leave, pension schemes, and sickness benefits.

Anti-discrimination legislation

Preventing discrimination has for a long time been considered an important policy objective. In 1980, legislation regarding gender discrimination in the workplace was enacted. This was followed by legislation regarding discrimination based on ethnic origin, religion and other beliefs, disability, and sexual orientation. In 2009, a new comprehensive Discrimination Act came into force. This act states that it is illegal to discriminate based on sex, transgender identity or expression, ethnicity, religion or other belief, disability, sexual orientation or age. Employers are required to take active steps to prevent discrimination, e.g. by formulating gender equality plans. The enforcement of the act is entrusted to an ombudsman called the Equality Ombudsman. Violations are punishable by compensation payments and fines. The act has received a lot of attention recently, and should be well-known by most firms.

¹⁶ Other studies of ethnic discrimination are Bertrand & Mullainathan (2004) and the studies cited in Riach & Rich (2002). Drydakis (2010) studies the effects of religious affiliation.

¹⁷ Other studies of age discrimination include Bendick *et al.* (1996, 1999) and Albert *et al.* (2011).

¹⁸ Roehling (1999) reviews laboratory experiments on how employers treat obese workers. Cawley (2004) use observational data to study the effects of obesity. Hammermesh & Biddle (1994) and Mobius & Rosenblat (2006) study the relationship between beauty and labor market outcomes.

Employment protection, parental leave and pension schemes

The rules governing who the employer can layoff/fire, and under which circumstances, are in the Employment Protection Act. Layoffs are possible for redundancy and personal (i.e. non-redundancy) reasons. However, the rules for laying off workers for personal reasons are very strict. The rules for layoffs for redundancies state that layoffs should follow the last in first out rule, but this rule can be replaced by agreements between the employer and the union. Also, for a number of years, workers aged over 45 were protected by special rules that made it more difficult to lay them off. Violations against these rules are punishable by compensation payments and fines.

An institution which may affect how employers perceive workers with children is the rules for parental leave. These rules are rather generous. Workers can be absent from work for 480 days per child.¹⁹ In addition, workers are allowed to be absent to care for sick children up to 120 days per child and year.

An institution which may affect how employers perceive older workers is that occupational pension schemes for some workers – e.g. state employees and white-collar workers – are calculated in a way so that the employers' costs increase with the employee's age. However, in recent years most of these schemes have been replaced by defined contribution plans where the employers' costs are the same irrespective of the employee's age.

Sickness benefit system

Sickness benefits substitute income for workers who cannot perform their usual work because of temporary illness. All employed persons are covered by the insurance, and unemployed persons are covered if they are registered as unemployed at the Public Employment Service. The first day of sickness absence is not compensated. During the first seven days, it is up to the workers to determine if they are too sick to work/search. Thereafter, a doctor's certificate is

¹⁹ The benefit is approximately 80 percent of the previous wage up to an income cap for the first 390 days. For the remaining 90 days, it is a fixed amount per day.

required. The employers pay sickness benefits equivalent to approximately 80 percent of the wage during day 2-14 of the period of sickness absence (i.e. ‘sick pay’).²⁰ After this period, the responsibility for sickness benefits is transferred to the Swedish Social Insurance Agency (SSIA). The SSIA is also responsible for sickness benefits for the unemployed.

Around the time of the experiment, there was a lot of debate about the firms’ co-payment in the system. Between 2003 and 2005 employers were responsible for sickness benefits for the first three weeks, but in 2005 this was reduced to two weeks. In 2005, the firms’ co-payment was extended by the Social Democratic government. In addition to paying the full sickness benefits for the first two weeks, the employers were also required to pay 15 percent of the sickness benefits for the rest of their employees’ absence. The objective of the reform was to increase the employers’ incentives to rehabilitate long-term sick workers and, hence, to speed up the process of getting them back to work. To make the reform cost neutral for the employers, they were compensated by a lower payroll tax, but the reform made it more costly to hire workers with a high risk of sickness absence.²¹ The reform was heavily criticized by the employers, and was abolished by the new center-conservative government after they won the election in 2006.

Combined these institutions are likely to make employers careful not to hire workers whose productivity, or total labor costs, they are uncertain of, such as workers who are immigrants, belong to ethnic or religious minorities, are older, have many children, are overweight, or have history of sickness absence.

3. The experiment

To study discrimination, we conduct a stated choice experiment. This means that the employers are first asked to describe an employee who recently and voluntarily left the firm, and then to choose between two hypothetical applicants to invite to a job interview or to hire as a replacement for their previous

²⁰ Only income up to an income cap of SEK 321,000 is used to calculate benefits.

²¹ This effect is similar to the effect of experience rating. Andersson & Meyer (2000) discuss the effects of experience rating.

employee.²² The applicants' characteristics are varied in a way that allows us to analyze different kinds of discrimination.

For the experiment to work as intended it is important that the participating recruiters perceive it as realistic; i.e. that it has high face validity (Carson *et al.*, 1994). To achieve this, we started by interviewing a number of personnel managers about their hiring strategies. Then, we designed a pilot survey which we tested in focus groups with personnel managers, and finally we designed a pre-test questionnaire which was sent to 40 firms. Based on the results of these preliminary tests, we designed the final questionnaire which was sent to 1,000 firms. The experiment was conducted in 2007. In the following sections, we describe the details of the experiment and discuss the important issue of validity.

3.1 The initial interviews and the pilot survey

To get a better understanding of the recruitment process, we interviewed nine personnel managers. They all had long experience of recruitment and were very familiar with the recruitment process used in their firms. In particular, we discussed which factors they used to screen applications in order to choose whom to invite to job interviews and how they assessed applicants in the interviews. Also, we discussed how familiar they were with the sickness benefit system, including their knowledge of the firms' co-payment. Based on the answers from the managers, we concluded that it would be possible to study the effects of gender, age, education, experience, ethnicity, religious beliefs, number of children, weight, history of sickness absence, wage, and type of firm co-payment in the sickness benefit system. These attributes include four factors mentioned in the Discrimination Act. The remaining factors in the act were considered either as unobservable (transgender identity/expression and sexual orientation) or difficult to assess (disability). It should be emphasized that all managers stated that they

²² The employers were also allowed not to choose between the applicants, i.e. to opt-out. The reason for including an out-out option is both that it is ethically questionable not to include such an option in a survey, and that the stated choice literature stresses the importance of including such an option (cf. Section 3.4).

never discriminated in their hiring decisions.

Then, we designed a pilot survey and tested it in focus groups. In the pilot questionnaire, the respondents were asked to choose between two hypothetical applicants based on the attributes mentioned above. We tested the questionnaire in three focus groups consisting of around nine participants in each group. To ensure that the participants had recent experience of recruitment, we chose to include personnel managers who were responsible for screening and interviewing prospective employees. The participants were first asked to fill out the questionnaire, and then to discuss its design. A researcher always chaired the sessions and another took notes. The researchers could discuss the questions, the response options and how realistic the questions were, but never commented on the participants' responses.

The main conclusions from the interviews in the focus groups were: (i) The recruiters did remember the last employee who left the workplace voluntarily and which qualifications this worker had, (ii) the recruiters answers suggested that they used signals of productivity to sort workers at the different stages of the recruitment process; typically they seemed to distinguish between two stages – whom to invite to a job interview and whom to offer a job – and seemed to use different signals at each stage, (iii) the recruiters indicated that they understood the experiment (e.g. the combinations of attributes and the alternatives), but that the choice between the two hypothetical applicants became difficult if too many attributes had to be assessed simultaneously.

3.2 The design of the experiment

We decided to study applicant characteristics which should matter in any of the stages of the hiring process; i.e. the invitation to a job interview and the hiring decision. The recruiters' choice always involved a replacement for an employee who had recently and voluntarily left the firm. The choice always involved applicants with a full set of characteristics: the same attributes as the previous employee in all dimensions except for the attributes which are varied in the experiment. In the first stage (interview), we decided to vary four attributes plus

the wage and the type of firm co-payment in the sickness benefit system. The attributes are gender, age, education and experience, which are all typical attributes included in a CV. In the second stage (hiring), we decided to vary seven attributes plus the wage and type of firm co-payment in the sickness benefit system. The attributes are gender, ethnicity (country of birth), religious beliefs, number of children, weight, and two measures of the applicant's history of sickness absence (number of sickness spells in the previous year and their length). These are all factors which may be observed or discussed in a job interview, but are usually not included explicitly in a CV.²³ Except for gender, each of the twelve attributes has three possible alternatives (levels); cf. Section 3.3.

A problem is that it is not feasible to non-parametrically estimate the joint effect of all these attributes. The pilot survey indicated that the respondents could handle four attributes in each game (a game is a choice between two hypothetical applicants), and even if the number of attributes is restricted to four, a full factorial design with three levels would result in a very large survey which is not feasible to implement in practice.²⁴ However, the number of games can be reduced if we focus on estimating a restricted number of effects, i.e. if we use a fractional factorial design. Therefore, we decided to focus on estimating the main effects of the attributes and a few interactions. When choosing which interactions to include, we strived to include combinations of attributes which are most likely to have important interaction effects (these are listed below). To simplify, we exclude all interactions between the wage and the other attributes. We believe this to be a reasonable assumption since wage setting in Sweden is relatively inflexible: factors such as union contracts and internal wage structures limit the possibility to set wages individually. Also, we assume that all three- and four-way interactions are zero (i.e. interactions between more than two attributes).

²³ In Sweden, there are no laws preventing employers to ask questions about these factors in job interviews. However, it is illegal to use the characteristics mentioned in the Discrimination Act in hiring or wage setting decisions.

²⁴ To non-parametrically estimate all effects (including the three- and four-way interaction effects), we would need to include at least 81 (3^4) games for each recruiter.

In the questionnaire, we decided to include twelve games for each respondent: four games with respect to the invitation to a job interview and eight games with respect to the job offer (more games are needed in the hiring stage since more factors are varied). In each game, the recruiter is asked to choose between two hypothetical applicants with different attributes: each applicant has the same attributes as the previous employee except for the four attributes which are varied, and, in most cases, each of the attributes which are varied can take on three levels.

To estimate the marginal value of each attribute (i.e. the wage reduction needed to make employers indifferent between applicants with and without a particular attribute), we always included the wage as one of the attributes we varied in all games. Moreover, since we are interested in varying the firms' cost of uncertainty in hiring (i.e. the extent of firm co-payment in the sickness benefit system), the type of firm co-payment was varied in all games. The other two variables which were varied in each game were chosen as follows: gender and age (games 1 and 2), education and experience (games 3 and 4), ethnicity and religious beliefs (games 5 and 6), gender and number of children (games 7 and 8), gender and weight (games 9 and 10), and frequency and intensity of sickness absence (games 11 and 12). This design allows us to estimate the main effect of each attribute as well as the interaction effects between the attributes within each game; i.e. the interaction effects between all applicant characteristics and the type of sickness benefit system, between gender and age, education and experience, etc. To estimate these effects, we need to include 162 hypothetical applicants.²⁵ However, some of the combinations are not relevant to include. As an example, consider a game where the attributes are gender, weight, the wage and the extent of firm co-payment in the sickness benefit system. In such a game, most recruiters would choose a normal weight man with a ten percent lower wage and the cheapest sickness benefit scheme. In the stated choice literature, it is emphasized that including such (unrealistic) choices can jeopardize the quality and credibility of

²⁵ The calculation of the number of hypothetical applicants required to estimate the relevant effects is based on standard statistical tables for fractional factorial experiments (c.f. Hahn & Shapiro, 1966).

the experiment. Therefore, we decided to drop six such combinations. We divided up the remaining 156 hypothetical applicants to 13 versions of the questionnaire with 12 games in each in order to alleviate the burden for the respondents. The questionnaire sent to each employer was a random draw from these 13 versions.

In addition to the games, the questionnaire also included detailed questions about the last employee who left the firm voluntarily, the recruiter, and the firm for which the recruiter worked (cf. Section 4.2 for descriptive statistics). An example of two games in the questionnaire is presented in the Appendix.²⁶

3.3 The attributes and their levels

When we chose the attributes and their levels, our objective was to include information which are typically mentioned in a CV, or are typically observed or discussed in a job interview. Moreover, we wanted to choose levels of the attributes which appeared realistic to the employers. Most of the attributes/alternatives are very straightforward.

For gender, we used male and female. For age, we used 29 years or younger, 30-55 years and 56 years or older. For education, the pilot survey indicated that it was best to relate this attribute to what was required for the job in question. Therefore, we used the lowest possible education (lowest quartile), the average education (middle quartiles), and the highest education for the job in question (highest quartile). For experience, we used 4 years or less, 5-7 years and 8 years or more. For ethnicity, we used born in the Nordic countries, born in Europe except the Nordic countries, and born in Africa, the Middle Eastern countries or South America. We chose this division since there is ample empirical evidence indicating that ethnic discrimination mainly affects immigrants from developing countries outside Europe. For religious beliefs, we used Christian, Jewish and Muslim. The majority population in Sweden is Christian and the two other groups are important minorities. For the number of children, we used no children, one child and two children or more. For weight, we used the silhouettes originally

²⁶ An example of all the games in a questionnaire is presented in Eriksson *et al.* (2012)

developed by Stunkard *et al.* (1983).²⁷ The silhouettes for men and women are illustrated in Figure 1. Silhouettes 1/2 represent underweight, silhouettes 3/4 normal weight, silhouettes 5/6 overweight and silhouettes 7/8/9 obese. We used normal weight, overweight and obese. For the two measures of sickness absence, we used 1-2, 3-5, and 6 or more times in the last year for frequency, and 7 or less, 8-14, and 15 or more days each time for intensity. For the wage, we decided to relate it to the wage of the previous employee. We used the alternatives the same wage, a ten percent lower wage, and a ten percent higher wage than the previous employee. Since we want to quantify the degree of discrimination in wage terms, it is important that we have enough variation in the wage alternatives. However, the alternatives must also be perceived as realistic by the employers. In the pre-test we used a five percent lower/higher wage, but the results convinced us that we should increase the variation to potentially increase the efficiency in the estimation of the marginal values. Even if wages are compressed in Sweden, a ten percent lower/higher wage should be realistic in most cases. For the extent of firm co-payment in the sickness benefit system, we used two weeks of full firm payment, three weeks of full firm payment, and two weeks of full firm payment plus 15 percent co-payment for the rest of the sickness absence. The firms' cost of uncertainty should be lowest in the first and highest in the last alternative. All these alternatives have been discussed in the policy debate, and were well-known and understood by the recruiters taking part in the pilot survey.

3.4 Validity

A potential concern with stated choice experiments is that the elicited preferences, or marginal values, may differ from what would be the case in real-world situations. This problem is known as strategic or hypothetical bias in the literature. The related method of contingent valuation has been severely criticized based on

²⁷ These silhouettes have previously been used in several studies (e.g. Berkey *et al.*, 1999, Munoz *et al.*, 1996), and have also been linked to body mass index (BMI) in Bulik *et al.* (2001). An alternative would be to use manipulated photos – as in e.g. Rooth (2009) –, but this approach is

these arguments (cf. Carson *et al.*, 1996, and Hausman, 2012). However, a number of methods for reducing this bias have been suggested in the literature (cf. List, 2001, Murphy *et al.*, 2004, Carson, 2012, and Kling *et al.*, 2012), and several recent studies show that stated and revealed preferences often coincide (cf. Murphy *et al.*, 2010, and Jacquemet *et al.*, 2011). The results in these studies suggest that the importance of hypothetical bias depends on the experimental setting (cf. Taylor *et al.*, 2001, and Ajzen *et al.*, 2004); the hypothetical valuation of a good is likely to exceed its actual valuation in situations which involve an important perceived ethical dimension and where a high value is considered ‘ethically commendable’, but not in other situations.²⁸ As is emphasized in the recent survey by Kling *et al.* (2012), there is a ‘current best practice for survey design’. Two important considerations are incentive compatibility (i.e. if the respondents have incentives to answer truthfully) and consequentiality (i.e. if the questions may affect outcomes which matter for the respondents).

The stated choice experiment approach differs from the ‘all or nothing’ dichotomous contingent valuation approach by adding realism to the experiment as the respondents are asked to choose between alternatives with different attributes in situations closely resembling individual purchasing – or as in our case hiring – decisions. In such experiments, several recent studies show that it is not possible to reject the hypothesis of equal valuation of the attributes in stated choice experiments and the real world (cf. Carlsson & Martinson, 2001, Cameron *et al.*, 2002, Backhaus *et al.*, 2005, and List *et al.*, 2006). A recent study by Deuchert *et al.* (2011) illustrates some of these points. They conduct an ex ante

likely to be inferior since the photos usually are only showing the applicants’ face, and thus is unlikely to be tightly linked to the applicants’ BMI.

²⁸ Ajzen *et al.* (2004) argue that the activation of positive attitudes may result in a discrepancy between intentions and actions in cases where the experiment involves the selling of desirable public goods (cf. Lusk & Norwood 2009). Guzman & Kolstad (2007) argue that the problem of hypothetical bias may be related to the respondents’ reluctance to invest in costly information acquisition in a hypothetical experiment. These arguments are based on the idea that people want to improve or maintain their self-image, but that it is more costly to do so when real money is involved. The idea that people derive value from the mere expression of attitudes and values that correspond to what they consider to be ‘good and worthy’ actions is however not new (cf. Katz, 1960, and Andreoni, 1989, 1990).

evaluation of Swiss firms' willingness to train disabled persons using a discrete choice stated preference design, where firms are asked if they accept to train hypothetical disabled persons with a number of assigned attributes (profiles). The results show that 22 percent of the evaluated profiles were accepted by firms. Since training disabled persons is most likely 'socially desirable', hypothetical bias is a concern. To address this issue, the authors use a follow-up question – where the firms were asked how certain they are that they would accept to train a person with the shown profile – to estimate a bias-corrected probability, which is 8.6 percent. Moreover, they estimate the marginal effect of the profiles' attributes, e.g. age, school performance and absence. Interestingly, the results show that the relative effects of the attributes are rather similar both with and without the bias-correction. This suggests that hypothetical bias may not be a major concern in a study focusing on relative effects.

The results in the stated choice literature suggest that there are several ways to minimize the problem with hypothetical bias. It is emphasized that the experiments should be preceded by interviews, focus groups and pre-tests, that the respondents' incentives to answer truthfully should be thoroughly analyzed, and that the questions should be relevant for the respondents. More specifically, the respondents should be given detailed information about the good (person) they are asked to value, they should be provided with a known reference which they can compare the alternatives against, and they should be allowed to make any choice between the suggested alternatives or to opt-out (cf. Hensher, 2010).

In the experiment, we incorporate all these features. We conducted an elaborate pre-study analysis, and made adjustments of the experimental design based on what we learned. Moreover, all the respondents handle personnel issues on a day-to-day basis, they are asked to consider well-defined replacements for a well-known previous employee, and they are allowed to choose or not to choose (i.e. to opt-out) between the applicants in each game. Also, they must consider several applicant attributes simultaneously, which makes it very difficult to behave strategically. Finally, since the 'politically correct' behavior is to not discriminate (all participants in the initial interviews stated that they never discriminated), we

expect any remaining hypothetical bias to affect the results downwards (i.e. reduce the estimates of discrimination).

4. Data

In this section, we describe the sample selection and present some descriptive statistics for the experiment.

4.1 Sample selection

We decided to focus on medium- and large-sized workplaces in Stockholm County, which is the largest Swedish county in terms of inhabitants. In this county, there are 2,048 workplaces with one location and 20 or more employees. From this population, we drew a sample of 1,000 workplaces to which we sent the questionnaire. Since we want to study potential differences in discrimination between different types of workplaces, we made a stratified sampling where the strata were based on the sector, size and gender composition of the workplace.²⁹ The survey was administered by Statistics Sweden and was sent to the employers by postal mail. An accompanying letter stated that the purpose of the study was to investigate the recruitment behavior of firms. The participants received no compensation for their participation, and were asked to return the questionnaire by postal mail. Two reminders were sent to non-respondents.

The response rate was around 46 percent. It was somewhat higher in the public sector than in the private sector.³⁰ A separate analysis of the non-respondents shows that the main reason why they did not participate seemed to be a lack of time rather than a reluctance to participate in a study of recruitment behavior.³¹ In

²⁹ For sector, we used the strata: (i) the private sector, and (ii) the public sector. For size, we used the strata: (i) 20-49 employees, (ii) 50-99 employees, (iii) 100-199 employees, (iv) 200-499 employees, and (v) 500 or more employees. For gender distribution, we used the strata: (i) less than 40 percent women, (ii) 40-60 percent women, and (iii) more than 60 percent women.

³⁰ The response rate in the private sector was 44 percent and in the public sector 50 percent. For the different size, categories the response rates were 43, 48, 56, 47 and 60 percent. For the gender distribution categories, the response rates were 46, 44 and 47 percent.

³¹ In the drop-out analysis, we contacted 33 employers who did not return their questionnaire. In this group, 21 stated that they were unwilling to answer the questionnaire, and eleven of them

total, 426 employers (recruiters) are included in the analysis (this corresponds to 4,895 observations).³²

4.2 Descriptive statistics

In this section, we describe the workplaces that participated in the experiment and the recruiters who responded to the questions. Then, we describe the characteristics of the last employee who left the firms voluntarily; i.e. the reference person in the experiment. Finally, we discuss to what extent employers try to gather information about their applicants' health and history of sickness absence. More detailed descriptive statistics are given in Eriksson *et al.* (2012).

The workplaces and the recruiters

The workplaces are quite diverse: nearly two thirds are in the private sector, one third in the public sector, and almost half have less than 50 employees.³³

Around one third of the recruiters who answered the questions were managing directors, one third personnel managers, and the rest held other positions tasked with personnel issues. Most of them worked with recruiting, personnel policy and rehabilitation, and almost all of them had worked with these issues for a number of years. Around three quarters were aged 30-55, around one quarter aged over 55, most were born in Sweden, almost two thirds were women, and most had several children. In general, they were highly educated with nearly 80 percent having a university education. They considered themselves to be Christians in three quarters of the cases and atheist/agnostic in most of the remaining cases. A majority considered themselves to be overweight or obese, and most of them had only limited sickness absence.

The previous employee

Nearly all respondents stated that at least one employee had left their workplace

stated a reason: eight said that the reason was a lack of time and one said that the reason was the purpose of the study.

³² Respondents who reported that no employee had left the firm are excluded from the analysis.

voluntarily in the two years preceding the experiment, and the majority stated that this employee had left within the last six months. Also, most respondents answered our rather detailed questions about the characteristics of this employee. Hence, we find it reasonable to assume that the respondents remembered the most recent employee who left the workplace.

In the questionnaire, the respondents were asked to describe this employee. The employees were in the majority of the cases men (51%), aged 30-55 (69%), born in Sweden (84%), and had a secondary (39%) or university (53%) education. Their qualifications for the positions they held were in the middle two or highest quartiles in 47 and 41 percent of the cases, respectively. Mostly, they had eight years or more of experience (52%), but had only spent part of this time in their current position. The clear majority was believed to be Christian, but in around a quarter of the cases their religious beliefs were unknown to the employer. Most of them had only been absent from work due to sickness on a few short occasions, and around 40 percent were judged to be overweight or obese. Their mean wage was SEK 26,800 (€3,200), and their median wage was SEK 25,000 (€2,900).

Health and history of sickness absence

An important issue is if it is reasonable to assume that recruiting employers try to gather information about their job applicants' health and history of sickness absence in the recruitment process. We find that around half of the recruiters claim that they try to gather information about their applicants' health (44%) or history of sickness absence (41%). They try to get information about health by asking the applicants or their references, asking about leisure activities, requiring health examinations, asking about smoking habits, evaluating physical appearance, and asking about previous occupational injuries. They try to get information about their applicants' history of sickness absence by asking the applicants or their references, and requesting the applicants to provide an excerpt

³³ Around 90 percent of the workplaces had a plan against gender, sexual, ethnic or religious discrimination, and 60 percent had a plan against age discrimination.

from the SSIA. Based on this evidence, we find it reasonable to assume that the recruiters often have some information about their applicants' health and history of sickness absence.

5. Theoretical framework and estimation

This section sketches a simple theoretical model of the employers' recruitment decision and discusses the empirical modeling.

5.1 Theoretical framework

It is reasonable to assume that most recruiting employers have access to only limited information about their job applicants' productivity – e.g. skills and turnover propensity – prior to hiring. In such cases, employers may find it optimal to base their hiring decisions on easily observed characteristics – e.g. gender or ethnicity – which they believe are correlated with productivity. Then, statistical discrimination may arise. However, employers, co-workers or customers may also have preferences regarding different groups which affect the hiring decisions. Then, preference-based discrimination may arise.

To capture the possibility of both types of discrimination, we can state the employer's maximization problem in the following way: the employer faces different applicants with observed characteristics given by the vector \mathbf{x} . The employer's utility depends on his or her expectations regarding the profits associated with the observed characteristics \mathbf{x} and his or her preferences regarding these characteristics. Therefore, the employer maximizes:

$$U(\mathbf{z}) = pf(\mathbf{x}) - [w(\mathbf{x}) + d(\mathbf{x})], \quad (1)$$

where $U(\mathbf{z})$ is the employer's utility function, $\mathbf{z} = (\mathbf{x}, w(\mathbf{x}))$, p is the price of the product the firm produces, $f(\mathbf{x})$ is the production function, $w(\mathbf{x})$ is the wage function, and $d(\mathbf{x})$ is the (dis)taste function, with $d(\mathbf{x}) \geq 0$. If an employer has negative preferences regarding some component of \mathbf{x} , then the employer will only hire the applicant if his or her marginal productivity is higher than the wage.

5.2 Empirical modeling

Let us now consider an employer who participates in the experiment. A previous employee with observed characteristics (and known productivity) $\mathbf{z}^0 = (\mathbf{x}^0, w^0)$ is supposed to be replaced by one of two hypothetical applicants with observed characteristics $\mathbf{z}_1 = (\mathbf{x}_1, w_1, I_1)$ and $\mathbf{z}_2 = (\mathbf{x}_2, w_2, I_2)$. \mathbf{x}_1 and \mathbf{x}_2 are a subset of \mathbf{x}^0 , which varies between the games, and I_i , $i = 1, 2$, are the type of firm co-payment in the sickness benefit system.

If we expand the utility associated with hiring applicant i in game g with observed characteristics \mathbf{z}_{ig} around the previous employee's set of observed characteristics, we get:

$$U_e(\mathbf{z}_{ig}) = \alpha_e + \mathbf{x}_{ig}\alpha + I_{ig}\gamma + (I_{ig}\mathbf{x}_{ig})\delta - w_{ig}\beta + \eta_{ige}, \quad (2)$$

where e denotes the employer and g the game. In this equation, we allow for a (linear) fixed employer effect α_e to take into account that the production, wage and taste functions are likely to differ between workplaces. Note that the equation include all relevant interaction effects, and that the wage, by design, is made orthogonal to the attributes \mathbf{x}_{ig} and I_{ig} .

In the choice between the two hypothetical applicants i and j , applicant i is invited to an interview or hired if:

$$U_e(\mathbf{z}_{ig}) - U_e(\mathbf{z}_{jg}) > 0 \text{ and } \partial U_e(\mathbf{z}_{ig}) / \partial \mathbf{z}_{ig} > 0. \quad (3)$$

We estimate $\theta = (\alpha', \gamma, \delta', \beta)'$, where the vector α measures the average degree of discrimination (both statistical and preference-based) with respect to the applicant characteristics.

From this specification, the marginal value of attribute x_k can be calculated as the ratio of parameters, hence:

$$\frac{\partial w}{\partial x_k} = \frac{\partial U_e(\mathbf{z}_{ig}) / \partial x_{igk}}{\partial U_e(\mathbf{z}_{ig}) / \partial w_{ig}} = \frac{\alpha_k + I_{ig}\delta_k}{-\beta}. \quad (4)$$

Note that if $\delta_k = 0$ we get the marginal value of attribute x_k when there is no interaction effect between attribute x_{igk} and I_{ig} . Henceforth, we refer to the case

where all interaction effects are zero as the baseline model.³⁴ It should be noted that we can calculate these marginal values of all characteristics even if there is no variation in real-world wages in these dimensions.

We estimate the parameters using the ordinary least square estimator. Thus, we are estimating (for applicants 1 and 2):

$$y_{1eg} - y_{2eg} = (\mathbf{x}_{1g} - \mathbf{x}_{2g})\boldsymbol{\alpha} + (I_{1g} - I_{2g})\gamma + (I_{1g}\mathbf{x}_{1g} - I_{2g}\mathbf{x}_{2g})\boldsymbol{\delta} - (w_{1g} - w_{2g})\boldsymbol{\beta} + \eta_{1eg} - \eta_{2eg}, \quad (5)$$

where y is the outcome variable.

In each game, the employer (recruiter) is asked to choose one of the two applicants or to opt-out.³⁵ If the respondent chooses an applicant, the dependent variable is either 1 or -1 if the first or the second applicant is chosen. We allow for heteroskedasticity by estimating the standard errors using a robust covariance matrix. We prefer to use the linear model rather than the logit model since the design is orthogonal for the linear model only. We have, however, also estimated fixed effect logit models and, qualitatively, the results are the same.

6. Results

In this section, we analyze the degree and nature of discrimination. First, we estimate the effects of each of the attributes included in the experiment; i.e. we estimate the baseline model with no interaction effects. We measure the degree of discrimination in terms of the probability of being invited to a job interview (the callback rate) and being offered a job (the job offer rate) as well as in terms of the marginal value of each attribute (the wage reduction needed to make the employers indifferent between applicants with and without a particular attribute). When comparing the magnitude of our estimates to the (callback) estimates in correspondence studies, it is important to note that our baseline probability is 100 percent (i.e. the employer can only choose one of the two applicants) whereas in a correspondence study this baseline probability is typically around 10 percent (i.e.

³⁴ I.e. the interaction effects between attribute k and the type of firm co-payment as well as all other applicant attributes are set to zero.

³⁵ In each of the games, around 30 percent of the respondents opted out. The results are not affected if we include these cases in the estimation.

the employer can choose another applicant). Hence, it is more illustrative to compare the relative effects. Second, we investigate if the degree of discrimination is affected by the firms' cost of uncertainty in hiring (the extent of firm co-payment in the sickness benefit system), i.e. we estimate the interaction effects. Third, we study if the degree of discrimination varies with the type of recruiter and firm; i.e. we estimate the model separately for different subgroups.

6.1 The degree of discrimination

Invitations to job interviews

Table 1 presents the estimates of the probability of being invited to a job interview (the callback rate). Almost all estimates of the effects of the applicants' attributes on the callback rate are statistically significant. The most striking result is the very large negative effect for applicants over 55 years old: the callback rate for such an applicant is 64 percentage points lower than the callback rate for an applicant who is less than 30 years old. The callback rate for a 30-55 year old applicant is 12 percentage points higher than for an applicant who is less than 30 years old. In contrast, there is no gender difference in the callback rate. Education and experience have the expected effects, i.e. a higher callback rate for applicants with more education or experience. In particular, education has a strong effect: an applicant with the highest education relevant for the job in question has an 82 percentage point higher callback rate than an applicant with the lowest education. The wage has a negative effect on the callback rate. Finally, the type of firm co-payment in the sickness benefit system has a clear effect on the callback rate: reducing the time the firms pay sickness benefits from three to two weeks would increase the callback rate with 9 percentage points whereas combining the same reduction with a 15 percent employer co-payment for the complete sickness spell would decrease the callback rate with 7 percentage points. These results are as expected since more firm co-payment in the sickness benefit system (for a given wage) implies higher costs for the firms and, therefore, less hiring.

Overall, our results confirm the results in previous correspondence studies. As mentioned above, Ahmed *et al.* (2012) find strong evidence of age discrimination,

and Carlsson (2011) find no evidence of gender discrimination. Moreover, the magnitude of the estimates is in line with the results in previous studies: for older workers, Ahmed *et al.* (2012) find a relative effect of 69 percent and Lahey (2008) a relative effect of 30 percent. Also, the results suggest that more firm co-payment in the sickness benefit system may have a negative effect on hiring.

Job offers

We now turn to the firms' hiring decision. In Table 2, we present the estimated effects of the applicants' attributes on the probability of offering a job (first column) and the marginal value of each attribute (second column).

For the job offer rate, we again find that most of the effects are statistically significant. A first striking result is the strong effects of ethnicity and religious beliefs. The job offer rate is similar for applicants born in Europe, while applicants born in Africa, the Middle Eastern countries and South America face a much lower job offer rate (minus 28 percentage points). Applicants who are Muslim or Jewish also have a much lower job offer rate than applicants who are Christian (minus 30 and 26 percentage points). A second striking result is the very large negative effect for obese applicants: being obese decreases the job offer rate by 83 percentage points compared to having normal weight. There is also, as expected, a lower job offer rate for workers with a history of sickness absence, especially for those with many spells and long durations. Moreover, applicants with two or more children have a 25 percentage points lower job offer rate.³⁶ Also, the wage has a negative effect. Finally, the type of firm co-payment in the sickness benefit system has the expected effect; i.e. if the employers' costs increase the job offer rate decreases.

In the second column, we have the corresponding estimates in terms of the marginal value of each attribute. These estimates can be interpreted as the wage reduction needed to make employers indifferent between applicants with and

³⁶ We have analyzed if this effect is different for men and women, but find that the interaction effect is not statistically significant.

without a particular attribute (all else equal). The results suggest that to eliminate the negative hiring effect, an applicant who is born in Africa, the Middle Eastern countries or South America would need to get a wage which is around 16 percent lower than an applicant who is born in the Nordic countries, an applicant who is Muslim (Jewish) would need to get a wage which is 17 (15) percent lower than an applicant who is Christian, an overweight (obese) applicant would need to get a wage which is 8 (48) percent lower than a normal weight applicant, and an applicant with a history of sickness absence would need to get up to a 48 percent lower wage. Since the mean of the previous employees' monthly wage is SEK 26,800 (€3,200), a 15 percent reduction would correspond to a SEK 4,000 (€470) lower wage, and a 48 percent reduction to a SEK 12,900 (€1,500) lower wage. The exact numbers should be interpreted with caution, but they indicate that the wage reductions needed to make employers indifferent between applicants with and without some of the attributes are substantial.

Overall, the results confirm the results from correspondence studies of the early stages of the hiring process, but also extend these results by considering attributes which have not been analyzed before. It is widely documented that non-European immigrants and Muslims face widespread discrimination in the Swedish labor market (cf. Carlsson & Rooth, 2007 who find a relative effect of 50 percent in callback rates). The negative effect for Jewish applicants is somewhat more surprising since this group has been in Sweden for a long time and is considered as well-established. However, recently there has been quite a lot of debate in Sweden about anti-Semitism and discrimination against Jews.³⁷ Also, it should be noted that we measure the effect given that the employers know that an applicant is a Jew. In contrast to most of the other characteristics we consider, the fact that an applicant is a Jew is probably more difficult to observe. It is also striking that applicants who are obese, have a history of sickness absence and/or have several

³⁷ See e.g. the Telegraph 21 Feb 2010, <http://www.telegraph.co.uk/news/worldnews/europe/sweden/7278532/Jews-leave-Swedish-city-after-sharp-rise-in-anti-Semitic-hate-crimes.html>. Discrimination against Jews in Sweden is discussed in e.g. U.S. Department of State (2005) and Bachner & Ring (2005).

children face substantially lower job offer rates. Moreover, the estimates of the size of the wage reductions needed to make employers indifferent between applicants with and without these attributes are substantial. The discrimination we find evidence of may reflect both preference-based and statistical discrimination: for example, discrimination against ethnic and religious minorities may reflect both types of discrimination, while discrimination against workers with several children or a history of sickness absence is likely to reflect statistical discrimination. Finally, the results again suggest that more firm co-payment in the sickness benefit system may have a negative effect on hiring.

6.2 Discrimination and the cost of uncertainty in hiring

We now turn to the analysis of what happens with the degree of discrimination when we vary the extent of firm co-payment in the sickness benefit system. In the experiment, we have three sickness benefit schemes; two weeks of full firm payment, three weeks of full firm payment, and two weeks of full firm payment followed by 15 percent firm co-payment. In the previous section, we saw that more co-payment resulted in lower callback and job offer rates. Now we want to investigate if different groups are affected differently when the extent of firm co-payment is changed. This analysis may help us understand the nature of the discrimination: if employers are statistically discriminating then the degree of this behavior should be correlated with the extent of firm co-payment. Hence, if employers are using such a hiring strategy, we should expect them to be less likely to hire applicants with attributes which signals a high risk of sickness absence when the extent of firm co-payment increases. In addition, more co-payment may also affect how employers perceive other attributes which signals uncertainty about total labor costs. The reason is that more co-payment implies that the employers' uncertainty about total labor costs increases. If employers are risk adverse, more uncertainty would make them less willing to hire all applicants they perceive as risky.

To estimate these effects, we interact the sickness benefit scheme variables (type of firm co-payment) with the attributes signaling a higher risk of sickness

absence (e.g. overweight/obese and a history of sickness absence) or uncertainty in the hiring decision in a more general sense (e.g. ethnic and religious minorities).³⁸ Tables 3 and 4 present the estimates of the interaction effects for the two stages of the hiring process; the invitation to a job interview (the callback rate) and the job offer (the marginal value in terms of the wage).

We find little evidence of any change in the degree of discrimination when the extent of firm co-payment is changed. In Table 3, only four interaction effects are statistically significant: there is a negative interaction effect between 30-55 years of age and two weeks of full firm payment, and positive interaction effects between the highest education, 5-7 years of experience, more than 7 years of experience and two weeks of full firm payment followed by 15 percent firm co-payment. The last three effects are reasonable; i.e. education and experience are considered as more valuable in a system with more co-payment. In Table 4, there are some statistically significant interaction effects, but most of them are difficult to interpret. For example, the estimates imply that applicants who are Muslims fare better both in a system with less co-payment (two weeks) and in a system with more co-payment (two weeks followed by 15 percent).

Overall, the results show that there is little evidence of any systematic relationship between the degree of discrimination and the extent of firm co-payment in the sickness benefit system. This may be interpreted as evidence against statistical discrimination, but there is an alternative explanation. It may be that the firms' total costs associated with an employee's absence are high in all three sickness benefit schemes. For most firms, the total costs when an employee is absent include not only the costs of their co-payment, but also many other costs associated with the disruption in production that the absence may cause. Moreover, firms may consider other negative effects associated with hiring the 'wrong' worker as more important than costs associated with sickness absence.

³⁸ In the regressions, we include all interactions described in Section 3, but only report the interactions between the sickness benefit schemes and the applicant characteristics. Most other interactions are not statistically significant. These estimates are available upon request.

Therefore, it may be that the (realistic) changes in the co-payment scheme we analyze are simply too small to affect the employers' hiring decisions.

6.3 Discrimination and the type of recruiters and firms

We now turn to the issue of whether the degree of discrimination is similar in all firms or if it differs depending on the type of the recruiter and/or firm. Again, this may help us understand the nature of the discrimination. We focus on the effects on the hiring decision, and estimate the models separately for different subgroups with respect to the characteristics of the recruiters and the workplaces.³⁹ As before, all regressions include controls for the firm's industry.

The characteristics of the recruiters

The characteristics of the recruiters which we consider are gender, age, education, experience, ethnicity, religious beliefs, children, weight, and history of sickness absence. Some of the results are presented in Table 5.

Comparing male and female recruiters, there are only three statistically significant differences.⁴⁰ Female recruiters, on average, are more reluctant than male recruiters to offer a job to an applicant who is obese or have a history of sickness absence of more than six spells per year. Also, female recruiters respond more to a change in the extent of firm co-payment in the sickness benefit system from three to two weeks than male recruiters. One interpretation of these results is that female recruiters, on average, are more risk averse than male recruiters.⁴¹

For age, we have estimated the model separately for recruiters who are less than 30 years, 30-55 years and over 55 years. We find three statistically significant differences, all of which between the oldest and youngest recruiters. The two groups respond differently to a change in the extent of firm co-payment

³⁹ The results for the callback rates are available upon request.

⁴⁰ Hensvik (2011) reports evidence that Swedish male and female employers do not use different hiring strategies.

⁴¹ Although we control for the industry the firms belong to, we cannot exclude the possibility that these differences are explained by other systematic differences in the characteristics of the firms male and female recruiters work in.

in the sickness benefit system and to applicants who have a history of more than six spells of sickness absence.

For education, we have estimated the model separately for recruiters with university, secondary and primary education. We only find two statistically significant differences; recruiters with secondary and university education differ in their treatment of applicants born in Europe outside the Nordic countries and applicants with one child. An analysis of recruiters with different lengths of experience yields no important differences.

For ethnicity, we have estimated the model separately for recruiters born in Sweden and recruiters born in other countries.⁴² We find that both types of recruiters treat applicants in the same way in most cases. The only statistically significant difference is that recruiters who are immigrants are less likely to discriminate against applicants who are Jews. There is a similar tendency for applicants who are Muslims, but this difference is not statistically significant. However, there is no difference in the recruiters' treatment of applicants born inside and outside Europe. Some of these results may be interpreted as suggestive of preference-based discrimination, but the results should be interpreted with caution since almost no recruiters are born outside Europe.

For religious beliefs, we have estimated the model separately for recruiters who state that they have a religious belief (mostly Christian) and recruiters who do not state that they have a religious belief (i.e. recruiters who answered atheist, agnostic or do not know). We find some statistically significant differences; 'religious recruiters' treat applicants who have two or more children less favorably and women more favorably. There is also a tendency that 'religious recruiters' view applicants who are Muslims (statistically significant at the ten percent level) or Jews (not statistically significant) less favorably.

For the recruiters' family situation, we have estimated the model separately for recruiters who have no children, one child and two or more children. We find only

⁴² Åslund *et al* (2009) find that Swedish employers tend to hire workers with similar background as their own.

two statistically significant differences; recruiters with two or more children treat applicants with one child more favorably than recruiters with no children. Moreover, recruiters with two or more children are more reluctant to hire applicants who are obese or have a history of many sickness spells in the past.

For weight, we have estimated the model separately for recruiters who report that they are underweight, normal weight, overweight, and obese. We find some evidence that obese recruiters treat applicants who are obese more favorably than other recruiters. Moreover, overweight recruiters seem to value European applicants lower and female applicants higher than normal weight recruiters. We have also run separate regressions based on the recruiters' history of sickness absence, but find no important differences.

Overall, our results indicate that the different types of recruiters treat applicants with different attributes in a rather similar way. This is more supportive of statistical discrimination than preference-based discrimination since, if preference-based discrimination is important, we would expect to find clear differences in recruitment behavior between different types of recruiters. However, some of the results are consistent with preference-based discrimination.

The characteristics of the workplaces

The characteristics of the workplaces which we consider are the sector, size and gender composition of the workplace. Some results are presented in Table 6.

For the sector of the workplaces, we have estimated the model separately for workplaces in the private and public sector, but find essentially no differences, even though there is some tendency that employers in the public sector treat applicants with attributes associated with risk more unfavorably.

For the size of the workplaces, we have estimated the model separately for four different categories: 20-49, 50-99, 100-249, and >249 employees. Our findings suggest that large workplaces are less likely to discriminate than small workplaces. For example, we find that workplaces with more than 249 employees do not discriminate against applicants who are born outside Europe, whereas workplaces with 20-49 and 50-99 employees attach a negative value to this

characteristic. Large workplaces also consider applicants who have two or more children or who are overweight more favorably than smaller firms.

For the gender composition of the workplaces, we have estimated the model separately for workplaces with less than 40 percent women, 40-60 percent women and more than 60 percent women. However, we find essentially no differences.

Overall, we find that different types of workplaces treat applicants with different attributes in a rather similar way, except that large firms seem less likely to discriminate than small firms. Again, this may be interpreted as supportive of statistical discrimination since it is likely that the consequences for a small firm of hiring the 'wrong' worker are more substantial than for a large firm. However, the scope for preference-based discrimination may be bigger in small firms where the recruiter and the employees interact more closely in the day-to-day operations.

7. Conclusions

Labor market discrimination is a major issue in the policy debate. Many policy attempts have been tried to reduce discrimination, but it has proven to be difficult to find effective policy measures. One interpretation of this is that policymakers do not have enough information about the degree and nature of the existing discrimination. In recent years, correspondence studies have extended our knowledge of discrimination, but this approach has some inherent weaknesses; e.g. focusing on only the initial stage of the hiring process. Hence, there is a need for new approaches that can help us get a better understanding of discrimination.

In this paper, we study if Swedish employers (recruiters) use information about job applicants' gender, age, ethnicity, religious beliefs, number of children, weight, or history of sickness absence when they recruit workers by conducting a stated choice experiment. In the experiment, the recruiters are first asked to describe an employee who recently and voluntarily left the firm, and then to choose between two hypothetical applicants to invite to a job interview or to hire as a replacement for their previous employee. The applicants always have a full set of characteristics: the same characteristics as the previous employee in all dimensions except for the four attributes which are varied.

Our results show that the recruiters prefer not to recruit applicants who are old, non-European, Muslim, Jewish, obese, have several children, or have a history of sickness absence. Some of these results confirm what we know from previous observational and correspondence studies of discrimination, but our results also extend the existing literature in several important ways. First, our results suggest that discrimination may be a major issue *both* in the invitation to an interview phase and in the hiring phase of the recruitment process. Our results show the importance of both attributes which are typically included in a CV and attributes which are typically observed or discussed in a job interview. Second, we can *quantify* the degree of discrimination in a new way by calculating the wage reduction needed to make employers indifferent between applicants with and without a particular attribute (all else equal). The exact magnitudes should be interpreted with caution, but our estimates indicate that substantial wage differentials – from 10 to 50 percent – are needed to compensate employers. Our estimates of discrimination are similar in magnitude to the relative effects found in correspondence studies of the effects of e.g. age, ethnicity and obesity.

The discrimination that we find evidence of may reflect both preference-based and statistical discrimination: discrimination based on e.g. ethnicity and religious beliefs may reflect both types of discrimination, while discrimination based on e.g. previous sickness absence and the number of children are likely to reflect statistical discrimination (employers may worry that such workers may be absent from work due to their own or their children's sickness). In reality, both types of discrimination are likely to co-exist. However, policymakers need to be informed about their relative importance since the policies needed to prevent discrimination are very different depending on the type of discrimination. We used two methods to try to disentangle the two types of discrimination. First, we analyzed if the degree of discrimination was affected by the firms' cost of uncertainty in hiring as measured by the extent of firm co-payment in the sickness benefit system. The idea being that, if the primary reason for discrimination is uncertainty, we would expect to find more discrimination when we increase the cost of uncertainty in hiring. Our results show that the degree of discrimination is *not* affected by

changing the cost of uncertainty. This may be interpreted as evidence against statistical discrimination, but may also reflect that the costs to firms of recruiting the ‘wrong’ worker are substantial in all the (realistic) sickness benefit schemes we analyzed and, therefore, did not affect their recruitment behavior. Moreover, firms may consider other negative effects associated with hiring the ‘wrong’ worker as more important than costs associated with sickness absence. Second, we analyzed if the degree of discrimination differed between different types of recruiters and firms. The idea being that, if the degree of discrimination is the same for different types of recruiters, it is likely to reflect statistical discrimination. Our results show that different types of recruiters – e.g. native Swedes and immigrants – respond in a rather similar way to most of the applicants’ attributes. This may be interpreted as supportive of statistical discrimination since it is unlikely that different types of recruiters should have the same preferences regarding the attributes we consider if preference based discrimination are important. Moreover, we find that large firms discriminate less than small firms. This is also supportive of statistical discrimination since the consequences of hiring the ‘wrong’ worker should be more substantial for small firms. Overall, our results indicate that both preference-based and statistical discrimination exist, but that statistical discrimination may be more important.

An important issue is what implications our results have for wage setting. It may be argued that workers who have a lower productivity should be paid a lower wage, and that wage setting should allow for such wage differentials to avoid a lower hiring rate for workers in low productivity groups. If productivity is difficult to assess prior to hiring, it may even be argued that workers with attributes which employers perceive as signals of low productivity should accept a lower wage until their true productivity can be verified. However, in most countries, it is considered both morally unacceptable and illegal for employers to use factors such as gender, ethnicity and religious beliefs as hiring and/or wage setting criteria. In contrast, in most countries it is not illegal for employers to use factors such as family situation, weight and health to sort workers. In fact, it may be argued that these attributes have a direct effect on productivity and, at least to

some extent, are personal choices. However, it is not obvious that it should be considered as fair that workers with these attributes are paid a lower wage. In addition, our results indicate that the wage differentials needed may be substantial: taken at face value, our calculations indicate that workers with certain characteristics would need to accept very low wages and/or that very large wage subsidies would need to be introduced. Both of these options are probably difficult to implement in real-world economies, especially in Europe. In the Swedish context, our results may also shed some light on the puzzling finding that some very generous wage subsidies – such as the entry recruitment incentives for newly arrived immigrants – have received very limited interest from employers.

Our results may also have implications for the design of firm co-payment schemes in social insurance systems, such as sickness benefits (cf. experience rating). Our results show that more firm co-payment in the sickness benefit system reduces the callback and job offer rates, but that changing the extent of co-payment does not seem to affect the degree of discrimination. These results suggest that policymakers should consider the effects on hiring carefully before introducing more firm co-payment, but that such changes may not affect vulnerable groups more than other groups.

Our study demonstrates that stated choice experiments may be a valuable complement to the existing approaches to study firms' hiring behavior. This approach has several important advantages, such as allowing us to study applicant characteristics which may be relevant in any of the stages of the hiring process and giving us a new way of measuring how firms' value different applicant characteristics. We have focused on how this approach can be used to measure discrimination, but it may also be a useful tool to study firms' hiring behavior in other dimensions. For example, it may be possible to use it to test the predictions of more elaborate theoretical models of firm-level hiring choices (cf. Oyer & Schaefer, 2011). In addition, our study illustrates that a potential way to distinguish between different types of discrimination is to use policy changes which involve changes in the firms' costs of hiring the 'wrong' worker. In our case, the variation may not have been sufficiently large to find any effects, but

future studies may be able to introduce more variation between the alternatives. However, it is important that the alternatives are perceived as realistic by the firms, and this – as in our case – may put restrictions on the alternatives that can be analyzed. This approach may also be useful for *ex ante* evaluations of policy reforms (cf. Deuchert *et al.*, 2011). For example, introducing experience rating in the unemployment benefits system may reduce layoffs, but could also make employers more reluctant to hire workers from groups they perceive as risky. Potentially, the stated choice approach could be used to study the effects on the hiring process of such reforms before they are introduced.

In summary, our results suggest that discrimination is prevalent in both stages of the hiring process, that the magnitude of this discrimination in terms of the wage reductions needed to make employers indifferent between applicants with and without some attributes are substantial, and that the discrimination, at least partially, is likely to reflect statistical discrimination.

References

- Ahmed, A., Andersson, L. and Hammarstedt, M. (2012) “Does age matter for employability? A field experiment on ageism in the Swedish labor market, *Applied Economics Letters*, 19, 403-406.
- Aigner, D.J. and Cain, G.G. (1977) “Statistical theories of discrimination in labor markets”, *Industrial and Labor Relations Review*, 30, 175-187.
- Ajzen, I., Brown, T. C. and Carjaval, F. (2004) “Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation”, *Personality and Social Psychology Bulletin*, 30, 1108–1120.
- Albert, R., Escot, L. and Fernández-Cornejo, J. A. (2011) “A field experiment to study sex and age discrimination in the Madrid labour market, *International Journal of Human Resource Management*, 22, 351-75.
- Albrecht, J., Björklund, A, and Vroman, S. (2003) “Is there a glass ceiling in Sweden”, *Journal of Labor Economics*, 21, 145-177.

- Altonji, J.G. and Blank, R.M. (1999) "Race and gender in the labor market, in: Ashenfelter, O., and Card, D. (eds), *Handbook of labor economics*, Vol. 3B, Elsevier.
- Anderson, P, and Meyer, B. (2000) "The effects of the unemployment insurance payroll tax on wages, employment, claims and denials", *Journal of Public Economics*, 78, 81-106.
- Andreoni, J. (1989) "Giving with impure altruism: Applications to charity and Ricardian equivalence", *Journal of Political Economy*, 97, 1447-58.
- Andreoni, J. (1990) "Impure altruism and donations to public goods: A theory of warm-glow giving", *Economic Journal*, 100, 464-77.
- Arrow, K. J. (1973) The theory of discrimination, in Ashenfelter, O. and Rees, A. (eds.), *Discrimination in labor markets*, Princeton University Press, Princeton, NJ.
- Åslund, O., Hensvik, L. and Nordström Skans, O. (2009) "Seeking similarity: How immigrants and natives manage at the labor market", IFAU working paper 2009:24.
- Backhaus, K. Wilken, R. Voeth, M. and Sichtmann, C. (2005) "An empirical comparison of methods to measure willingness to pay by examining the hypothetical bias", *International Journal of Market Research*, 47, 543-562
- Bachner, H. and Ring, J. (2005) Antisemitic images and attitudes in Sweden, The Living History Forum, Stockholm.
- Becker, G. (1957) *The economics of discrimination*, Chicago University Press.
- Bendick, M., Brown, L. E. and Kennington, W. (1999) "No foot in the door: An experimental study of employer discrimination against older workers", *Journal of Aging & Social Policy*, 10, 5-23.
- Bendick, M., Jackson, C.W. and Romero, J.H. (1996) "Employment discrimination against older workers: An experimental study of hiring practices, *Journal of Aging & Social Policy*, 8, 25-46.
- Berkey, C.S., Frazier, A.L., Gardner, J.D. and Colditz, G.A. (1999) "Adolescence and breast carcinoma risk", *Cancer*, 1999, 2400-2409.

- Bertrand, M. (2011) New perspectives on gender, in: Ashenfelter, O., and Card, D. (eds) *Handbook of labor economics*, Volume 4B, Elsevier.
- Bertrand M. and Mullainathan, S. (2004) “Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination”, *American Economic Review*, 94, 991-1013.
- Bulik, C., Wade, T., Heath, A., Martin, N., Stunkard, A. and Eaves, L. (2001) “Relating body mass index to figural stimuli: Population-based normative data for Caucasians”, *International Journal of Obesity*, 10, 1517-1524.
- Cameron, T.A. Poe, G.L. Ethier, R.G. and Schulze, W.D. (2002) “Alternative non-market value-elicitation methods: Are the underlying preferences the same?”, *Journal of Environmental Economics and Management*, 44, 391–425
- Carlsson, M. (2011) “Does hiring discrimination cause gender segregation in the Swedish labor market?”, *Feminist Economics*, 17, 71-102.
- Carlsson, F. and Martinsson, P. (2001) “Do hypothetical and actual marginal willingness to pay differ in choice experiments?”, *Journal of Environmental Economics and Management*, 41, 179-192.
- Carlsson, M. and Rooth, D.-O. (2007), “Evidence of ethnic discrimination in the Swedish labor market using experimental data”, *Labour Economics*, 14, 716-729.
- Carson, R.T. (2012) “Contingent valuation: A practical alternative when prices aren’t available”, *Journal of Economic Perspectives*, 26, 27-42.
- Carson, R. Flores, E. Martin, K. and Wright, J. (1996) “Contingent valuation and revealed preference methodologies: Comparing the estimates for quasi-public goods”, *Land Economics*, 72, 80–99
- Carson, R., Louviere, J., Anderson, D., Arabie, P., Bunch, D., Hensher, D., Johnson, R., Kuhfeld, W., Steinberg, D., Swait, J., Timmermans, H. and Wiley, J. (1994) “Experimental analysis of choice”, *Marketing Letters*, 5, 351-368.
- Cawley, J. (2004) “The impact of obesity on wages”, *Journal of Human Resources*, 39, 451–74.

- Deuchert, E., Kauer, L. and Meisen Zannol, F. (2011) “Would you train me with my mental illness? Evidence from a discrete choice experiment” Working paper Center for Disability and Integration, University of St. Gallen.
- Drydakis, N. (2010) “Religious affiliation and labour bias”, *Journal for the Scientific Study of Religion*, 49, 472–488.
- Edin, P.-A. and Åslund, O. (2001) ”Invandrare på 1990-talet arbetsmarknad, in Bergmark, Å. (ed) *Ofärd i välfärden*, SOU 2001:54, 101-142.
- Edin, P.-A. and Lagerström, J. (2006) “Blind dates: Quasi-experimental evidence on discrimination, Working paper 2006:4, IFAU, Uppsala.
- Eriksson, S., Johansson, P. and Langenskiöld, S. (2012) “What is the right profile for getting a job? A stated choice experiment of the recruitment process”, Discussion paper no. 6691, IZA, Bonn.
- Eriksson, S. and Lagerström, J. (2012) “Detecting discrimination in the hiring process: Evidence from an Internet-based search channel”, *Empirical Economics*, 43, 537-563.
- Guzman, R. and Kolstad, C. (2007) “Researching preferences, valuation and hypothetical bias”, *Environmental and Resource Economics*, 37, 465–487.
- Hahn, G.J. and Shapiro, S.S. (1966) “A catalog and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments”, General Electric research and development center, Schenectady, New York.
- Hamermesh, D.S. and Biddle. J.E. (1994) “Beauty and the labor market”, *American Economic Review*, 84, 1174–94.
- Hausman, J. (2012) “Contingent valuation: From dubious to hopeless”, *Journal of Economic Perspectives*, 26, 43-56.
- Heckman, J. (1998) “Detecting discrimination”, *Journal of Economics Perspectives*, 12, 101-116.
- Heckman, J. and Siegelman, P. (1993) The Urban Institute audit studies: Their methods and findings, in Fix, M.E. and Struyk, R.J. (ed) *Clear and convincing evidence: Measurement of discrimination in America*, 187-258, Washington, Urban Institute.

- Hensher, D.A. (2010) “Hypothetical bias, choice experiments and willingness to pay” *Transportation Research Part B: Methodological*, 44, 735–752.
- Hensvik, L. (2011), “Manager impartiality? Worker-firm matching and the gender wage gap”, Working paper 2011:22, IFAU, Uppsala.
- Jacquemet, N., Joule, R.-V., Luchini, S. and Shogren J. F. (2011) “Do people always pay less than they say? Testbed laboratory experiments with IV and HG values”, *Journal of Public Economic Theory*, 13, 857-882.
- Katz, D. (1960) “The functional approach to the study of attitudes”, *Public Opinion Quarterly*, 24, 163-204.
- Kling, C.L., Phaneuf, D.J. and Zhao, J. (2012) “From Exxon to BP: Has some number become better than no number?”, *Journal of Economic Perspectives*, 26, 3-26.
- Lahey, J.N. and Beasley, R.A. (2009), “Computerizing Audit Studies”, *Journal of Economic Behavior & Organization*, 70, 508-514.
- le Grand, C. (1997) “Occupational segregation and the gender gap in Sweden”, Stockholm University.
- List, J.A. (2001) “Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards”, *American Economic Review*, 91, 1498–1507.
- List, J.A., Sinha, P. and Taylor M. (2006) “Using choice experiments to value non-market goods and services: Evidence from field experiments”, *Advances in Economic Analysis & Policy*, 6.
- Lusk, J.L. and Norwood F.B. (2009) “An inferred valuation method”, *Land Economics*, 85, 500-514.
- Mobius, M.M. and Rosenblat, T.S. (2006) “Why beauty matters”, *American Economic Review*, 96, 222–35.
- Munoz, K.A., Ballard-Barbash, R., Graubard, B., Swanson, C.A., Schairer, C. and Kahle, L.L. (1996) “Recall of body weight and body size estimation in women enrolled in the breast cancer detection and demonstration project (BCDDP)”, *International Journal of Obesity Related and Metabolic Disorders*, 1996, 854-859.

- Murphy, J., Allen, P., Stevens, T. and Weatherhead, D. (2004) "A meta-analysis of hypothetical bias in stated preference valuation, Department of Resource Economics, University of Massachusetts, Amherst.
- Murphy, J., Stevens, T. and Yadav, L (2010) "A comparison of induced value and home-grown value experiments to test for hypothetical bias in contingent valuation", *Environmental and Resource Economics*, 47, 111-123.
- Neumark, D. (2012) "Detecting discrimination in audit and correspondence studies", *Journal of Human Resources*, 47, 1128-1157.
- Neumark, D. with the assistance of Bank, R. and Van Nort, K. (1996) "Sex discrimination in restaurant hiring: An audit study", *Quarterly Journal of Economics*, 111, 915-941.
- Oyer, P. and Schaefer, S. (2011) Personnel Economics: Hiring and incentives, in: Ashenfelter, O. and Card, D. (ed) *Handbook of labor economics*, volume 4B, Elsevier.
- Petersen, T., Snartland, V. and Meyerson Milgrom, E.M. (2007) "Are female workers less productive than male workers?", *Research in Social Stratification and Mobility*, 25, 13-37.
- Phelps, E.S. (1972) "The statistical theory of racism and sexism", *American Economic Review*, 62, 659-661.
- Riach, P. and J. Rich, (1987) "Testing for sexual discrimination in the labour market", *Australian Economic Papers*, 26, 165-178.
- Riach, P. and Rich, J. (2002) "Field experiments of discrimination in the marketplace", *Economic Journal*, 112, 480-518.
- Riach, P. and Rich, J. (2004) "Deceptive field experiments of discrimination: Are they ethical?", *Kyklos*, 57, 457-470.
- Riach, P. and Rich, J. (2006) "An experimental investigation of sexual discrimination in hiring in the English labor market", *Advances in Economic Analysis & Policy*, 2.
- Roehling, M.V. (1999) "Weight-based discrimination in employment: Psychological and legal aspects", *Personnel Psychology*, 52, 969-1016.

- Rooth, D.-O. (2009) "Obesity, attractiveness and differential treatment in hiring: A field experiment", *Journal of Human Resources*, 44, 710-735.
- Rooth, D.-O. (2011), "Work out or out of work: The labor market return to physical fitness and leisure sports activities", *Labour Economics*, 18, 399-409.
- Stunkard, A.J., Sorensen, T. and Schulsinger, F. (1983) "Use of the Danish adoption register for the study of obesity and thinness", in: Kety, S.S., Rowland, L.P., Sidman, R.L. and Mathysse, S.W. (ed), *The Genetics of Neurological and Psychiatric Disorders*, Raven Press, 115-120.
- Taylor, L. Mckee, O.M., Laury, S.K. and Cummings R.G. (2001) "Induced value tests of the referendum voting mechanism", *Economics Letters*, 71, 61-65.
- US Department of State (2005) Report on Global Anti-Semitism, US Department of State, Washington.
- Van Beek, K.W.H., Koopmans, C.C. and van Praag, B.M.S. (1997) "Shopping at the Labour Market: A Recent Tale of Fiction", *European Economic Review*, 41, 295-317.
- Weichselbaumer, D. (2004) "Is it sex or personality? The impact of sex-stereotypes on discrimination in applicant selection", *Eastern Economic Journal*, 30, 159-86.

Appendix: Examples of the games in the questionnaire

Who do you invite to an interview if you must choose one of the following two applicants?
The only differences between the two applicants and the employee you described in Part 1 are summarized in the table.

APPLICANT A:		APPLICANT B
Woman	Gender	Man
29 years or younger	Age	56 years or older
10% lower wage than the employee	Wage (non-negotiable)	10% higher wage than the employee
Day 2-21	Sickness insurance	Day 2-14

1 Invite A 2 A and B similar 3 Invite B

Who do you hire if you must hire one of the following two applicants?
The only differences between the two applicants and the employee you described in Part 1 are summarized in the table.

APPLICANT A:		APPLICANT B
Europe excluding the Nordic countries	Country of birth	Africa/Middle East/South America
Christian	Religion	Muslim
Same wage as the employee	Wage (non-negotiable)	10% lower wage than the employee
Day 2-21	Sickness insurance	Day 2-14 + COFIN

1 Hire A 2 A and B similar 3 Hire B

Figure 1 Silhouettes as a way to illustrate weight (reprinted from Stunkard *et. al.*, 1983)

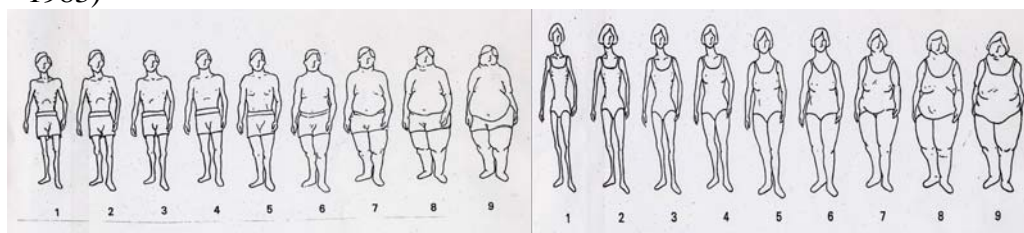


Table 1 Callback rate to a job interview

<i>Gender (ref: male):</i>	
Female	-0.01 (0.04)
<i>Age (ref: 29 years or younger):</i>	
30-55 years	0.12** (0.05)
>55 years	-0.64*** (0.05)
<i>Education (ref: lowest quartile):</i>	
Middle quartiles	0.63*** (0.06)
Highest quartile	0.82*** (0.05)
<i>Experience (ref: <5 years):</i>	
5-7 years	0.13** (0.06)
>7 years	0.12** (0.06)
<i>Wage:</i>	
Wage	-0.01*** (0.00)
<i>Sickness benefits (ref: 3 weeks full firm payment):</i>	
2 weeks full firm payment	0.09** (0.04)
2 weeks full firm payment plus 15% co-payment	-0.07** (0.04)

Notes: Estimated with ordinary least squares. The dependent variable is the probability of being invited to a job interview (the callback rate). Also included are a constant and dummy variables for the industry the workplace belongs to. Robust standard errors are in parentheses. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 2 Job offer rate and the marginal value of each attribute

	Job offer rate	Marginal value
<i>Gender (ref: male):</i>		
Female	0.00 (0.03)	0.10 (1.62)
<i>Ethnicity (ref: Nordic):</i>		
Other European	-0.02 (0.06)	-0.94 (3.34)
Africa, Middle East, South America	-0.28*** (0.06)	-16.24*** (3.67)
<i>Religious belief (ref: Christian):</i>		
Jewish	-0.26*** (0.06)	-15.02*** (3.61)
Muslim	-0.30*** (0.06)	-17.19*** (3.78)
<i>Children (ref: no children):</i>		
1 child	-0.09 (0.06)	-5.06 (3.32)
2 or more children	-0.25*** (0.06)	-14.58*** (3.40)
<i>Weight (ref: normal weight):</i>		
Overweight	-0.13** (0.06)	-7.68** (3.29)
Obese	-0.83*** (0.05)	-48.08*** (4.91)
<i>Wage:</i>		
Wage	-0.02*** (0.00)	-
<i>Frequency of sickness absence (ref: 1-2 times per year):</i>		
3-5 times per year	-0.41*** (0.05)	-23.85*** (3.33)
6 or more times per year	-0.83*** (0.05)	-48.03*** (4.57)
<i>Intensity of sickness absence (ref: 7 or less days):</i>		
8-14 days each time	-0.29*** (0.04)	-16.57*** (2.94)
15 or more days each time	-0.55*** (0.05)	-31.69*** (3.63)
<i>Sickness benefit (ref: 3 weeks full firm payment):</i>		
2 weeks full firm payment	0.13*** (0.03)	7.24*** (1.69)
2 weeks of full firm payment plus 15% copayment	-0.09*** (0.03)	-4.97*** (1.53)

Notes: The estimates in the first column are estimated with ordinary least squares, and the dependent variable is the probability of offering a job. Also included are a constant and dummy variables for the industry the workplace belongs to. In the second column are the implied marginal values of the attributes in terms of the wage. They are calculated as the ratio of the coefficient for the attribute and the coefficient for the wage. Robust standard errors are in parentheses (in the second column calculated using the delta method). ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 3 Change in the callback rate to a job interview from changing the firm co-payment in the sickness benefit system (reference 3 weeks of full firm payment)

	2 weeks	2 weeks plus 15%
<i>Gender (ref: male):</i>		
Female	0.02 (0.06)	-0.05 (0.05)
<i>Age (ref: 29 years or younger):</i>		
30-55 years	-0.14** (0.07)	0.04 (0.07)
>55 years	-0.08 (0.08)	-0.03 (0.07)
<i>Education (ref: lowest quartile):</i>		
Middle quartiles	0.08 (0.11)	0.08 (0.10)
Highest quartile	0.12 (0.09)	0.23*** (0.09)
<i>Experience (ref: <5 years):</i>		
5-7 years	0.02 (0.10)	0.28** (0.12)
>7 years	0.00 (0.11)	0.18* (0.10)

Notes: Estimated with ordinary least squares. The table presents the interaction effects between the attributes and the type of firm co-payment in the sickness benefit system. The regression also includes all applicant attributes in Table 1, all other interaction effects, a constant and dummy variables for the industry the workplace belongs to. Robust standard errors are in parentheses. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 4 Change in the job offer rate from changing the firm co-payment in the sickness benefit system (reference 3 weeks of full firm payment)

	2 weeks	2 weeks plus 15%
<i>Gender (ref: male):</i>		
Female	-5.52** (2.63)	2.19 (2.28)
<i>Ethnicity (ref: Nordic):</i>		
Other European	-8.31 (5.46)	-5.56 (4.93)
Africa, Middle East, South America	6.52 (5.34)	6.60 (4.89)
<i>Religious belief (ref: Christian):</i>		
Jewish	11.67** (5.43)	-0.27 (4.78)
Muslim	18.99*** (6.53)	14.82*** (5.47)
<i>Children (ref: no children):</i>		
1 child	-3.53 (4.70)	-4.18 (4.77)
2 or more children	4.27 (4.53)	0.85 (4.66)
<i>Weight (ref: normal weight):</i>		
Overweight	-8.38* (5.05)	-1.89 (4.36)
Obese	-8.71* (4.65)	2.75 (4.07)
<i>Frequency of sickness absence (ref: 1-2 times per year):</i>		
3-5 times per year	3.99 (4.28)	2.80 (3.93)
6 or more times per year	-5.42 (3.98)	-7.69** (3.77)
<i>Intensity of sickness absence (ref: 7 or less days):</i>		
8-14 days each time	-0.13 (3.98)	-2.33 (3.77)
15 or more days each time	-2.83 (4.50)	-1.92 (3.92)

Notes: The table presents the implied marginal values in terms of the wage for the interaction terms in the job offer regression; i.e. a regression with all applicant characteristics in Table 2, all interaction effects, a constant and dummy variables for the industry the workplace belongs to (the calculation of the marginal effects is described in the notes to Table 2). The standard errors are calculated using the delta method and the heteroskedastic robust covariance matrix. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 5 The marginal value of the applicants' attributes, subgroups recruiters

	Gender		Age			Education		Ethnicity		Religious beliefs	
	Male	Female	<30 years	30-55 years	>55 years	Secondary	University	Native Swede	Immigrant	Non- religious	Religious
Female	-0.93 (1.91)	1.01 (2.65)	1.07 (4.31)	-1.17 (1.62)	6.43 (7.51)	-5.09 (4.18)	1.16 (1.78)	0.91 (1.71)	-7.11 (4.96)	-5.55* (3.02)	1.96 (1.91)
Other European	-1.40 (3.93)	1.84 (5.47)	-1.93 (10.78)	-0.95 (3.41)	-2.68 (14.16)	-22.00** (9.90)	2.43 (3.66)	0.41 (3.51)	-11.79 (10.64)	-5.69 (5.83)	0.84 (3.99)
Non-European	-11.04*** 4.20	-20.03*** (6.21)	13.33 (15.73)	-11.66*** (3.60)	-59.44** (25.29)	-25.31*** (9.58)	-14.33*** (4.06)	-15.96*** (3.86)	-23.02* (11.99)	-13.11* (6.95)	-17.10*** (4.26)
Jewish	-15.17*** (4.14)	-13.95** (6.03)	1.41 (9.42)	-14.17*** (3.64)	-15.04 (15.63)	-6.91 (8.53)	-17.99*** (4.15)	-17.67*** (3.86)	7.14 (10.74)	-6.30 (6.33)	-17.73*** (4.30)
Muslim	-16.95*** (4.32)	-16.26*** (6.27)	20.37 (19.38)	-15.23*** (3.80)	-27.06 (17.12)	-18.64*** (9.86)	-18.93*** (4.23)	-19.00*** (4.04)	-4.10 (11.08)	-5.44 (6.64)	-19.79*** (4.52)
1 child	-2.38 (3.76)	-9.20 (5.65)	-5.91 (8.03)	-3.15 (3.32)	-20.13 (16.30)	15.50* (9.43)	-8.19 (3.62)	-4.40 (3.50)	-10.45 (9.79)	-7.15 (6.28)	-5.15 (3.83)
2 or more	-10.01** (3.94)	-19.09*** (5.71)	-2.18 (7.61)	-9.91*** (3.36)	-49.35** (21.23)	-10.57 (8.37)	-15.47** (3.78)	-13.50*** (3.55)	-19.84* (10.84)	-3.58 (5.92)	-17.68*** (4.04)
Overweight	-0.54 (3.98)	-12.83** (5.42)	-0.82 (7.10)	-6.18* (3.36)	-23.40 (16.03)	-8.19 (8.02)	-7.04** (3.62)	-7.01** (3.46)	-8.86 (9.75)	-4.28 (6.14)	-8.83** (3.83)
Obese	-33.23*** (4.85)	-61.77*** (9.36)	-31.42*** (10.19)	-40.11*** (4.46)	-109.76*** (40.11)	-65.23*** (15.25)	-44.17*** (5.12)	-47.81*** (5.14)	-42.38*** (13.97)	-42.05*** (8.41)	-48.93*** (5.79)
3-5 times per year	-19.31*** (3.94)	-30.05*** (5.76)	-5.02 (7.83)	-20.92*** (3.20)	-49.89** (21.17)	-30.59*** (9.51)	-22.53*** (3.56)	-22.79*** (3.45)	-31.57*** (11.11)	-15.90*** (5.70)	-25.39*** (3.95)
6 or more times per year	-34.20*** (4.52)	-63.06*** (9.00)	-16.30** (7.98)	-42.15*** (4.29)	-97.44*** (34.15)	-44.94*** (11.69)	-48.78*** (4.99)	-47.81*** (4.78)	-49.07*** (14.62)	-45.56*** (8.03)	-47.73*** (5.35)
8-14 days each time	-14.28*** (3.44)	-19.26*** (4.92)	-11.88 (8.88)	-13.70*** (2.82)	-30.21* (15.95)	-17.68** (7.77)	-16.24*** (3.18)	-15.95*** (3.08)	-17.73** (8.72)	-14.91*** (5.53)	-16.50*** (3.39)
15 or more days each time	-25.85*** (4.09)	-37.66*** (6.39)	-26.16** (10.17)	-28.36*** (3.51)	-59.43** (23.08)	-36.73*** (9.81)	-30.09*** (3.90)	-32.30*** (3.81)	-20.10** (10.14)	-28.54*** (6.29)	-32.08*** (4.31)
2 weeks full firm payment	3.34* (1.92)	12.09*** (3.00)	-4.09 (4.24)	7.13*** (1.71)	12.29 (8.14)	13.12*** (4.82)	6.17*** (1.82)	6.98*** (1.77)	8.58 (5.34)	3.34 (3.02)	8.52*** (2.01)
2 weeks of full firm payment plus 15%	-2.28 (1.83)	-7.50*** (2.53)	-15.71*** (5.10)	-4.30*** (1.53)	-6.94 (6.83)	1.79 (3.88)	-5.68*** (1.70)	-4.75*** (1.60)	-7.16 (4.87)	-6.04** (2.83)	-4.37** (1.77)

Notes: The table reports the implied marginal values of the attributes in terms of the wage. They are calculated as the ratio of the coefficient for the attribute and the coefficient for the wage from regressions of the job offer rate (see the notes to Table 2 for the details). The reference categories are the same as in Table 2. The standard errors are calculated using the delta method and the heteroskedastic robust covariance matrix. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 6 The marginal value of the applicants' attributes, subgroups firms

	Sector		Size				Percentage women		
	Private	Public	20-49	50-99	100-249	>249	<40%	40-60%	>60%
Female	-1.59 (1.75)	5.47 (4.65)	-2.89 (2.34)	0.37 (5.22)	1.22 (3.39)	4.96 (3.40)	-1.09 (2.49)	2.89 (3.01)	-0.23 (3.27)
Other European	-2.77 (3.59)	3.04 (9.44)	-8.59* (4.64)	6.45 (11.41)	17.27** (8.50)	5.17 (7.60)	-0.91 (5.11)	-2.19 (6.19)	0.82 (6.50)
Africa, Middle East, South America	-18.53*** (3.96)	-6.81 (10.82)	-22.22*** (5.38)	-26.16** (12.22)	-2.93 (7.42)	3.19 (9.13)	-18.32*** (5.75)	-15.31** (6.75)	-14.97** (7.39)
Jewish	-12.39*** (3.71)	-33.00** (12.96)	-18.54*** (5.07)	-10.23 (11.02)	2.95 (8.45)	-13.83* (8.20)	-13.81** (5.46)	-10.49* (6.24)	-24.00*** (8.02)
Muslim	-14.40*** (3.98)	-28.95** (12.35)	-18.06*** (5.20)	-3.16 (12.47)	-7.85 (7.61)	-20.93** (9.91)	-16.92*** (5.96)	-17.35** (7.63)	-20.56*** (7.33)
1 child	-8.94** (3.67)	9.96 (9.44)	2.21 (4.62)	-17.79 (11.56)	-12.69* (6.94)	-3.34 (7.44)	-6.12 (5.04)	-10.18 (6.36)	2.69 (6.68)
2 or more children	-14.59*** (3.68)	-20.47** (9.85)	-16.50*** (4.82)	-22.16** (11.09)	-12.19* (7.24)	-1.65 (7.02)	-18.14*** (5.40)	-14.10** (5.96)	-9.83 (6.67)
Overweight	-9.04** (3.61)	-4.54 (8.85)	-8.12* (4.84)	-14.11 (10.55)	10.77 (6.70)	-18.35** (7.35)	-11.48** (5.17)	-7.94 (5.78)	-3.48 (6.61)
Obese	-42.22*** (4.98)	-78.18*** (18.85)	-49.81*** (7.39)	-86.14*** (22.75)	-31.08*** (7.86)	-34.56*** (8.37)	-51.35*** (7.94)	-38.58*** (7.68)	-54.06*** (10.69)
3-5 times per year	-20.51*** (3.53)	-34.32*** (10.44)	-25.80*** (5.02)	-30.62*** (11.24)	-20.99*** (6.58)	-17.27*** (6.19)	-22.87*** (5.46)	-21.18*** (5.64)	-27.54*** (6.70)
6 or more times per year	-40.83*** (4.49)	-76.74*** (17.78)	-43.70*** (6.31)	-69.25*** (18.04)	-44.41*** (9.58)	-38.35*** (8.05)	-47.78*** (7.24)	-40.16*** (7.36)	-56.02*** (9.94)
8-14 days each time	-16.15*** (3.30)	-20.97*** (8.13)	-17.03*** (4.21)	-38.43*** (12.54)	-13.72** (5.80)	-2.57 (5.52)	-13.72*** (4.62)	-18.99** (5.42)	-16.69*** (5.63)
15 or more days each time	-31.87*** (4.19)	-38.69*** (10.74)	-32.85*** (5.33)	-52.18*** (14.29)	-17.73** (7.54)	-22.72*** (6.67)	-30.08*** (5.78)	-27.77*** (6.30)	-36.22*** (7.35)
2 weeks full firm payment	5.63*** (1.80)	14.14*** (5.30)	8.64*** (2.46)	9.12* (5.37)	3.65 (3.56)	4.58 (3.54)	7.84*** (2.68)	8.69*** (3.07)	5.69* (3.27)
2 weeks of full firm payment plus 15%	-3.99** (1.64)	-7.13 (4.40)	-4.44** (2.17)	-11.52** (5.38)	-1.76 (3.22)	-6.39** (3.24)	-3.28 (2.41)	-3.35 (2.73)	-7.93*** (3.06)

Notes: The table reports the implied marginal values of the attributes in terms of the wage. They are calculated as the ratio of the coefficient for the attribute and the coefficient for the wage from regressions of the job offer rate (see the notes to Table 2 for the details). The reference categories are the same as in Table 2. The standard errors are calculated using the delta method and the heteroskedastic robust covariance matrix. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.