CrossMark

What is the right profile for getting a job? A stated choice experiment of the recruitment process

Stefan Eriksson[1] · Per Johansson[2] ·
Sophie Langenskiöld[3,4]

Abstract We study the recruitment behaviour 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, work experience, ethnicity, religious beliefs, family situation, weight, and health, but otherwise have similar characteristics as the previous employee. 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. We also calculate the reduction in wage costs needed to make employers indifferent between applicants with and without these characteristics, and find that wage costs would have to be reduced by up to 50 % for applicants with some characteristics.

✉ Per Johansson
per.johansson@statistik.uu.se

[1] Department of Economics, Uppsala University, PO Box 513, 751 20 Uppsala, Sweden

[2] Department of Statistics, Uppsala University, IFAU and IZA, PO Box 51 751 20, Uppsala, Sweden

[3] Department of Health and Caring Sciences, Uppsala University, PO Box 564, 751 22 Uppsala, Sweden

[4] Department of Learning, Informatics, Management and Ethics, Medical Management Centre, Karolinska Institutet, 171 77 Stockholm, Sweden

🖉 Springer

1 Introduction

Labour 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 (OECD 2014a, b). Moreover, it is often claimed that workers who are old, have several children, are obese, or have a history of sickness absence are avoided by employers. In response to such concerns, most countries have introduced anti-discrimination laws and other policies to prevent some types of discrimination, but it has been difficult to find effective measures.

To explain the observed differences in labour market outcomes between groups, it is crucial to understand employers' hiring decisions. 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 (statistical discrimination). Such behaviour 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 productivity,[1] and where strict employment protection laws make it difficult to fire low productivity workers. In addition, the preferences of employers, employees, and customers concerning, e.g. gender and ethnicity, may affect the hiring decisions (taste-based discrimination).

Empirically, it is very difficult to show the existence of discrimination.[2] The traditional approach is to use administrative or survey data to analyse differences between groups. However, with 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, mostly in the form of correspondence studies.[3] In these studies, fictitious written 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 groups are compared, and any difference is interpreted as discrimination (Riach and Rich 2002; Bertrand and Mullainathan 2004; Carlsson and Rooth 2007). With this approach, the researcher has control over the applicant information observed by the firms, and hence, it is possible to estimate causal effects. Clearly, correspondence studies have increased our knowledge of discrimination, but this approach has some limitations: it is only possible to study discrimination in the initial stage of the hiring process (i.e. the invitation to a job interview decision), it is difficult to separate between taste-based and statistical discrimination,[4] and it has been argued that it is ethically questionable to

---

[1] 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.

[2] Altonji and Blank (1999) and Bertrand (2011) survey the discrimination literature.

[3] Another, more unusual, approach–called audit studies—uses fictitious job applicants (cf. the critique in Heckman 1998).

[4] Taste-based and statistical discrimination are described in Becker (1957), Phelps (1972), Arrow (1973) and Aigner and Cain (1977). The problem in the empirical studies is that detailed information about the characteristics of the recruiters is seldom available. Most correspondence studies only have information about the firms' sector and size and, in some cases, can guess the employers' gender and ethnicity from

subject employers to fictitious job search (Riach and Rich 2004).[5] Hence, there is still room for additional experimental methods to study discrimination.

The main purpose of this study is to use a stated choice experiment to investigate whether 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, i.e. if they discriminate.[6] Moreover, we quantify the degree of discrimination by calculating the reduction in wage costs needed to make employers indifferent between applicants with and without these characteristics. Finally, we interpret the discrimination in terms of taste-based and statistical discrimination.

Stated choice experiments are often used in transport, tourism, and environmental economics, but our study is the first which uses it to study employers' hiring behaviour.[7] In the experiment, which was conducted in the Swedish labour market, 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 similar characteristics as the previous employee, except for four attributes which are varied in each choice. The attributes varied include gender, age, education, work experience, ethnicity, religious beliefs, number of children, weight, history of sickness absence, the wage, and the type of firm co-payment in the sickness benefit system. The wage is varied to quantify the degree of discrimination in wage costs terms. The type of firm co-payment is varied to measure the effect of changing the cost of uncertainty in hiring. 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 the relevant interaction effects (cf. Sect. 2.2).

The stated choice approach shares some advantages with correspondence studies. Most importantly, we know which information that is 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. In addition, the stated choice approach has some additional advantages. First, we can consider 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

---

Footnote 4 continued

their name. For policymakers, the distinction between taste-based and statistical discrimination is crucial since the appropriate policy response is likely to vary depending on the nature of the discrimination.

[5] Another problem is that the estimates of the degree of discrimination may be affected by the choice of the fixed attributes in the applications. Both the choice of the attributes and their levels may affect the estimates of the degree of discrimination (Heckman and Siegelman 1993). Some recent studies try to handle these issues (Neumark 2012).

[6] 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.

[7] 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 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 costs terms.

decision. This is not possible in correspondence studies which, by construction, only consider the first stage of the hiring process. The fact that we vary many applicant characteristics simultaneously also means that we can compare the degree of the various types of discrimination. This is different from traditional (paired) correspondence studies, but is similar to some recent studies (e.g. Rooth 2011; Eriksson and Rooth 2014). Second, the fact that we include the wage as one of the attributes that are varied makes it possible to calculate the reduction in wage costs needed to make employers indifferent between applicants with and without a particular characteristic (all else equal).[8] Importantly, we can do this even if there is no flexibility in actual wage setting in these dimensions. In addition to these two main advantages, our approach gives us two new ways to try to distinguish between taste-based and statistical discrimination since we have variation in the cost of uncertainty in hiring (i.e. firm co-payment in the sickness benefit system) and know the characteristics of the recruiters. Finally, since all recruiters know that they participate in an experiment and are allowed to opt-out, there are no ethical concerns. There are two main limitations with the stated choice approach. The first is that it is based on stated rather than observed behaviour. A concern is, therefore, that the employers may not give answers which are consistent with their actual hiring behaviour; i.e. a hypothetical bias may arise. We use several methods, based on the stated choice literature and what we learned from our own pre-study to minimize this problem; e.g. the recruiters are asked to assess several attributes simultaneously. Also, since the 'politically correct' behaviour is to not discriminate (all employers in our pre-study stated that the never discriminated), we expect any remaining bias to reduce the estimates of discrimination. The second limitation is the risk of a low response rate since the experiment 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 make employers indifferent between applicants with and without these characteristics, wage costs would have to be reduced by up to 50 % for applicants with some characteristics.

The rest of the paper is organized as follows. The experimental design is outlined in Sect. 2. Section 3 presents the data. Section 4 provides a theoretical framework and describes the empirical modelling. Section 5 contains the results, and Sect. 6 concludes.

2 The experiment

2.1 The planning of the experiment

In the stated choice literature, it is strongly emphasized that an experiment should be proceeded by a thorough pre-study with interviews, focus groups, pilot question-

---

[8] It should be possible to vary the wage requirement in a correspondence study, but this is very unusual. Drydakis and Vlassis (2010) informally ask the employers responding positively to their fictitious applications about wage levels.

naires, etc. In particular, it is important that the participants perceive the experiment as realistic, i.e. has high face validity (Carson et al. 1994). To get a better understanding of the recruitment process, we started by interviewing a few experienced personnel managers. In particular, we discussed which factors they used to screen applications to choose whom to invite to job interviews, and how they assessed applicants in the interviews. We also discussed how familiar they were with the details of the sickness benefit system. Based on these interviews, we concluded that it would be possible to study the effects of gender, age, education, work experience, ethnicity, religious beliefs, number of children, weight, history of sickness absence, the wage, and the type of firm co-payment in the sickness benefit system. Four of these characteristics are mentioned in the Swedish Discrimination Act,[9] while the rest have been discussed in the policy debate. It should be emphasized that all managers stated that they never discriminated in their hiring decisions.

We then designed a pilot questionnaire and tested it in focus groups consisting of experienced personnel managers. The participants were first asked to fill in the questionnaire (i.e. choose between hypothetical applicants with different characteristics) and then to discuss its design. From these discussions, we concluded that: (i) the recruiters did remember the last employee who left the workplace voluntarily and which characteristics 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, and (iii) the recruiters indicated that they understood the combinations of attributes and the alternatives in the questionnaire, but that the choice became difficult if too many attributes had to be assessed simultaneously. This may be one explanation why recruiters seem to use signals to sort job applicants.

2.2 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 reference person). The choice always involved applicants with a full set of attributes; similar attributes as the previous employee except for the attributes that 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 work 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,

---

[9] The Swedish Discrimination 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. The factors excluded from the experiment were considered either as unobservable (transgender identity/expression and sexual orientation) or difficult to assess (disability).

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).[10] These are all factors which may be observed or discussed in a job interview, but are usually not included explicitly in a CV.[11] Except for gender, each of the twelve attributes has three possible alternatives (levels); cf. Sect. 2.3.

A problem is that it is not feasible to nonparametrically estimate the joint effect of all these attributes. The pre-study indicated that the respondents could handle four attributes in each choice (i.e. the choice between the 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.[12] However, the number of choices 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 seemed relevant according to our pre-study and previous studies. To simplify, we exclude all interactions between the wage and the other attributes. This implies that the wage distribution is independent of these attributes, which should be the case under the null hypothesis that there is no discrimination. We also assume that all three- and four-way interactions are zero (i.e. interactions between more than two attributes). This is reasonable since the marginal effects are likely to decrease as the order of the interactions increases.

In the questionnaire, we decided to include twelve choices for each respondent: four choices with respect to the invitation to a job interview and eight choices with respect to the job offer (more choices are needed in the hiring stage since more factors are varied). In each choice, the recruiter is asked to choose between two hypothetical applicants with different attributes; each applicant has similar attributes as the previous employee except for the four attributes which are varied. Each choice is made independently of the other choices; in particular, there is no link between the choices in the two stages of the hiring process.

To calculate the reduction in wage costs needed to make employers indifferent between applicants with and without a particular attribute, the wage is varied in all choices. Moreover, since we are interested in varying the firms' cost of uncertainty in hiring, the type of firm co-payment in the sickness benefit system is varied in all choices. The other two variables that are varied are: gender and age (choices 1 and 2), education and work experience (choices 3 and 4), ethnicity and religious beliefs (choices 5 and 6), gender and number of children (choices 7 and 8), gender and weight

---

[10] We did not include age, education, and work experience in the second stage for two reasons; our pre-study indicated that these factors should matter most in the first stage of the hiring process and we wanted to limit the number of characteristics.

[11] 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.

[12] To nonparametrically estimate all effects (including the three- and four-way interaction effects), we would need to include at least 81 (3^4) choices for each recruiter.

(choices 9 and 10), and frequency and intensity of sickness absence (choices 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 choice, i.e. the interaction effects between all applicant characteristics and the type of sickness benefit system, between gender and age, education and work experience, etc. To estimate these effects, we need to include 162 hypothetical applicants.[13] However, a few of the combinations are clearly not relevant to include. As an example, consider a choice where the attributes are gender, weight, the wage, and the type of firm co-payment in the sickness benefit system. In such a choice, we expect that most recruiters would choose a normal weight man with the lowest wage costs and the least expensive sickness benefit scheme. In the stated choice literature, it is emphasized that including such (unrealistic) choices can jeopardize the quality and credibility of the experiment. Therefore, we decided to drop six such combinations. We divided the remaining 156 hypothetical applicants into 13 versions of the questionnaire with 12 choices 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 choices, the questionnaire included detailed questions about the last employee who left the firm voluntarily, the recruiter, and the firm for which the recruiter worked (cf. Sect. 3.2 for descriptive statistics).

2.3 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 which appeared realistic to the employers. Table 1 presents the attributes and their levels. Most of these are very straightforward, but three deserve some additional justification. For weight, we used silhouettes originally developed by Stunkard et al. (1983). The silhouettes for men and women are illustrated in Appendix; silhouettes 1–2 represent underweight (not used), silhouettes 3–4 normal weight, silhouettes 5–6 overweight, and silhouettes 7–9 obese.[14] For the wage, we decided to relate it to the wage of the previous employee: the same wage, a 10 % lower wage, or a 10 % higher wage. We use relative rather than absolute wages since the wage level varies between jobs. Since we want to quantify the degree of discrimination in wage costs terms, it is important that we have enough variation in the wage alternatives, but the alternatives must also be perceived as realistic by the employers. Even if wages are compressed in Sweden, a 10 % lower/higher entry wage should be realistic in most cases. This was also confirmed by the recruiters in the

---

[13] The calculation of the number of hypothetical applicants required to estimate the relevant effects is based on statistical tables for fractional factorial experiments (cf. Hahn and Shapiro 1966).

[14] 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 likely to be inferior since the photographs usually are only showing the applicants' face and thus is unlikely to be tightly linked to the applicants' BMI.

Table 1 The attributes and their levels

Gender	Woman; man
Age	29 years or younger; 30–55 years; 56 years or older
Education	The lowest possible education (lowest quartile), average education (middle quartiles), and the highest education for the job in question (highest quartile)
Experience	4 years or less; 5–7 years; 8 years or more
Ethnicity	The Nordic countries; Europe except the Nordic countries; Africa/Middle East/South America
Religious belief	Christian; Jew; Muslim
Family situation	No children; one child; two or more children
Weight	Normal weight (3–4); overweight (5–6); obese (7–9)
Frequency of sickness absence	1–2 times per year; 3–5 times per year; 6 or more times per year
Intensity of sickness absence	7 or less days each time; 8–14 days each time; 15 or more days each time
Wage	10 % lower wage than the previous employee; the same wage as the previous employee; 10 % higher wage than the previous employee
Firm co-payment	3 weeks of full firm payment (day 2–21); 2 weeks of full firm payment (day 2–14); 2 weeks of full firm payment plus 15 % co-payment (day 2–21 + COFIN)

pre-study.[15] For the type of firm co-payment in the sickness benefit system, we used the three most recent schemes, i.e. two weeks of full firm payment, three weeks of full firm payment, and two weeks of full firm payment plus 15 % co-payment for the rest of the sickness absence. The firms' cost of uncertainty in hiring should be lowest in the first and highest in the last alternative. All these alternatives have been used in Sweden in the years before the experiment, have been widely discussed in the policy debate, and were well known and understood by all the recruiters taking part in the pre-study.[16]

Two examples of the choices are in Appendix.[17] In the first example (invitation to a job interview), gender, age, the wage, and the type of firm co-payment are assigned a level, and the recruiters are told that the applicants otherwise have similar attributes

---

[15]   In the pre-test, we used a 5 % lower/higher wage, but our discussion with the participants convinced us that we should increase the variation to potentially increase the efficiency in the estimation of the marginal values.

[16]   In Sweden, employers pay for the first two weeks of their employees' sick leave (no benefits the first day), then the Swedish Social Insurance Agency (SSIA) pays the benefits for the rest of the spell. 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 changed so that employers had to pay both the full sickness benefits for the first two weeks and 15 % of the sickness benefits for the rest of their employees' absence. To make the reform cost neutral for the employers, the payroll tax was reduced, but still the reform made it more costly to hire workers with a high risk of sickness absence. In 2006, the 15 % co-payment rule was abolished. Since then, the co-payment rules have not been changed. Employees get around 80 % of their previous wage up to an income cap and must submit a doctor's certificate after seven days.

[17]   An example of all the choices in a questionnaire is presented in Eriksson et al. (2012).

as the previous employee in all dimensions which are typically mentioned in a CV. In the second example (job offer), gender, weight, the wage, and the type of firm co-payment are assigned a level, and the recruiters are told that the applicants otherwise have similar attributes as the previous employee in all dimensions which are typically observed or discussed during a job interview.

2.4 Validity

A 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 hypothetical bias in the literature. The related method of contingent valuation has been severely criticized based on these arguments (Carson et al. 1996; Hausman 2012). However, a number of methods for reducing this bias have been suggested in the literature (List 2001; Murphy et al. 2005; Carson 2012; Kling et al. 2012), and several recent studies show that stated and revealed preferences often coincide (Murphy et al. 2010; Jacquemet et al. 2011). The results in these studies suggest that the importance of hypothetical bias depends on the experimental setting (Taylor et al. 2001); 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 (Ajzen et al. 2004; Guzman and Kolstad 2007). 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 (Carlsson and Martinsson 2001; Cameron et al. 2002; Backhaus et al. 2005; List et al. 2006).[18]

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

---

[18] A recent study by Deuchert et al. (2013) illustrates some of these points. They conduct an ex ante evaluation of Swiss firms' willingness to train disabled persons using a discrete choice stated preference design, where firms are asked whether they accept to train hypothetical disabled persons with a number of assigned attributes (profiles). The results show that 22 % 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 %. 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.

should be preceded by interviews, focus groups, and pilot questionnaires, that the respondents' incentives to answer truthfully should be thoroughly analysed, 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 (Hensher 2010).

In the experiment, we incorporate all these features. We conducted an elaborate pre-study analysis (i.e. interviews and focus groups) 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. answer that they consider both applicants similar) between the applicants in each choice. Also, they must consider several applicant attributes simultaneously, which makes it difficult to behave strategically. Finally, given the reasonable assumption that the 'politically correct' behaviour is to not discriminate (all recruiters in the pre-study stated that they never discriminated), we expect any remaining hypothetical bias to affect the results downwards (i.e. reduce the estimates of discrimination).

An objection against our design may be that we make the attributes that are varied more visible to employers than in a real-world hiring situation and that this risks that they respond more to these attributes than they normally would. However, the attributes that we vary are all highly visible in a CV and/or a job interview (except maybe religious beliefs). Moreover, the questions always stress that all applicant attributes that are not varied are the same as those of the previous employee. In addition, in the experiment, the recruiter considers several attributes simultaneously and then chooses the one with the best combined attributes. As a consequence, if all attributes are equally visible to the recruiter, this would not flaw an analysis of the relative effects of the attributes.

3 Data

3.1 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 2048 workplaces with one location and 20 or more employees. From this population, we drew a sample of 1000 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 workplaces.[19] The survey was administered by

---

[19] 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+ employees. For gender distribution, we used the strata: (i) <40 % women, (ii) 40–60 % women, and (iii) >60 % women.

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 behaviour of firms. The participants received no compensation for their participation, and two reminders were sent to non-respondents.

The response rate was around 46 %.[20] 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 behaviour.[21] In total, 426 employers (recruiters) are included in the analysis (this corresponds to 4895 observations).[22]

3.2 Descriptive statistics

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.

Around one-third of the recruiters who answered the questionnaire were managing directors, one-third personnel managers, and the rest held other positions tasked with personnel issues. Most of them worked with recruitment, 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. Nearly 80 % of them had 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 viewed themselves as 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 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. The majority had 8 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

---

[20] The response rate in the private and public sector was 44 and 50 %, respectively. For the size categories, the response rates were 43, 48, 56, 47, and 60 %. For the gender categories, the response rates were 46, 44, and 47 %.

[21] 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 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.

[22] Respondents who reported that no employee had left the firm are excluded from the analysis.

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 % were judged to be overweight or obese. Their mean wage was SEK 26,800 (€2900), and their median wage was SEK 25,000 (€2700).

Health and history of sickness absence

An important issue is if it is reasonable to assume that recruiting employers gather information about their job applicants' health and history of sickness absence. Half of the respondents claimed 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 even requesting the applicants to provide an excerpt from the Swedish Social Insurance Agency (SSIA). We therefore find it reasonable that the recruiters often have some information about their applicants' health and history of sickness absence.

4 Theoretical framework and empirical modelling

4.1 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 propensity to quit—prior to hiring. Hence, employers may find it optimal to base their hiring decisions on easily observed characteristics—e.g. ethnicity—which they believe are correlated with productivity (statistical discrimination). In addition, employers, employees, or customers may have preferences concerning different groups which affect the hiring decisions (taste-based discrimination). To capture the possibility of both types of discrimination, we assume that a recruiter in a firm is maximizing the expected value of the following utility function:

$$U(q, \mathbf{x}, y, e) = pf(q(\mathbf{x}), y) - w - d(\mathbf{x}, y, e), \tag{1}$$

where p is the price of the firm's product and f is the production function. Production depends on the worker's productivity, q, which is unobservable, and on idiosyncratic firm characteristics affecting productivity, y. Since q is unobservable, the recruiter uses the observable worker characteristics $\mathbf{x}$—a vector including education, work experience, gender, ethnicity, age, etc.—as a real or perceived signal of q. $d(\mathbf{x}, y, e) > 0$ is the recruiter's distaste function, where e is idiosyncratic firm characteristics affecting preferences. We assume that the wage, w, is exogenous. Since q is not observed, the recruiter maximizes the expected value of $U(q, \mathbf{x}, y, e | \mathbf{x}, y, e)$. Statistical discrimination is captured by the fact that $\mathbf{x}$ is used as a signal of q, and taste-based discrimination is captured by the distaste function.

4.2 Empirical modelling

Let us now consider a recruiter who participates in the experiment. A previous employee is supposed to be replaced by one of two hypothetical applicants with observed attributes $\mathbf{z_1} = (\mathbf{x_1}, w_1, I_1)$ and $\mathbf{z_2} = (\mathbf{x_2}, w_2, I_2)$, where I_i, $i = 1, 2$, is the type of firm co-payment in the sickness benefit system. Assume that $q(\mathbf{x}) = q(\omega, \mathbf{x^0})$, where $\mathbf{x^0}$ is the attributes of the previous employee that we are varying (i.e. $\mathbf{x_1}$ and $\mathbf{x_2}$), while ω is the, for us, unobserved terms. The previous employee with observed attributes (and known productivity) is then completely described by the vector $\boldsymbol{v^0} = (\omega, \mathbf{z^0})$, where $\mathbf{z^0} = (\mathbf{x^0}, w^0, I^0)$, w^0 is the wage, and I^0 is the institutions at the time of employment.

In each choice, the recruiter is asked to choose between applicants 1 and 2. He/she will choose applicant 1 if the expected value of choosing applicant 1 is larger than the expected value of choosing applicant 2. In the experiment, we stress that all other attributes are the same as of the previous employee, i.e. ω is held constant and hence is equal for both applicants. This enables us to remove ω since we can view the recruiter's utility functions for applicants 1 and 2 as local approximations of the recruiter's utility function for the previous employee.

To be specific, let $U(\omega, \mathbf{z_j})$ be the utility function of applicant $j = 1, 2$. A first-order Taylor expansion of $U(\omega, \mathbf{z_j})$ around $\mathbf{z^0}$ is given by $U(\omega, \mathbf{z_j}) \approx U(\boldsymbol{v^0}) + (\mathbf{z_j} - \mathbf{z^0})U_z(\boldsymbol{v^0})$, where $j = 1, 2$ and $U_z(\boldsymbol{v^0})$ is the derivative of the utility function with respect to the vector $\mathbf{z}$. The recruiter, hence, chooses applicant 1 if $E(U(\omega, \mathbf{z_1}) - U(\omega, \mathbf{z_2})) > 0$, where the expectation is taken over ω. Since $U(\boldsymbol{v^0})$ is constant between the different choices, and the variables are fixed by design, this means that applicant 1 is chosen if $(\mathbf{z_1} - \mathbf{z_2})E(U_z(\boldsymbol{v^0})) > 0$. Given our experimental design, the difference in utility for the comparison is equal to:

$$
\begin{aligned}
E(U(\omega, \mathbf{z_1}) - U(\omega, \mathbf{z_2})) &= (\mathbf{x_1} - \mathbf{x_2})E(U_x(\boldsymbol{v^0})) + (w_1 - w_2)E(U_w(\boldsymbol{v^0})) \\
&+ (I_1 - I_2)E(U_I(\boldsymbol{v^0})) + (I_1\mathbf{x_1} - I_2\mathbf{x_2})E(U_{xI}(\boldsymbol{v^0})) + (\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_2)E(U_{\tilde{x}\tilde{x}}(\boldsymbol{v^0})),
\end{aligned}
\tag{2}
$$

where $\tilde{\mathbf{x}}_\mathbf{j}$ is the small subset of covariates that potentially could have interactions effects (the interactions included are those between gender and age, education and work experience, ethnicity and religious beliefs, gender and number of children, gender and weight, and frequency and intensity of sickness absence; all other covariates are by design orthogonal to the other covariates). $U_r(\boldsymbol{v^0})$ is the derivative of the utility function with respect to $r = \mathbf{x}, w, I, xI$ and $\tilde{x}\tilde{x}$.

By taking the expectation over the whole population of recruiters, we hence get:

$$
\begin{aligned}
EE(U(\omega, \mathbf{z_{1g}}) - U(\omega, \mathbf{z_{2g}})) &= (\mathbf{x_{1g}} - \mathbf{x_{2g}})'\boldsymbol{\alpha} + (w_{1g} - w_{2g})\beta + (I_{1g} - I_{2g})'\boldsymbol{\gamma} \\
&+ (\mathbf{x_{1g}}I_{1g} - \mathbf{x_{2g}}I_{2g})'\boldsymbol{\delta} + \tilde{\mathbf{x}}_{1g}\tilde{\mathbf{x}}_{1g} - \tilde{\mathbf{x}}_{2g}\tilde{\mathbf{x}}_{2g})'\boldsymbol{\rho},
\end{aligned}
\tag{3}
$$

where g denotes the choice. The parameters are defined as the average response in the population of recruiters, hence $\boldsymbol{\alpha} = EE(U_x(\boldsymbol{v^0}))$, $\beta = EE(U_w(\boldsymbol{v^0}))$, $\boldsymbol{\gamma} = EE(U_I(\boldsymbol{v^0}))$, $\boldsymbol{\delta} = EE(U_{xI}(\boldsymbol{v^0}))$ and $\rho = EE(U_{\tilde{x}\tilde{x}}(\boldsymbol{v^0}))$, which follows from the ran-

domization of $\mathbf{x}$, w and I being independent of ω, $\mathbf{x^0}$, w^0 and I^0. Note that equation (3) includes all relevant interaction effects and that the wage, by design, is made orthogonal to the attributes $\mathbf{x_{ig}}$ and I_{ig}. From this specification, the marginal value of attribute x_k can be calculated as the ratio of parameters, hence:

$$\frac{\partial w}{\partial x_k} = \frac{\partial EEU(\omega, \mathbf{z})/\partial x_{igk}}{\partial EEU(\omega, \mathbf{z})/\partial w_{igk}} = \frac{\alpha_k + I\delta_k}{-\beta}. \tag{4}$$

Note that if $\delta_k = 0$, we get the marginal value of attribute x_k when there is no interaction effect between attribute x_k and the type of firm co-payment. Henceforth, we refer to the case where all interaction effects are zero as the baseline model.[23] It should be emphasized that we can calculate the 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):

$$\begin{aligned} Y_{1eg} - Y_{2eg} = (\mathbf{x_{1g}} - \mathbf{x_{2g}})'\boldsymbol{\alpha} + (w_{1g} - w_{2g})\beta + (I_{1g} - I_{2g})'\boldsymbol{\gamma} \\ + (\mathbf{x_{1g}}I_{1g} - \mathbf{x_{1g}}I_{2g})'\boldsymbol{\delta} + (\mathbf{\tilde{x}_{1g}\tilde{x}_{1g}} - \mathbf{\tilde{x}_{2g}\tilde{x}_{2g}})'\rho + \eta_{1eg} - \eta_{2eg}, \end{aligned} \tag{5}$$

where Y is the outcome variable and e denotes the recruiter/employer. In each choice, the recruiter is asked to choose one of the two applicants or to opt-out.[24] If the respondent chooses an applicant, the dependent variable is either 1 or -1 if the first or the second applicant is chosen. We cluster the standard errors at the recruiter/employer level. We do not include fixed effects, but get very similar results if we do.[25] An alternative estimation method would be to use the logit model. We prefer to use the linear model rather than the logit model since the design is orthogonal for the linear model only. However, we have also estimated logit models (with and without fixed effects) and, qualitatively, the results are very similar.[26]

5 Results

In this section, we analyse and interpret the degree of discrimination. 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). When comparing the magnitude of our estimates to the (callback) estimates in correspondence studies, it should be noted that our baseline probability is 100 % (i.e. a recruiter can only choose one of the two hypothetical applicants), whereas in a correspondence study this baseline probability is typically around 10 % (i.e. a recruiter can choose another (real) applicant). Therefore, a comparison must be based on the relative effects. We

---

[23] i.e. the interaction effects between attribute k and the type of firm co-payment are set to zero.

[24] In each of the choices, around 30 % of the respondents opted out. The results are not affected if we include these cases in the estimation.

[25] These results are available upon request.

[26] These results are available upon request.

then calculate the marginal value of each attribute; i.e. the reduction in wage costs needed to make employers indifferent between applicants with and without a particular attribute. Finally, we interpret the discrimination in terms of taste-based and statistical discrimination.

5.1 Invitation to job interviews

In Table 2, we present 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 % points lower than the callback rate for an applicant who is less than 30 years old. The callback rate for a 30- to 55-year-old applicant is 13 % 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 work experience have the expected effects, i.e. a higher

Table 2 Callback rate to a job interview

Gender (ref: male):	
Female	0.00
	(0.04)
Age (ref: 29 years or younger):	
30–55 years	0.13**
	−0.06
>55 years	−0.64***
	(0.06)
Education (ref: lowest quartile):	
Middle quartiles	0.63***
	(0.06)
Highest quartile	0.83***
	(0.05)
Experience (ref: <5 years):	
5–7 years	0.13**
	(0.06)
>7 years	0.13**
	(0.05)
Wage:	
Wage	−0.01***
	(0.00)
Sickness benefits (ref: 3 weeks of full firm payment):	
2 weeks of full firm payment	0.09**
	(0.04)
2 weeks of full firm payment plus 15 % co-payment	−0.08*
	(0.04)

Estimated with ordinary least squares. The dependent variable is the probability of being invited to a job interview (the callback rate). Also included is a constant. Standard errors clustered at the recruiter/employer level are in parentheses. ***, **, * denote statistical significance at the 1, 5 and 10 % levels, respectively

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 83 % 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 % points, whereas combining the same reduction with a 15 % employer co-payment for the complete sickness spell would decrease the callback rate with 8 % 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.

Our results confirm the results in previous correspondence studies; Ahmed et al. (2012) find strong evidence of age discrimination (the relative effect is 69 %), and Carlsson (2011) finds no evidence of gender discrimination. The results also suggest that more firm co-payment in the sickness benefit system may have a negative effect on hiring.

5.1.1 Job offers

In the first column in Table 3, we present the estimated effects of the applicants' attributes on the probability of offering a job (the job offer rate).

Again, we 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 % points). Applicants who are Muslim or Jewish also have a much lower job offer rate than applicants who are Christian (minus 29 and 26 % points). A second striking result is the very large negative effect for obese applicants: being obese decreases the job offer rate by 83 % points compared to having normal weight. Moreover, applicants with two or more children have a 25 % points lower job offer rate.[27] There is also a lower job offer rate for workers with a history of sickness absence, especially for those with many spells and long durations. Also, the wage has a negative effect. Finally, the type of firm co-payment in sickness benefits has the expected effect; i.e. if the employers' costs increase, the job offer rate decreases.

Our results confirm the results in previous correspondence studies, but also extend these results by considering attributes which have not been analysed before. Previous studies show that non-European immigrants and Muslims face widespread discrimination in the first stage of the hiring process (e.g. Carlsson and Rooth 2007, find a relative effect of 50 %).[28] The negative effect for Jewish applicants is somewhat more surprising. However, recently anti-Semitism and anti-discrimination against Jews have received a lot of attention in Sweden.[29] It should also be noted that we measure the

---

[27] We have analysed whether this effect is different for men and women, but do not find that the interaction effect is significant.

[28] Evidence of ethnic discrimination is also reported in Bertrand and Mullainathan (2004), and the studies cited in Riach and Rich (2002). Drydakis (2010) studies the effects of religious affiliation.

[29] This is discussed in, for example, US Department of State (2005) and Bachner and Ring (2005).

Table 3 Job offer rate and the marginal value of each attribute

	Job offer rate	Marginal value
Gender (ref: male):		
Female	0.00	0.12
	(0.03)	(1.95)
Ethnicity (ref: Nordic):		
Other European	−0.02	−0.92
	(0.06)	(3.74)
Africa, Middle East, South America	−0.28***	−16.28***
	(0.06)	(4.10)
Religious belief (ref: Christian):		
Jewish	−0.26***	−14.85***
	(0.06)	(3.44)
Muslim	−0.29***	−17.05***
	(0.07)	(4.51)
Children (ref: no children):		
1 child	−0.09	−5.08
	(0.06)	(3.67)
2 or more children	−0.25***	−14.62***
	(0.06)	(3.80)
Weight (ref: normal weight):		
Overweight	−0.13**	−7.58**
	(0.06)	(3.50)
Obese	−0.83***	−48.05***
	(0.04)	(5.38)
Frequency of sickness absence (ref: 1-2 times per year):		
3-5 times per year	−0.41***	−23.90***
	(0.04)	(3.32)
6 or more times per year	−0.84***	−48.30***
	(0.04)	(4.86)
Intensity of sickness absence (ref: 7 or less days):		
8-14 days each time	−0.29***	−16.73***
	(0.04)	(2.93)
15 or more days each time	−0.55***	−31.79***
	(0.04)	(5.48)
Wage:		
Wage	−0.02***	–
	(0.00)	

Table 3 continued

	Job offer rate	Marginal value
Sickness benefit (ref: 3 weeks of full firm payment):		
2 weeks of full firm payment	0.13***	7.30***
	(0.03)	(1.83)
2 weeks of full firm payment plus 15% copayment	−0.09***	−5.02***
	(0.03)	(1.76)

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 is a constant. 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. Standard errors clustered at the recruiter/employer level are in parentheses (in the second column calculated using the delta method). ***, **, * denote statistical significance at the 1, 5 and 10 % levels, respectively

effect given that the recruiters know that an applicant has a certain characteristic, and religious beliefs are probably less likely to be observed than the other characteristics. It is also striking that applicants who are obese, have a history of sickness absence, and/or have several children face substantially lower job offer rates. These factors have received much less attention in correspondence studies,[30] probably because they are typically not mentioned in a CV. Finally, more firm co-payment in the sickness benefit system again seems to have a negative effect on hiring.

5.2 Quantifying the degree of discrimination in wage costs terms

In the second column in Table 3, we have the estimates measured as the marginal value in wage costs terms of each attribute; i.e. the reduction in wage costs needed to make employers indifferent between applicants with and without a particular attribute (all else equal). The results suggest that to eliminate the negative hiring-effect wage costs would need to be reduced by 16 % for an applicant born in Africa, the Middle Eastern countries, or South America compared with an applicant born in the Nordic countries, by around 17 (15) % for a Muslim (Jewish) applicant compared with a Christian applicant, by around 8 (48) % for an overweight (obese) applicant compared with a normal weight applicant, and by up to 48 % for an applicant with a history of sickness absence.

To get a sense of the magnitude of these effects, it is illustrative to re-calculate them as wage levels. Since the mean of the previous employees' monthly wage is SEK 26,800 (€2,900), a 15 % reduction would correspond to SEK 4,000 (€435) lower wage costs and a 48 % reduction to SEK 12,900 (€1,400) lower wage costs. The exact numbers should of course be interpreted with caution (in particular because they are based on an extrapolation of a rather small difference in wage demands in the experiment), but they indicate that the wage costs reductions needed to make employers indifferent between applicants are substantial.

---

[30] This is similar to the finding in Rooth (2009).

5.3 Taste-based or statistical discrimination

The discrimination we find evidence of may reflect both taste-based and statistical discrimination; discrimination against, for example, ethnic and religious minorities may reflect both types of discrimination, while discrimination against, for example, workers with several children, obesity, or a history of sickness absence is likely to reflect statistical discrimination.

Our approach gives us two new ways to try to distinguish between taste-based and statistical discrimination. First, the design of the experiment allows us to investigate what happens with the degree of discrimination if we vary the extent of firm co-payment in the sickness benefit system. If statistical discrimination is important, we expect that when the extent of firm co-payment increases, employers should be less likely to hire both applicants with attributes that signal a high risk of sickness absence, and, if they are risk averse, applicants who have other characteristics that signal uncertainty about total labour costs. To investigate this, we interact the type of firm co-payment variables with the attributes signalling 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). The results show that there is little evidence of any systematic relationship between the degree of discrimination and the extent of firm co-payment.[31] This may be interpreted as evidence against the importance of statistical discrimination, but it may also reflect 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. In addition, firms may consider other negative effects associated with hiring the 'wrong' worker as more important than costs associated with sickness absence. Therefore, it may be that the (realistic) changes in co-payment that we analyse are simply too limited to affect the employers' hiring decisions.

Second, we analyse whether the degree of discrimination is similar in all firms or if it differs depending on the type of recruiter and/or firm; i.e. we run the baseline regression on subsamples of various subgroups. For the recruiter characteristics (e.g. gender, ethnicity, and age), our results indicate that different types of recruiters treat applicants with different attributes in a rather similar way; e.g. the effect of the ethnicity is very similar among ethnic majority and minority recruiters.[32] This is more supportive of statistical discrimination than taste-based discrimination since, if taste-based discrimination is important, we would expect to find clear differences in recruitment between different types of recruiters (cf. Åslund et al. 2014). For the firm characteristics (e.g. sector, size, and gender composition), we find that recruiters in the different types of workplaces treat applicants with different attributes in a rather similar way, except that employers in large firms seem less likely to discriminate than employers in small firms.[33] Again, this may be interpreted as supportive of statistical discrimination since

---

[31] Only a few of the interactions are statistically significant, and there is no systematic pattern in these effects. The details of the estimation are presented in Eriksson et al. (2012).

[32] These results are presented in detail in Eriksson et al. (2012).

[33] These results are presented in detail in Eriksson et al. (2012).

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 taste-based discrimination may be bigger in small firms where the recruiter and the employees interact more closely in the day-to-day operations. Another explanation is that, since large firms often have human resources departments, they may be more aware of anti-discrimination laws.

Overall, our results suggest that both types of discrimination exist, but that statistical discrimination may be more important.

6 Conclusions

Labour market discrimination is a major issue in the policy debate. Many policy attempts have been tried to reduce discrimination, but it has been difficult to find effective policy measures. In recent years, correspondence studies have extended our knowledge of discrimination, but still there is room for additional experimental approaches to study discrimination.

In this study, we use a stated choice experiment to study whether Swedish employers use information about job applicants' gender, age, ethnicity, religious beliefs, number of children, weight, or history of sickness absence when they recruit workers. 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: similar characteristics as the previous employee except for the four attributes that are varied in each choice.

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 discrimination studies, but our results also extend the existing literature in several 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 reduction in wage costs needed to make employers indifferent between applicants with and without a particular attribute (all else equal). The exact numbers should of course be interpreted with caution, but our estimates indicate that substantial wage costs differentials—from 10 to 50 %—are needed to compensate employers.

The discrimination we find evidence of may reflect both taste-based and statistical discrimination. Discrimination based on, for example, ethnicity and religious beliefs may reflect both types of discrimination, while discrimination based on e.g. previous sickness absence, obesity, and 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). Our results indicate that both taste-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 costs 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 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 costs 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 subsides would need to be introduced at least until the true productivity can be verified. 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 (employers can get up to an 80 % wage costs subsidy)—have received very limited interest.

Our study demonstrates that stated choice experiments may be a valuable complement to the existing approaches to study firms' hiring behaviour. This approach has several advantages, such as giving us a feasible way to study applicant characteristics which may be relevant in any of the stages of the hiring process and to measure how firms' value different applicant characteristics in wage costs terms. 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 behaviour 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 and Schaefer 2011). This approach may also be useful for ex ante evaluations of policy reforms (cf. Deuchert et al. 2013). 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 that 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.

Overall, our results suggest that discrimination is prevalent in both stages of the hiring process, that the magnitude of the discrimination in terms of the reductions in wage costs needed to make employers indifferent between applicants with and without some attributes are substantial, and that many of our results are consistent with statistical discrimination.
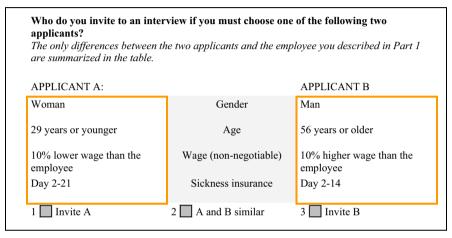
Appendix: Two examples of the choices in the questionnaire

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

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. The weight scale is given below:

APPLICANT A: APPLICANT B

APPLICANT A		APPLICANT B
Man	Gender	Woman
7-9	Weight	3-4
10% higher wage than the employee	Wage (non-negotiable)	Same wage as the employee
Day 2-14 + COFIN	Sickness insurance	Day 2-21

1 ☐ Hire A   2 ☐ A and B similar   3 ☐ Hire B

References

Ahmed A, Andersson L, Hammarstedt M (2012) Does age matter for employability? A field experiment on ageism in the Swedish labor market. Appl Econ Lett 19:403–406

Aigner DJ, Cain GG (1977) Statistical theories of discrimination in labor markets. Ind Lab Relat Rev 30:175–187

Ajzen I, Brown TC, Carjaval F (2004) Explaining the discrepancy between intentions and actions: the case of hypothetical bias in contingent valuation. Personal Soc Psychol B 30:1108–1120

Altonji JG, Blank RM (1999) Race and gender in the labor market. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol 3B. Elsevier, Amsterdam

Arrow KJ (1973) The theory of discrimination. In: Ashenfelter O, Rees A (eds) Discrimination in labor markets. Princeton University Press, Princeton

Åslund O, Hensvik L, Nordström Skans O (2014) Seeking similarity: how immigrants and natives manage at the labor market. J Lab Econ 32:405–441

Backhaus K, Wilken R, Voeth M, Sichtmann C (2005) An empirical comparison of methods to measure willingness to pay by examining the hypothetical bias. Int J Mark Res 47:543–562

Bachner H, Ring J (2005) Antisemitic images and attitudes in Sweden. The Living History Forum, Stockholm

Becker G (1957) The economics of discrimination. Chicago University Press, Chicago

Berkey CS, Frazier AL, Gardner JD, Colditz GA (1999) Adolescence and breast carcinoma risk. Cancer 1999:2400–2409

Bertrand M (2011) New perspectives on gender. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol 4B. Elsevier, Amsterdam

Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. Am Econ Rev 94:991–1013

Bulik C, Wade T, Heath A, Martin N, Stunkard A, Eaves L (2001) Relating body mass index to figural stimuli: population-based normative data for Caucasians. Int J Obes 10:1517–1524

Cameron TA, Poe GL, Ethier RG, Schulze WD (2002) Alternative non-market value-elicitation methods: are the underlying preferences the same? J Environ Econ Manag 44:391–425

Carlsson M (2011) Does hiring discrimination cause gender segregation in the Swedish labor market? Fem Econ 17:71–102

Carlsson F, Martinsson P (2001) Do hypothetical and actual marginal willingness to pay differ in choice experiments? J Environ Econ Manag 41:179–192

Carlsson M, Rooth DO (2007) Evidence of ethnic discrimination in the Swedish labor market using experimental data. Lab Econ 14:716–729

Carson RT (2012) Contingent valuation: a practical alternative when prices aren't available. J Econ Perspect 26:27–42

Carson R, Flores E, Martin K, Wright J (1996) Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Econ 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, Wiley J (1994) Experimental analysis of choice. Mark Lett 5:351–368

Deuchert E, Kauer L, Meisen Zannol F (2013) Would you train me with my mental illness? Evidence from a discrete choice experiment. J Ment Health Policy 16:67–80

Drydakis N (2010) Religious affiliation and labour bias. J Sci Study Relig 49:472–488

Drydakis N, Vlassis M (2010) Ethnic discrimination in the Greek labour market: occupational access, insurance coverage and wage offers. Manch Sch 78:201–218

Eriksson S, Rooth DO (2014) Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. Am Econ Rev 104:1014–1039

Eriksson S, Johansson P, 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

Guzman R, Kolstad C (2007) Researching preferences, valuation and hypothetical bias. Environ Resour Econ 37:465–487

Hahn GJ, Shapiro SS (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

Hausman J (2012) Contingent valuation: from dubious to hopeless. J Econ Perspect 26:43–56

Heckman J (1998) Detecting discrimination. J Econ Perspect 12:101–116

Heckman J, Siegelman P (1993) The Urban Institute audit studies: their methods and findings. In: Fix ME, Struyk RJ (eds) Clear and convincing evidence: measurement of discrimination in America. Urban Institute, Washington, DC, pp 187–258

Hensher DA (2010) Hypothetical bias, choice experiments and willingness to pay. Transp Res B Methodol 44:735–752

Jacquemet N, Joule RV, Luchini S, Shogren JF (2011) Do people always pay less than they say? Testbed laboratory experiments with IV and HG values. J Public Econ Theory 13:857–882

Kling CL, Phaneuf DJ, Zhao J (2012) From Exxon to BP: has some number become better than no number? J Econ Perspect 26:3–26

List JA (2001) Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards. Am Econ Rev 91:1498–1507

List JA, Sinha P, Taylor M (2006) Using choice experiments to value non-market goods and services: evidence from field experiments. Adv Econ Anal Policy 6:1–37

Munoz KA, Ballard-Barbash R, Graubard B, Swanson CA, Schairer C, Kahle LL (1996) Recall of body weight and body size estimation in women enrolled in the breast cancer detection and demonstration project (BCDDP). Int J Obes Relat Metab Disord 1996:854–859

Murphy J, Allen P, Stevens T, Weatherhead D (2005) A meta-analysis of hypothetical bias in stated preference valuation. Environ Resour Econ 30:313–325

Murphy J, Stevens T, Yadav L (2010) A comparison of induced value and home-grown value experiments to test for hypothetical bias in contingent valuation. Environ Resour Econ 47:111–123

Neumark D (2012) Detecting discrimination in audit and correspondence studies. J Hum Res 47:1128–1157

OECD (2014a) Employment outlook 2014. OECD, Paris

OECD (2014b) International migration outlook 2014. OECD, Paris

Oyer P, Schaefer S (2011) Personnel economics: hiring and incentives. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol 4B. Elsevier, Amsterdam

Phelps ES (1972) The statistical theory of racism and sexism. Am Econ Rev 62:659–661

Riach P, Rich J (2002) Field experiments of discrimination in the marketplace. Econ J 112:480–518

Riach P, Rich J (2004) Deceptive field experiments of discrimination: are they ethical? Kyklos 57:457–470

Rooth DO (2009) Obesity, attractiveness and differential treatment in hiring: a field experiment. J Hum Res 44:710–735

Rooth DO (2011) Work out or out of work: the labor market return to physical fitness and leisure sports activities. Lab Econ 18:399–409

Stunkard AJ, Sorensen T, Schulsinger F (1983) Use of the Danish Adoption Register for the study of obesity and thinness. In: Kety SS, Rowland LP, Sidman RL, Mathysse SW (eds) The genetics of neurological and psychiatric disorders. Raven Press, New York, pp 115–120

Taylor L, Mckee OM, Laury SK, Cummings RG (2001) Induced value tests of the referendum voting mechanism. Econ Lett 71:61–65

US Department of State (2005) Report on global anti-semitism. US Department of State, Washington, DC

Van Beek KWH, Koopmans CC, van Praag BMS (1997) Shopping at the labour market: a recent tale of fiction. Eur Econ Rev 41:295–317