

Eriksson, Stefan; Lagerström, Jonas

**Working Paper**

## Detecting discrimination in the hiring process: Evidence from an internet-based search channel

Working Paper, No. 2007:19

**Provided in Cooperation with:**

IFAU - Institute for Evaluation of Labour Market and Education Policy, Uppsala

*Suggested Citation:* Eriksson, Stefan; Lagerström, Jonas (2007) : Detecting discrimination in the hiring process: Evidence from an internet-based search channel, Working Paper, No. 2007:19, Institute for Labour Market Policy Evaluation (IFAU), Uppsala

This Version is available at:

<https://hdl.handle.net/10419/78677>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



IFAU – INSTITUTE FOR  
LABOUR MARKET POLICY  
EVALUATION

# **Detecting discrimination in the hiring process: evidence from an Internet- based search channel**

Stefan Eriksson  
Jonas Lagerström

WORKING PAPER 2007:19

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out: evaluations of the effects of labour market policies, studies of the functioning of the labour market, evaluations of the labour market effects of measures within the educational system and evaluations of the effects of the social insurance on the labour market. Besides research, IFAU also works on: spreading knowledge about the activities of the institute through publications, seminars, courses, workshops and conferences; influencing the collection of data and making data easily available to researchers all over the country.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The authority has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council is responsible for giving proposals for decisions about what research grants should be approved. A reference group including representatives for employers and employees as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala

Visiting address: Kyrkogårdsgatan 6, Uppsala

Phone: +46 18 471 70 70

Fax: +46 18 471 70 71

[ifau@ifau.uu.se](mailto:ifau@ifau.uu.se)

[www.ifau.se](http://www.ifau.se)

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

# Detecting discrimination in the hiring process: evidence from an Internet-based search channel\*

by

Stefan Eriksson<sup>a</sup> and Jonas Lagerström<sup>b</sup>

August 25, 2007

## Abstract

This paper uses data from an Internet-based CV database to investigate how factors which may be used as a basis for discrimination, such as the searchers' ethnicity, gender, age and employment status, affect the number of contacts they receive from firms. Since we have access to essentially the *same* information as the firms, we can handle the problems associated with unobserved heterogeneity better than most existing studies of discrimination. We find that, even when we control for all other differences, searchers who have non-Nordic names, are old or unemployed receive significantly fewer contacts. Moreover, we find that this matters for the hiring outcome: Searchers who receive more contacts have a higher probability of actually getting hired.

Keywords: Job Search, Unobserved Heterogeneity, Discrimination

JEL-codes: J64, J71

---

\*We are grateful for comments from Per-Anders Edin, Erik Grönqvist, Tuomas Pekkarinen, Dan-Olof Rooth, Peter Skogman-Thoursie, and seminar participants at the ESPE Annual Conference in Chicago, the Nordic Summer Institute in Labour Economics, the Institute for Labour Market Policy Evaluation (IFAU) and Växjö University. Thanks to Eva Granath, Anders Wellman and AMS for providing the data. Financial support from the Institute for Labour Market Policy Evaluation (IFAU) is gratefully acknowledged.

<sup>a</sup> Department of Economics, Uppsala University, PO Box 513, SE-751 20 Uppsala, Sweden, Stefan.Eriksson@nek.uu.se.

<sup>b</sup> Institute for Labour Market Policy Evaluation (IFAU), PO Box 513, SE-751 20 Uppsala, Sweden, jonas.lagerstrom@abo.fi.

## Table of contents

1	Introduction .....	3
2	Data.....	6
3	The number of contacts received.....	13
3.1	Identification.....	13
3.2	Variables and estimation .....	13
3.3	Results .....	15
3.4	Robustness .....	21
4	The probability of getting hired.....	23
4.1	Identification and estimation .....	24
4.2	Results .....	25
5	Conclusions .....	27
	References.....	29
	Appendix 1: The questionnaire.....	31
	Appendix 2: Comparison of the characteristics of the searchers .....	33
	Appendix 3: The probability of receiving a contact .....	36

# 1 Introduction

The emergence of the Internet has changed the nature of job search. In addition to the traditional job search channels, a person who wants to find a new job now has the opportunity to register his or her CV on websites where employers search for prospective employees. The emergence of Internet-based search channels introduces new opportunities for researchers to investigate how the job seekers' characteristics affect their job search success, which can help us to uncover discrimination. Much of the existing literature on discrimination can be criticized by arguing that the results do not reflect actual discrimination, but instead are explained by factors which are observable to the firms but unobservable to the researcher. Thus the issue of unobserved heterogeneity is often crucial. Using data from Internet-based search channels can help us solve this problem, since the initial contact between job seeker and firm is based only on the information in the CV. Thus in principle the researcher will have access to the *same* information as the firms, and given that this data is handled correctly, can obtain better estimates of discrimination in the initial stages of the hiring process than most existing studies.

The purpose of this paper is to study how factors which may be used as a basis for discrimination, such as the job searchers' ethnicity, gender, age and employment status, affect the number of contacts they receive from firms using data from a Swedish Internet-based CV database. We also investigate if searchers who receive contacts through this search channel have a higher probability of getting hired; i.e. if the number of contacts received matter for the actual hiring outcome.

Firms' hiring decisions are often difficult, since firms rarely have access to full information on how productive the applicants will be if they are hired. Instead, employers use all available information to predict how productive the job seekers are. This typically includes information on education, labor market experience and other factors directly affecting productivity. However, there are many other factors that may be equally important, such as motivation and social skills, but these factors are often impossible to observe prior to hiring. Therefore, we may expect that employers use other easily observable factors that they believe are correlated with those unobservable characteristics as sorting criteria. These factors may include ethnicity, gender, age and employment status. Such behavior from information-constrained firms may thus give rise to statistical discrimination against e.g. immigrants, women, older

searchers and the long-term unemployed. In addition, searchers may experience preference-based discrimination based on e.g. ethnicity.

To empirically investigate whether such discrimination occurs, data is needed on the search activities and outcomes for a large number of job seekers. In addition, control variables have to be included for all other factors that affect the firms' hiring decisions. If this is not possible, it is difficult to know whether the results should be attributable to discrimination or to the unobservable factors; i.e. the issue of unobserved heterogeneity is crucial. Most existing studies of discrimination cannot claim to have access to the *same* information as the firms have, and thus can be criticized on these grounds.

This paper uses data from 'My CV' which is an Internet-based search channel provided by the Swedish Public Employment Office since the late 1990s. Anyone who wants to find a job is invited to submit details — personal characteristics and requirements about the jobs — over the Internet to the database. Employers are then invited to search in this database for applicants that they find interesting, and can contact them for interviews etc. by e-mail within the system. The data covers all applicants remaining as active searchers in December 2004 who agreed to participate in a research project on the recruitment behavior of firms. The sample we use includes 18 167 active searchers. The dataset includes the information that the job seekers have entered about themselves, the number of contacts they have received from firms, and the number of contact attempts they have responded to. This data is combined with data on hiring from the Employment Office.

The dataset has several advantages. First and most importantly, we know that we have access to essentially the *same* information as the employers have when they choose which applicants to contact.<sup>1</sup> This means that, given that we insert properly defined control variables for all other relevant characteristics, we can isolate the effects of the characteristics that we are interested in. Second, the sample is large (compared with previous work) and includes a diverse pool of searchers. Also, in contrast to studies based on data from field experiments such as e.g. Bertrand & Mullainathan (2004), our data reflects genuine job search and the types of jobs covered have not been chosen by us. Third, even though our focus is on the initial stages of the hiring process, we can also study whether the number of contacts received matter for the hiring outcome; i.e. if a high contact rate increases the hiring probability.

---

<sup>1</sup> We have access to all information except for the content of the personal letters; see Section 2.

We start by analyzing how the number of contacts received is affected by the job seekers' characteristics. In the regression analysis, we include control variables for all observable characteristics. We find that searchers from some ethnic groups, older searchers and searchers who are unemployed (both openly and participating in labor market programs) have a significantly lower contact rate. Also, we find that the results differ somewhat across different subgroups. These results show that firms use ethnicity, age, and employment status as sorting criteria, and thus that discrimination based on these characteristics may be a feature in the Swedish labor market. We then analyze whether contacts received through 'My CV' affect hiring, and find that searchers who receive more contacts have a higher probability of getting hired.

Two papers closely related to ours are Edin & Lagerström (2006) and Eriksson & Lagerström (2006) which use data for 2001 from an earlier version of 'My CV' called the 'Applicant Database'. Edin & Lagerström study discrimination based on gender and ethnicity, and find that women have a lower contact probability than men, while searchers with foreign names do not have a lower contact probability than searchers with Swedish names. Eriksson & Lagerström investigate how employment status affects the contact probability, and find that being unemployed reduces it. The results in this paper confirm some of these results, but also extend the results in several important respects. First, the dataset used here is much larger, 18 167 compared with 8 043 job seekers, and is more representative of the whole Swedish labor market. This may explain the different results for women (see Section 3.3). Second, we have access to more detailed data on the job seekers' ethnicity. This is important since we find that the results differ for searchers belonging to different ethnic groups. Third, we can link our data with other data on the job seekers from the Employment Office, and this means that we can study whether contacts received through 'My CV' matter for the actual hiring outcome.

Our paper is also related to the literature using field experiments to study discrimination. Typically in such studies, two applications, which are designed to be identical in all respects except for the ethnicity or gender of the applicants, are sent to each firm, and the responses of the firms are analyzed. This method can help solve the problems associated with unobserved heterogeneity, but it has also been criticized; it is difficult to construct applications which are identical in all other dimensions, many studies are limited in size and focus on just a few selected occupations, and it may be argued that it is unethical to subject employers to fake job search (see e.g. Heckman (1998) and Riach &



Rich (2002)). Ethnic discrimination is studied e.g. in Bertrand & Mullainathan (2004), who find that searchers with White-sounding names receive more callbacks for interviews than searchers with African-American-sounding names in the US, and Carlsson & Rooth (2006), who find similar effects for searchers with Arabic-sounding names in Sweden. Riach & Rich (2002) summarize the results of several studies finding evidence of discrimination against minorities. Gender discrimination is studied in e.g. Neumark et al (1996), Riach & Rich (1997), Weichselbaumer (2004) and Carlsson (2007), which all find evidence of discrimination against women in particular occupations. A related study is Goldin & Rouse (2000). Discrimination based on employment status, using regression based methods, is studied in Belzil (1996) and Blau & Robins (1990).

The rest of this paper is organized as follows. Section 2 introduces the dataset, presents descriptive statistics and discusses selection issues. Section 3 contains the analysis of how the job seekers' characteristics affect the number of contacts they receive, and includes a discussion of identification and estimation issues, a presentation of the results and a discussion of robustness issues. Section 4 contains the analysis of how the number of contacts received affects the probability of getting hired. Section 5 concludes.

## 2 Data

The database 'My CV' is a search channel offered to job seekers by the Swedish Public Employment Office since 1997. Anyone who wants to find a job, irrespective of current employment status, is invited to submit details to the database about their personal characteristics and the requirements they have about the jobs they want to find. This can be done either from home over the Internet or at the Employment Office. The searchers submit their information by entering their personal details into a number of standardized forms. In the forms, the searchers are asked to enter information about their education, labor market experience, other skills, the requirements they have about the jobs they want to find, and are asked to write a short personal letter.<sup>2</sup> The information is

---

<sup>2</sup> Due to privacy concerns, the Employment Office did not give us access to the personal letters except for their lengths. The letters may contain both information that have been registered elsewhere and new information. Obviously, the letters may affect the firms' contact decisions.

only made visible to employers if all forms have been completed, so there are no missing values. Employers are invited to search in this database for applicants that they find interesting, and can contact them for interviews etc. by e-mail within the system. The searchers can then respond to the contacts by e-mail, also within the system.<sup>3</sup>

In late 2004, 'My CV' contained over 100 000 searchers. All searchers who logged into the system in December 2004 were asked about whether they wanted to participate in a research project on the recruitment behavior of firms. Nearly 40 percent agreed. Those who agreed to participate were also asked to answer a short questionnaire (see Appendix 1). The sample we use includes 18 167 searchers. Table 1 presents some descriptive statistics about these searchers and the jobs they hope to find.

Table 1 Descriptive statistics about the characteristics of the searchers and the jobs they want to find (in fractions)

---

<b>Gender:</b>	
Female	0.53
<b>Ethnicity:</b>	
Foreign name <i>of which</i>	0.18
Nordic name	0.06
African name	0.003
Arabic name	0.02
Asian name	0.01
Other foreign name	0.09
<b>Age:</b>	
Mean (years)	34.3
Age 20-25	0.27
Age 26-35	0.33
Age 36-50	0.28
Age 51-	0.12

---

However, even if we had gotten access to the letters it would have been very difficult to control for their contents in an objective way. We check whether the length of the letters matter, since their length may be correlated with quality, but find that this variable does not affect our results (see Section 3.4).

<sup>3</sup> We have data on the searchers' responses to the firms' contacts, but we do not use this data in the empirical analysis since we know that many searchers respond in other ways, e.g. by phone, rather than by using the system of e-mails. However, doing a similar analysis as in Section 3 for the response rate we find some evidence that groups that have a low contact rate have a higher response rate.

---

<b><i>Employment status:</i></b>	
Employed	0.27
Unemployed	0.53
Participate in a labor market program	0.05
University student	0.05
Participate in other adult education	0.02
High school student	0.01
On parental leave	0.01
Other	0.06
<b><i>Highest level of completed education:</i></b>	
Primary	0.13
Secondary	0.33
Post-secondary	0.54
<b><i>Work experience:</i></b>	
Less than 1 year	0.15
1-2 years	0.12
2-5 years	0.19
5-10 years	0.14
10-15 years	0.10
15-20 years	0.09
More than 20 years	0.21
Almost no experience in desired occupation	0.28
Some experience in desired occupation	0.39
Almost all experience in desired occupation	0.33
<b><i>Other skills:</i></b>	
Managerial experience	0.32
Foreign experience	0.10
Telecommuting experience	0.11
Research experience	0.05
Driving license	0.77
Good language skills - Swedish	1.00
Good language skills - English	0.61
Good language skills – French	0.04
Good language skills - German	0.13
Good language skills - Spanish	0.04
Number of languages	3.5
Number of computer programs	2.3
Number of other skills	4.7
<b><i>Region:</i></b>	
Stockholm	0.27
Uppsala	0.08
Södermanland	0.07
Östergötland	0.07
Jönköping	0.05
Kronoberg	0.04
Kalmar	0.04
Gotland	0.01
Blekinge	0.03
Skåne	0.17
Halland	0.07

---

Västra Götaland	0.19
Värmland	0.04
Örebro	0.06
Västmanland	0.06
Dalarna	0.04
Gävleborg	0.05
Västernorrland	0.05
Jämtland	0.02
Västerbotten	0.04
Norrbottn	0.03
<b>Occupation:</b>	
Legislators, senior officials and managers (Amsyk1)	0.04
Professionals (Amsyk2)	0.32
Technicians and associate professionals (Amsyk3)	0.31
Clerks (Amsyk4)	0.31
Service workers and shop sales workers (Amsyk5)	0.28
Skilled agricultural and fishery workers (Amsyk6)	0.04
Craft and related trades workers (Amsyk7)	0.13
Plant and machine operators and assemblers (Amsyk8)	0.13
Elementary occupations (Amsyk9)	0.31
Mean number of weeks in 'My CV'	50.4
Median number of weeks in 'My CV'	36.9

Notes: The ethnicity variables are based on a question in the questionnaire (see note 4). All other variables are taken directly from 'My CV', except for employment status and experience which are from the questionnaire (this data is compared with the information registered in 'My CV' to make sure it is similar). Experience in desired occupation refers to experience only in occupations where the searcher is looking for work. A searcher may look for work in several regions/occupations.

In Table 1, we see that the searchers in the sample are quite diversified with respect to age, gender, ethnicity, education, experience, region and occupation: The average age is 34 years, 53 percent are women, 18 percent have foreign names, 27 percent are employed, 53 percent are unemployed, more than half have a post-secondary education, most workers have some labor market experience, and most search for work in the metropolitan areas.<sup>4</sup>

The searchers in the sample have received 12 994 contacts from firms during their time in the database. Table 2 gives descriptive statistics on the

<sup>4</sup> The ethnicity of the searchers is not registered in 'My CV'. However, in the questionnaire, the searchers were asked if they believe employers perceive their name as 'Swedish' or 'foreign'. Those who answered 'foreign' were also asked whether they believe employers perceive their name as 'Nordic', 'African', 'Arabic', 'Asian' or 'other foreign'. Since firms can observe the name of the searcher, this variable should capture how ethnicity affects the contact decision. Post-secondary education includes all types of education that is above the secondary level; i.e. university education, training in crafts etc.

contacts received and the responses these contacts have yielded. Also included is data from the Employment Office on the fraction of searchers who have found jobs – using any search channel – after they have registered in ‘My CV’.

Table 2 Descriptive statistics about the contacts received, the responses to these contact and the probability of finding a job (in fractions)

<b>Group</b>	<b>Fraction receiving at least one contact</b>	<b>Average number of contacts received</b>	<b>Average number of responses per contact</b>	<b>Probability of finding a job after registration</b>
All	0.28	0.72	0.21	0.22
Men	0.28	0.84	0.21	0.23
Women	0.28	0.61	0.21	0.22
Swedish name	0.28	0.73	0.20	0.23
Foreign name	0.27	0.65	0.23	0.18
Nordic name	0.30	0.80	0.22	0.21
African name	0.25	0.43	0.16	0.05
Arabic name	0.23	0.44	0.28	0.18
Asian name	0.18	0.38	0.36	0.12
Other foreign name	0.27	0.64	0.22	0.19
Age 20-25	0.22	0.46	0.17	0.21
Age 26-35	0.28	0.69	0.19	0.27
Age 36-50	0.32	0.89	0.24	0.22
Age 51-	0.34	1.00	0.23	0.14
Employed	0.35	0.98	0.20	0.40
Unemployed	0.24	0.57	0.21	0.18
Labor market program	0.24	0.61	0.26	0.12
University student	0.32	0.87	0.16	0.20
Other adult student	0.31	0.71	0.16	0.11
High school student	0.18	0.42	0.16	0.10
On parental leave	0.35	0.84	0.22	0.16
Primary education	0.33	0.97	0.19	0.17
Secondary education	0.23	0.48	0.21	0.21
Post-secondary education	0.30	0.80	0.21	0.25

Notes: The ‘average number of responses’ is the number of responses divided by the number of contacts received. The ‘probability of finding a job’ is the probability of being deregistered by the Employment Office because the searcher has found a job at any time after he or she registered in ‘My CV’ (before May 9, 2005). The last column only includes searchers who are registered both in ‘My CV’ and at the Employment Office (for details see Section 4.1).

In Table 2 we see that 28 percent of the searchers have been contacted by an employer. The average number of contacts received is 0.72. Searchers have, on average, responded 0.14 times, implying that the average response rate is around 21 percent. The average probability of finding a job at any time after registration (but before May 9, 2005) is 0.22. Looking at the different

subgroups, we see that: (1) men and women have the same probability of being contacted at least once, but men receive more contacts, (2) searchers with Swedish names receive more contacts, and have a higher probability of finding jobs, than searchers with foreign names, (3) older searchers receive more contacts than younger searchers, and (4) employed searchers receive more contacts than unemployed searchers.

The results in Table 2 show that searchers with foreign names, women and searchers who are unemployed receive fewer contacts. However, these results may be explained by other differences in observable characteristics across groups. Thus, these results should be interpreted with caution until we control for all differences simultaneously in a regression analysis.

Turning now to the issue of how representative the sample is there are four potentially important selection issues that we need to consider.<sup>5</sup> First we must ask whether the searchers who agreed to participate in our study differ from those who did not, and if this may affect our results. Unfortunately, we do not have any information about the searchers who did not agree to participate. However, this should only affect our results if we fail to include important (observable) variables that are correlated with ethnicity, gender, age or employment status in our regressions, or if our regressions do not fully capture the potentially very complex way the employers use the information. We will investigate this issue further in the robustness analysis in Section 3.4. It should also be noted that when the searchers were asked about whether or not they agreed to participate the question did not reveal much about the exact purpose of the study. Still, even though there are no strong indications that this selection effect should significantly affect the results, this possibility must be kept in mind when interpreting the results.

Second, since both workers and firms can choose whether or not to use ‘My CV’, we might wonder whether those that do use it differ from those that do not. To illustrate whether the searchers in our dataset differ from the typical job searcher in Sweden, Table A 1 in Appendix 2 compares the characteristics of our data with data on typical job searchers from the Employment Office. Comparing these two datasets we see that the searchers are quite similar.

---

<sup>5</sup> Compared with the data used in Edin & Lagerström (2006) and Eriksson & Lagerström (2006) the searchers in our sample are more representative of the average job searcher in Sweden in most dimensions (e.g. ethnicity, age and region), except for gender and education; in our sample the majority of the searchers are women and the fraction with a post-secondary education is high.

The most noteworthy differences are that our sample contains more women and highly educated searchers, and that our searchers are less likely to search for work as service workers (amsyk 5). Regarding the representability of the firms, our dataset does not include direct information about the firms that use it; we only have data on the offers received by the searchers. However, to get a rough sense of whether the vacancies that firms try to fill using our data differ from other vacancies, Table A 2 in Appendix 2 compares the preferred occupation/region of the searchers in ‘My CV’ who have received at least one offer with the inflow of vacancies to the Employment Office. This is obviously not an ideal measure, but it should give us a rough idea since searchers who receive contacts probably are looking for work in occupations/regions corresponding to the occupations/regions of the vacancies. The comparison shows that the searchers in ‘My CV’ who have been contacted are more likely to search for jobs as clerks (amsyk 4), and less likely to search for jobs as service workers (amsyk 5). To see if these differences matter for our results, we will perform the empirical analysis for different subgroups in Section 3.3.

Third, it may be that employers who use ‘My CV’ have access to less information than employers using other search channels, and therefore are more likely to use indicators, such as e.g. ethnicity, to sort workers. If this is the case, our results may overestimate the true effects for the whole labor market. However, it is not obvious that firms which use ‘My CV’ have access to less information than firms that e.g. choose among written applications.

Fourth, it is possible that some employers who find interesting candidates in ‘My CV’ choose to contact them in other ways than by using the system of mailboxes; e.g. contacting the searchers by phone if they have written their phone numbers in the personal letters. However, according to the Employment Office, most employers contact searchers within the system of mailboxes. Since we do not have any data on contacts outside the system of mailboxes, it is difficult to know whether this is an important issue or not, but we should keep it in mind when interpreting the results.

## 3 The number of contacts received

We want to investigate how the job searchers' characteristics affect the number of contacts they receive. In this section, we discuss identification issues, define the variables and the econometric specification, present the results, and discuss robustness issues.

### 3.1 Identification

Suppose that an employer has chosen to use 'My CV' to fill a vacancy. The employer obviously wants to locate the most productive worker. However, a lot of factors will affect the productivity of an applicant in a particular job, and only some of these factors are directly observable in the database. Which characteristics should we expect such an employer to consider relevant? Probably, the employer will consider two types of information. First, all factors that he or she believes directly will affect the productivity of the applicants; e.g. education and work experience. Second, all factors (e.g. ethnicity, gender, age and employment status) that the employer believes are indicators for other important factors (e.g. ability to co-operate, motivation and other social skills) that are unobservable when the decision is made. Typically, these indicators are not important for productivity by themselves, but instead serve as indicators for unobservable characteristics. Also, we cannot ignore the possibility that employers have preferences over the indicators directly.

Now, how can we identify the effect of the searchers' characteristics on the number of contacts received? Here it is important to note that the information of the econometrician coincide with the information set of the firm which could potentially contact the applicant. Thus, we can write the econometric model as a regression function with regressors that are orthogonal to the error term. We can do this because the firm acts on the basis of the expected value of the applicants' ability conditional on observable attributes. The latent variable may therefore be viewed as the expected ability of the applicant conditional on his or her observable characteristics.

### 3.2 Variables and estimation

To identify the effects from ethnicity, gender, age and employment status on the contact rate, we need to control for all other factors which may affect these decisions. We define the variables in a way analogue to the way we described them in Table 1.



For education, we include dummy variables for the highest level of completed education; primary, secondary and post-secondary. For experience, we use dummy variables for seven lengths of experience; less than 1, 1-2, 2-5, 5-10, 10-15, 15-20 and more than 20 years. Also we have dummy variables for the fraction of that experience which is in the occupations where the searcher is looking for work; almost none, some or almost all. For other skills we use dummy variables for managerial experience, foreign work experience, telecommuting experience, research experience, driving skills, computer skills, language skills, and other skills. To control for the quality of the personal letter, we include a variable for its length.<sup>6</sup>

For age, we divide the searchers into five groups; 20-25, 26-35, 36-50 and 51- years old. For gender we use a dummy variable for women. For ethnicity we use dummy variables for Swedish and foreign names, dividing the second group into searchers with Nordic, African, Arabic, Asian and other foreign names. Sometimes we combine the last four groups into searchers with non-Nordic names. For employment status, we divide the searchers into eight groups; employed, unemployed, participants in labor market programs, university students, participants in other adult education, high school students, on parental leave and others.

To capture differences across occupational and regional labor markets we include variables for occupation and region. This is important since we expect that an employer's choice usually will be limited to the searchers who have stated that they are interested in a particular occupation at a particular location. For occupation we use the Employment Office's classification system to divide the searchers into 114 occupations. For location we divide the searchers into 23 regions (21 counties, all of Sweden and abroad).

We also need to include controls for the length of time searchers have been registered in 'My CV', since those that have been in the database longer, on average, have received more contacts. Thus, we include a vector of the variables time and time squared (see the discussion in Section 3.4).

Concerning the estimation, our data are count data, and we use the Poisson model to estimate models for the number of contacts received.<sup>7</sup>

---

<sup>6</sup> This variable does not affect the results.

<sup>7</sup> Other alternatives are the negative binomial model or the Poisson QML model (see the discussion below).

### 3.3 Results

Table 3 presents the results for the number of contacts received.

Table 3 Poisson estimates of the number of contacts received

	(1)	(2)	(3)	(4)
Nordic name	0.014 (0.075)	-0.020 (0.068)	0.023 (0.063)	-0.030 (0.063)
African name	0.008 (0.383)	0.084 (0.376)	0.030 (0.341)	-0.110 (0.358)
Arabic name	-0.311** (0.128)	-0.234* (0.127)	-0.216* (0.131)	-0.286** (0.134)
Asian name	-0.312* (0.187)	-0.183 (0.189)	-0.140 (0.172)	-0.184 (0.173)
Other foreign name	-0.065 (0.067)	-0.233*** (0.067)	-0.218*** (0.057)	-0.228*** (0.054)
Female	-0.179*** (0.034)	-0.072** (0.034)	-0.100** (0.041)	0.018 (0.038)
Age 26-35	-0.016 (0.049)	-0.295*** (0.059)	-0.116** (0.057)	-0.124** (0.054)
Age 36-50	0.065 (0.052)	-0.495*** (0.083)	-0.146** (0.078)	-0.132* (0.074)
Age 51-	0.004 (0.066)	-0.565*** (0.106)	-0.174** (0.096)	-0.165* (0.091)
Unemployed	-0.203*** (0.041)	-0.135*** (0.040)	-0.129*** (0.037)	-0.125*** (0.035)
Labor market program	-0.278*** (0.093)	-0.159* (0.092)	-0.103* (0.081)	-0.128* (0.077)
University student	0.074 (0.082)	0.065 (0.082)	0.076 (0.077)	0.071 (0.075)
Other adult training	-0.029 (0.097)	0.076 (0.094)	-0.077 (0.099)	0.003 (0.095)
High school student	0.136 (0.247)	0.254 (0.243)	-0.029 (0.193)	0.045 (0.189)
On parental leave	0.044 (0.146)	0.028 (0.146)	0.102 (0.131)	0.162 (0.126)
Other	0.034 (0.083)	0.046 (0.075)	-0.050 (0.069)	-0.067 (0.065)
Secondary education	-	0.121** (0.056)	0.088* (0.053)	0.063 (0.049)
Post-secondary education	-	0.164*** (0.050)	0.142*** (0.049)	0.086* (0.045)
1-2 years experience	-	0.111 (0.072)	0.095 (0.066)	0.130** (0.065)
2-5 years experience	-	0.347*** (0.068)	0.317*** (0.063)	0.332*** (0.061)
5-10 years experience	-	0.387*** (0.075)	0.321*** (0.072)	0.333*** (0.069)
10-15 years experience	-	0.474*** (0.084)	0.316*** (0.079)	0.377*** (0.078)

	(1)	(2)	(3)	(4)
15-20 years experience	-	0.646*** (0.101)	0.447*** (0.088)	0.534*** (0.085)
20- years experience	-	0.557*** (0.099)	0.317*** (0.089)	0.402*** (0.087)
Some relevant exp.	-	0.179*** (0.045)	0.104** (0.042)	0.118*** (0.041)
Almost only relevant exp.	-	0.127*** (0.047)	0.108** (0.043)	0.131*** (0.043)
Other skills	No	Yes	Yes	Yes
Dummies for occupation	No	No	Yes	Yes
Dummies for region	No	No	No	Yes
Number of observations	18 167	18 167	18 167	18 167
R <sup>2</sup>	0.26	0.31	0.39	0.42

Notes: The number of contacts received is estimated using the Poisson model. The time vector and a constant are always included. Other skills include the skill variables listed in Table 1. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level. The reference category is a 20-25 years old employed man with a Swedish-sounding name having primary education, and 0-1 year non-relevant experience.

In column 1 in Table 3 we only include factors which firms may use as indicators for unobservable factors; i.e. we include ethnicity, gender, age and employment status (the time vector is included in all regressions). We see that searchers who have Arabic and Asian names, are women, unemployed or participate in labor market programs receive fewer contacts. This indicates that firms use these factors to sort workers, but we cannot exclude the possibility that these effects are explained by other differences across groups. In columns 2 to 4 we therefore successively introduce variables for other observable factors. In column 2 we add the skill variables, and see that many of these factors have strong positive effects on the contact rate; all measures of education and experience are highly significant. In columns 3 and 4 we add the variables for the requirements searchers have about the occupation and region of the desired jobs, and see that these factors also matter. In column 4 we see that the contact rate is lower for searchers with Arabic or other foreign names, older searchers, and searchers who are unemployed or participate in labor market programs, even though the effect from being old or a program participant is significant only at the 10 percent level. For searchers with African and Asian names the negative effects are insignificant, but this may be because these groups are small. If we combine all searchers with non-Nordic names into one category,

we find a strong negative effect (see column 1 in Table 4). The negative effect for women in column 1 disappears when we control for other differences.<sup>8</sup> Since we have data on essentially everything employers observe prior to hiring, these results are a strong indication that firms use these factors to sort workers, and thus that discrimination may be a feature in the Swedish labor market. Concerning the magnitude of the effects, we find that having an Arabic name, being older than 50 or being unemployed reduce the number of contacts received by around 0.2, 0.14 and 0.11 respectively. Since the average number of contacts received is 0.72, the relative effects from these reductions are substantial (15 to 28 percent).

Our results are statistically significant at conventional levels and remain stable across different specifications and estimation methods (e.g. using the negative binomial model or the Poisson QML model). Also, the results are similar if we restrict the sample to searchers who have received at least one contact, or if we treat the unemployed and program participants as one category.<sup>9</sup>

We have also run separate regressions for different subgroups based on gender, age, employment status, ethnicity, place of birth, education, and occupation/region of the desired jobs. Table 4 presents some of these results.

---

<sup>8</sup> The negative effect for women disappears when we include the regional variables. Here it is important to note that this variable captures in which region(s) the searcher is *looking for work*, and not the region of *residence*. Since it is likely that labor market conditions differ across *both* occupations and regions, this indicates that women to a larger extent than men search for work in weak labor markets (defined as particular occupations in particular regions). Also, the women in our dataset, on average, search for work in fewer regions than men. If we instead use the region of residence in the regressions, we find that the negative effect for women is similar to the result in column 3. Thus the fact that women tend to search in fewer regions may, at least partially, explain why women get fewer contacts than men.

<sup>9</sup> Since unemployed searchers do not always update their information when they enter a program, firms may not be able to distinguish program participants from other unemployed searchers.

Table 4 Poisson estimates of the number of contacts received for different subgroups

	All	Women	Men	Age ≥ 40	Age < 40	Un- employed	Non- Nordic name	Nordic name	Born in Sweden
Non-Nordic name	-0.228*** (0.050)	-0.139** (0.066)	-0.261*** (0.071)	-0.325*** (0.080)	-0.129** (0.060)	-0.266*** (0.071)	-	-	-0.159* (0.092)
Female	0.019 (0.038)	-	-	-0.004 (0.059)	0.082* (0.049)	0.026 (0.057)	-0.042 (0.082)	-0.021 (0.025)	-0.010 (0.051)
Age 26-35	-0.125** (0.054)	-0.088 (0.063)	-0.110 (0.091)	-	-	0.051 (0.086)	-0.242** (0.117)	-0.126*** (0.036)	-0.178*** (0.069)
Age 36-50	-0.134* (0.074)	-0.153 (0.094)	-0.103 (0.115)	-	-	0.128 (0.112)	-0.290** (0.143)	-0.139*** (0.050)	-0.167*** (0.105)
Age 51-	-0.166* (0.091)	-0.230* (0.121)	-0.152 (0.136)	-	-	0.031 (0.132)	-0.024 (0.189)	-0.161*** (0.060)	-0.111 (0.125)
Unemployed	-0.125*** (0.035)	-0.102** (0.047)	-0.129*** (0.050)	-0.110** (0.051)	-0.165*** (0.045)	-	-0.265*** (0.078)	-0.120*** (0.022)	-0.173*** (0.048)
Labor market program	-0.129* (0.077)	-0.181* (0.106)	-0.030 (0.103)	-0.066 (0.122)	-0.269*** (0.094)	-	0.130 (0.165)	-0.162*** (0.051)	-0.272*** (0.104)
Number of obs.	18 167	9 618	8 549	5 591	12 576	9 683	2 251	15 916	11 384
R <sup>2</sup>	0.42	0.37	0.47	0.45	0.40	0.41	0.44	0.40	0.38

	Low education	High education	White- collar	Blue- collar	Female occupation	Male occupation	Swedish occupation	Immigrant county	Non-imm. county
Non-Nordic name	-0.164** (0.075)	-0.259*** (0.063)	-0.299*** (0.062)	-0.112** (0.056)	-0.098 (0.090)	-0.224** (0.098)	-0.500** (0.247)	-0.180*** (0.058)	-0.336*** (0.083)
Female	0.186*** (0.060)	-0.058 (0.047)	-0.043 (0.046)	0.051 (0.047)	0.051 (0.087)	0.024 (0.088)	0.052 (0.183)	0.011 (0.047)	0.070 (0.061)
Age 26-35	-0.082 (0.082)	-0.125* (0.071)	-0.181** (0.075)	-0.086 (0.058)	-0.021 (0.088)	-0.002 (0.174)	-0.242 (0.211)	-0.185*** (0.068)	-0.062 (0.079)
Age 36-50	-0.132 (0.111)	-0.133 (0.097)	-0.172* (0.100)	-0.179** (0.084)	-0.004 (0.125)	0.156 (0.189)	-0.361 (0.245)	-0.100 (0.093)	-0.201* (0.109)
Age 51-	-0.263** (0.132)	-0.123 (0.118)	-0.160 (0.117)	-0.212** (0.106)	-0.220 (0.165)	0.136 (0.203)	-0.445 (0.281)	-0.148** (0.113)	-0.245* (0.136)
Unemployed	-0.196*** (0.052)	-0.070 (0.047)	-0.114*** (0.044)	-0.158*** (0.041)	-0.133** (0.061)	-0.164** (0.073)	-0.289*** (0.111)	-0.145*** (0.044)	-0.117** (0.055)
Labor market program	-0.347*** (0.117)	0.025 (0.099)	0.023 (0.099)	-0.171* (0.089)	-0.352*** (0.131)	0.061 (0.159)	-0.395 (0.286)	-0.094 (0.109)	-0.106 (0.107)
Number of obs.	8 377	9 790	9 330	13 491	5 555	4 466	1 649	10 234	7 933
R <sup>2</sup>	0.44	0.41	0.42	0.42	0.36	0.47	0.45	0.41	0.45

Note: Estimated using the Poisson model. The regressions also include the time vector, a constant and all other variables listed in Table 1. ‘Non-Nordic name’ includes all foreign names except Nordic names. ‘Born in Sweden’ only includes searchers who are registered at the Employment Office and who, according to their records, are born in Sweden. ‘Low education’ is primary/secondary education, and ‘high education’ is post-secondary education. ‘White-collar’ is amsyk 1-3, and ‘blue-collar’ is amsyk 4-9. ‘Female occupation’ / ‘male occupation’ / ‘Swedish occupation’ include searchers who only search in occupations where at least 70 percent of the searchers belong to the respective groups. ‘Immigrant counties’ are the 5 counties with most immigrants, and ‘Non-immigrant counties’ are all other counties. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level.

From the results in Table 4 the following are worth noting: (1) In most cases, searchers with non-Nordic names receive fewer contacts. The particular ethnic groups most affected differ somewhat, but since some ethnic groups get very small in the subsamples it is hard to get statistically significant results. However, we find some evidence that the ethnic groups mostly affected among men are searchers with African, Asian and other foreign names, and among women searchers with Arabic and other foreign names.<sup>10</sup> Focusing on differences which are statistically significant, we find that the ethnic effects are stronger for searchers who are over 40 years old, are looking for work in high-skill occupations (Amsyk 1-3), or are looking for work in occupations where most searchers (in 'My CV') have Swedish names. We also see that the negative effect from having a non-Nordic name is substantial for searchers born in Sweden (second-generation immigrants). (2) There is always a negative effect from being unemployed, and mostly for participating in labor market programs. (3) Compared with men, there are some indications that women who have a post-secondary education, or look for work in white-collar occupations, have a lower contact rate, while women who have less education, or look for work in blue-collar occupations, have a higher contact rate. The result that the gender effect is stronger for women looking for white-collar work is statistically significant. These results are similar to results in other studies of gender discrimination. This may also explain the difference between our results and the results in Edin & Lagerström (2006), who find that being a woman has a negative effect on the contact probability. Our sample, which is more representative of the whole Swedish labor market, includes more women with low education who are looking for jobs in low skill occupations.

In addition to the results presented in Table 4 the following are worth noting from the analysis of differences across subgroups: (1) If we run the regressions separately for occupations which typically involve extensive contacts with customers, we find no noteworthy differences. (2) Concerning the other skill variables included in the regressions, we find that most coefficients are similar to the results in Table 3, except that the effect from education is stronger for workers with foreign names. This may reflect that the average primary education is shorter in many other countries.

---

<sup>10</sup> The results for the specific ethnic groups are not displayed in Table 4, but available from the authors upon request.

To summarize the results so far, we find that searchers with non-Nordic names (especially Arabic names), older searchers and searchers who are unemployed (openly unemployed or participate in labor market programs) receive fewer contacts. Moreover, these negative effects appear stable and remain even if we control for all other observable differences across searchers.

### **3.4 Robustness**

To evaluate the robustness of the results we have performed a series of robustness checks.

First, we may be concerned that we have not managed to properly control for all the information firms use when they make their contact decisions. As mentioned before, we have access to the same information as the firms, except for the personal letters. Thus this can only be a problem if the letters matter a lot, or if firms use the information in the database in a way that we cannot capture in a regression analysis. To test whether the letters matter, we have estimated the model both with and without the variable for the length of the letter, as well as run separate regressions for searchers with short and long letters. We find that a longer letter has a positive effect on the number of contacts received, but that the inclusion of this variable does not affect any of the other results. A related concern is that searchers with foreign names have less language skills in Swedish, and that this is reflected in the quality of their letters. However, in Table 4 we showed that most of the negative effect from having a non-Nordic name remains even if we estimate the model only on searchers who are born in Sweden. Since it is natural to assume that second-generation immigrants have stronger Swedish skills than first-generation immigrants, this indicates that the lower contact rate is not primarily explained by non-Nordic immigrants writing lower quality letters. Finally, we have experimented with including a variable measuring the searchers' previous wage as a regressor. We can do this for the subsample consisting of all searchers who have been unemployed in 2004 (this data is collected to calculate unemployment benefits and are taken from the register used to handle these payments; Astat). This variable is not observable to firms, and given that we have managed to control for all observable differences, should not affect the firms' contact decisions. This variable turns out to be statistically insignificant, thus indicating that our regressions capture all relevant differences.

Second, an issue that may cause problems is the stock-flow aspect of the sample. The searchers in our sample have been in the database for different



lengths of time as people enter and leave the database continuously. In the estimation, we have included a time vector consisting of time and time squared to take into account the fact that a searcher who has been in the database longer is more likely to have received more contacts. To test whether the way we control for the time in the database matters, we have tried a number of alternatives to the baseline regression:<sup>11</sup> (1) Dividing the sample into groups based on the searchers' time in the database (e.g. ten week-groups) and then including dummies for the groups as regressors instead of the time vector. (2) Splitting the sample into searchers who have been in the database longer ( $\geq 52$  weeks) or shorter ( $< 52$  weeks) and then running separate regressions including the time vector. Some results of this sensitivity analysis are presented in Table 5.

Table 5 Poisson estimate of the number of contacts received using alternative ways of controlling for time in the database

	Baseline	(1)	(2)
Non-Nordic name	-0.228*** (0.050)	-0.224*** (0.048)	-0.170** (0.070)
Female	0.019 (0.038)	0.026 (0.037)	0.026 (0.060)
Age 26-35	-0.125** (0.054)	-0.120** (0.052)	-0.056 (0.078)
Age 36-50	0.128 (0.112)	-0.138* (0.072)	0.034 (0.112)
Age 51-	-0.166* (0.091)	-0.187** (0.089)	-0.119 (0.144)
Unemployed	-0.125*** (0.035)	-0.108*** (0.034)	-0.255*** (0.057)
Labor market program	-0.129* (0.077)	-0.137* (0.077)	-0.120 (0.132)
Time vector	Yes	Yes	Yes
Number of obs.	18 167	18 167	12 102

Notes: Estimated using the Poisson model. The regressions also include all variables listed in Table 1. In Column 1 we use dummy variables for time in the database (10 week periods) instead of the time vector. Column 2 only includes searchers that have been in the base less than 52 weeks. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level.

<sup>11</sup> Another alternative would be to run the regressions including only contacts received during a short period before or after the time of the study. However, since we do not have data on the time of the contacts we are unable to do this.

From Table 5, we see that most results remain qualitatively unchanged irrespective of how we control for time. The coefficients on non-Nordic name, unemployed and age 51- are always negative, and mostly significant. We have also estimated the model on a sample only including searchers who have been in the database for a very short time (less than 2, 4, 6, 8 and 10 weeks). Using this approach, we can be fairly certain that very few searchers have left the database during the period considered. Here we find that the point estimates that we are interested in always remain negative, but often turn out to be insignificant. However, there is no clear trend in the size of the point estimates when we reduce the time that we allow the searchers to have been in the database. Taken together all this indicates that the way we control for time is not that crucial for our main results.<sup>12</sup>

Third, we may worry that the same contact attempt results in more than one contact, and thus that our dependent variable is biased upwards. To test whether this is a problem, we have run all regressions with the probability of getting at least one contact as the dependent variable using the Probit model. The results are in Appendix 3, and we see that most results are similar, except that being a woman has a positive effect on the probability of receiving a contact, while it has no significant effect in our baseline regressions.

To summarize, the robustness analysis indicates that we do have managed to properly control for all differences across searchers, and thus that our main results are qualitatively robust.

## 4 The probability of getting hired

In the previous section, we have shown that searchers who have non-Nordic names, are old or are unemployed receive fewer contacts from firms than other searchers. However, from these results we cannot be sure that the contacts received through ‘My CV’ have a positive effect on the probability of actually getting hired or, to put it differently, if searchers who receive fewer contacts, all else equal, have a lower chance of getting a job.

---

<sup>12</sup> Another interesting issue is whether searchers that have been unemployed long are treated differently than other unemployed searchers. This issue is hard to analyze with our data since we do not know the time of the contacts (see the discussion in Eriksson & Lagerström (2006).

## 4.1 Identification and estimation

We have data on the complete employment histories for the relevant period for 12 169 of our searchers who are also registered at the Employment Office (we have this data until May 9, 2005).<sup>13</sup> Thus we can observe if, and when, these searchers are deregistered by the Employment Office because they have found jobs. Using these data we can construct a variable measuring the probability of getting a job, and then investigate whether this variable is affected by contacts received through ‘My CV’. To do this we construct a variable that measures if the searcher has found a job at *any time* after the date he or she registered.

A crude way of measuring the effect of the contacts received on the hiring probability is to run a regression with the number of contacts received as the only explanatory variable.<sup>14</sup> Running such a regression using the Probit model, we find that the coefficient is 0.016 (0.007).<sup>15</sup> However, the coefficient in this regression is likely to be biased, since it may simply reflect the fact that factors such as education, which are observable to employers using ‘My CV’, have a positive effect on the hiring probability.

Thus, a crucial issue is how we can identify the effect from a contact registered in ‘My CV’ on the probability of getting hired. The answer is that since we know that we have access to the *same* information as the firms which use the database, *the number of contacts received cannot be affected by any other factors*; i.e. there are no unobservable factors which may affect the number of contacts received. This means that we can specify a regression equation where the probability of finding a job is a function of the contacts received and all other observable factors in ‘My CV’ which may affect the contact rate, and then interpret the coefficient for contacts received as an unbiased measure of the casual effect these contacts have on the hiring probability. The fact that there are other variables, such as e.g. motivation, which may affect the hiring probability, does not bias this estimate as long as they are unobservable to firms in ‘My CV’.

---

<sup>13</sup> Searchers who register in ‘My CV’ are not required to also register at the Employment Office. However, searchers who want to receive unemployment benefits must register at the Employment Office, and are also strongly encouraged to register in ‘My CV’. Thus most unemployed searchers in our sample are in both registers, while many non-unemployed searchers in our sample are not registered at the Employment Office.

<sup>14</sup> An alternative is to use whether searchers have received at least one contact.

<sup>15</sup> The robust standard error is reported within the parenthesis.

## 4.2 Results

Table 6 presents the results of the estimation of the probability of getting hired.

Table 6 Probit estimates of the probability of getting hired

	Number of contacts	At least one contact
Contacts	0.016* (0.010)	0.094*** (0.036)
Nordic name	-0.029 (0.058)	-0.028 (0.058)
African name	-0.936*** (0.354)	-0.935*** (0.352)
Arabic name	-0.149 (0.098)	-0.148 (0.098)
Asian name	-0.418*** (0.154)	-0.417*** (0.154)
Other foreign name	-0.149*** (0.048)	-0.149*** (0.048)
Female	-0.023 (0.033)	-0.026 (0.033)
Age 26-35	-0.047 (0.041)	-0.047 (0.041)
Age 36-50	-0.249*** (0.061)	-0.250*** (0.061)
Age 51-	-0.569*** (0.083)	-0.570*** (0.083)
Unemployed	-0.650*** (0.030)	-0.649*** (0.030)
Labor market program	-0.890*** (0.072)	-0.887*** (0.072)
University student	-0.632*** (0.071)	-0.631*** (0.071)
Other adult training	-0.890*** (0.099)	-0.887*** (0.099)
High school student	-0.943*** (0.366)	-0.947*** (0.364)
On parental leave	-0.794*** (0.142)	0.795*** (0.142)
Other	-0.658*** (0.061)	-0.658*** (0.061)
Secondary education	0.097** (0.046)	0.097** (0.046)
Post-secondary education	0.130*** (0.046)	0.129*** (0.046)
Experience	Yes	Yes
Other skills	Yes	Yes
Dummies for region/occupation	Yes	Yes

	Number of contacts	At least one contact
Number of observations	12 969	12 969
R <sup>2</sup>	0.09	0.09

Notes: Estimated using the Probit model. The dependent variable is the probability of finding a job at any time (before May 9, 2005). All regressions also include all other explanatory variables listed in Table 1 and the time vector. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level.

From Table 6, we see that the variables measuring the number of contacts received, and having received at least one contact, have positive and significant effects on the probability that a searcher has found a job, even though the number of contacts received is significant only at the 10 percent level. This can be interpreted as an indication that searchers who use ‘My CV’ benefit from this in terms of their search success. The magnitudes of the effects are such that having received at least one contact increases the probability of being hired with 0.03 percentage points, and each additional contact results in a 0.005 percentage points higher probability of getting a job. Since the average probability of finding a job is 0.22, the relative effect from this increase is around 13 percent for being contacted at least once. Also, we see that most other variables have the expected signs; i.e. it is negative to have a foreign name, be old or unemployed. However, these results have to be treated with caution since all estimates, except the estimates for contacts, are affected by a lot of factors which we cannot control for (e.g. motivation). As a robustness check we have also run the regressions only including jobs received after December 2004. The results are similar.

We have also run separate regressions for different subgroups to see if the effects differ across groups. Table 7 reports some of these results.

Table 7 Probit estimates of the probability of getting hired for different subgroups

	All	Women	Men	Non-Nordic name	Nordic name	Un-employed
Number of contacts	0.016* (0.010)	0.071*** (0.018)	-0.008 (0.012)	0.086** (0.038)	0.012 (0.010)	0.027* (0.015)
Number of obs.	12 969	6 802	6 075	1 524	11 361	7 658
R <sup>2</sup>	0.09	0.10	0.11	0.18	0.09	0.06

Notes: Estimated using the Probit model. The dependent variable is the probability of finding a job at any time (before May 9, 2005). ‘Contacts’ is the number of contacts received. All regressions also include all other explanatory variables listed in Table 1 and the time vector. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level.

In Table 7, we see that having received contacts in ‘My CV’ has a significant positive effect on the probability of finding a job for women, searchers with non-Nordic names and unemployed searchers. For men the effect is close to zero, and for searchers with Nordic names it is positive, but insignificant. These results may be interpreted as an indication that ‘My CV’ matters more for groups that have a difficult time finding a job using other search channels.

To summarize, we find that searchers who have received contacts during their time in ‘My CV’, all else equal, have a somewhat higher probability of actually getting a job. Thus the lower contact rate for e.g. searchers with non-Nordic names seems to matter for the actual hiring outcome by decreasing their chance of getting hired. This is problematic since our results indicate that searchers who belong to these groups actually have a *higher* probability than other searchers of getting hired *if they are contacted*.

## 5 Conclusions

The emergence of Internet-based search channels introduces new opportunities for research on discrimination. The fact that firms using these search channels base their choices only on factors which are registered in the database gives researchers an opportunity to bypass some of the problems associated with unobserved heterogeneity. Since in principle the same information is observed by the recruiting firms and the researcher, it becomes easier to identify discrimination in the initial stage of the hiring process. Thus the use of such data is an

attractive alternative to field experiments. The main advantages being that it is easier to perform these studies on large datasets which are likely to be representative of the whole labor market, and that these data involve genuine job search.

This paper uses data from a Swedish version of an Internet-based database to investigate how the number of contacts received is affected by factors such as ethnicity, gender, age and employment status. We find that the number of contacts received is lower for searchers from some ethnic groups, older searchers, and searchers who are unemployed (both openly and participating in labor market programs). Some of these differences are explained by differences in education, experience, other skills and the requirements searchers have about the jobs, but even when we control for all such differences most of the negative effects remain. Moreover, these differences matter for the hiring outcome, since we also find evidence that searchers who receive more contacts have a higher probability of actually finding jobs.

Our results show that firms using ‘My CV’ sort their applicants based on easily observable characteristics such as ethnicity, age and employment status. An important issue is if this should be labeled ‘discrimination’. As mentioned earlier, discrimination can be of two types; statistical discrimination, where firms use easily observable factors as indicators for important unobservable factors, and preference-based discrimination, where employers have preferences over these characteristics. In the first case it may be argued that it is completely rational for firms to sort workers based on characteristics functioning as indicators and thus that no actual ‘discrimination’ occurs, whereas in the second case most people would agree that actual ‘discrimination’ occurs. We cannot separate these two types of behavior, even though statistical discrimination is a much more likely explanation for some of the effects, such as the negative effect of being unemployed. However, irrespective of why firms use ethnicity, age and employment status to sort their applicants, the implications of this behavior are that searchers who belong to these groups get fewer contacts and, as a consequence of this, have a lower probability of finding jobs.

## References

- Belzil, C. (1996), Relative Efficiencies and Comparative Advantages in Job Search, *Journal of Labor Economics* 14, 154-173.
- Bertrand, M. and S. Mullainathan (2004), Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, *American Economic Review*, 94, 991-1013.
- Blau, D. M. and Robins, P. K. (1990), Job Search Outcomes for the Employed and Unemployed, *Journal of Political Economy* 98, 637-655.
- Carlsson, M. (2007), Sex Discrimination and Segregation in the Swedish Labour Market: A Field Experiment, Mimeo, Kalmar University College, Kalmar.
- Carlsson, M. and D.-O. Rooth (2006), Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data, IZA Working Paper 2281, IZA, Bonn.
- Edin, P.-A. and J. Lagerström (2006), Blind Dates: Quasi-experimental Evidence on Discrimination, Working Paper 2006:4, IFAU, Uppsala.
- Eriksson, S. and J. Lagerström (2006), Competition between Employed and Unemployed Job Applicants: Swedish Evidence, *Scandinavian Journal of Economics*, 108, 373-396.
- Goldin, C., and C. Rouse (2000), Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians”, *American Economic Review*, 90, 715-741.
- Heckman, J. (1998), Detecting Discrimination, *Journal of Economic Perspectives* 12, 101-116.
- Neumark, D. with the assistance of R. J. Bank and K. D. van Nort (1996), Sex Discrimination in Restaurant Hiring: An Audit Study, *Quarterly Journal of Economics*, 111, 915-941.
- Riach, P.-A. and J. Rich (1997), Testing for Sexual Discrimination in the Labor Market, *Australian Economic Papers*, 26, 165-178.
- Riach, P.-A. and J. Rich (2002), Field Experiments of Discrimination in the Market Place, *Economic Journal*, 112, F480-F518.



Weichselbaumer, D. (2004), Is It Sex or Personality? The Impact of Sex-stereotypes on Discrimination in Applicant Selection, *Eastern Economic Journal*, 30, 159-186.

# Appendix 1: The questionnaire

The searchers who agreed to participate in this study were asked to answer the following questions.

So far, have you received a contact with an employer from using 'My CV'?

- Yes
- No

What is your main employment status at present?

- I have a job
- I'm unemployed
- I participate in a labour market program
- I'm a university student
- I participate in adult education
- I'm a high school student
- I'm on parental leave
- None of the above

How long is your total labour market experience?

- Less than one year
- 1-2 years
- 2-5 years
- 5-10 years
- 10-15 years
- 15-20 years
- More than 20 years

How much of your total labour market experience are in those occupations you are looking for work in?

- Nothing or almost nothing (less than 25 %)
- Some (25-75 %)
- All or almost all (more than 75 %)

Are you registered as unemployed at the Employment Office?

- Yes
- No

Do you think that an employer in general perceives your name as Swedish?

- Yes
- No

If you answered no to the previous question: Do you think that an employer generally perceives your name as...

- Nordic
- Asian
- African
- Arabic
- None of the above

THANK YOU FOR YOUR PARTICIPATION!

## Appendix 2: Comparison of the characteristics of the searchers

Table A 1 Comparison of the characteristics of the searchers in 'My CV' and at the Swedish Public Employment Office (in fractions)

	Unemployed searchers		All searchers	
	'My CV'	Employ. Office	'My CV'	Employ. Office
<b>Ethnicity:</b>				
Foreign name	0.18	0.19	0.18	0.19
<b>of which</b>				
African name	0.004	0.002	0.003	0.001
Arabic name	0.02	0.04	0.02	0.05
Asian name	0.01	0.05	0.01	0.05
Other foreign name	0.09	0.09	0.09	0.09
<b>Gender:</b>				
Female	0.50	0.40	0.53	0.50
<b>Age:</b>				
Mean (years)	35.5	35.5	34.3	39.2
Age 20-25	0.25	0.26	0.27	0.18
Age 26-35	0.30	0.31	0.32	0.26
Age 36-50	0.29	0.28	0.28	0.33
Age 51-	0.15	0.15	0.12	0.23
<b>Highest level of completed education:</b>				
Primary	0.14	0.20	0.13	0.20
Secondary or post-secondary	0.86	0.80	0.87	0.80
<b>Work experience:</b>				
No or almost no experience	0.27	0.25	0.26	0.22
Some or almost all experience	0.73	0.75	0.74	0.78
<b>Region of residence:</b>				
Stockholm	0.21	0.19	0.21	0.14
Uppsala	0.03	0.03	0.03	0.03
Södermanland	0.03	0.03	0.03	0.03
Östergötland	0.06	0.05	0.05	0.05
Jönköping	0.02	0.03	0.02	0.03
Kronoberg	0.02	0.02	0.02	0.02
Kalmar	0.02	0.03	0.02	0.03
Gotland	0.01	0.01	0.01	0.01
Blekinge	0.01	0.02	0.01	0.02
Skåne	0.16	0.13	0.15	0.14
Halland	0.03	0.03	0.03	0.03
Västra Götaland	0.15	0.16	0.15	0.17
Värmland	0.02	0.03	0.02	0.04
Örebro	0.03	0.03	0.03	0.04

	Unemployed searchers		All searchers	
	'My CV'	Employ. Office	'My CV'	Employ. Office
Västmanland	0.03	0.03	0.03	0.03
Dalarna	0.03	0.04	0.03	0.03
Gävleborg	0.03	0.04	0.04	0.04
Västernorrland	0.04	0.03	0.04	0.03
Jämtland	0.01	0.02	0.01	0.02
Västerbotten	0.03	0.03	0.03	0.03
Norrbottn	0.03	0.03	0.03	0.04
<b>Occupation:</b>				
Amsyk 1	0.04 (0.02)	0.02 (0.01)	0.04 (0.02)	0.02 (0.01)
Amsyk 2	0.30 (0.17)	0.23 (0.15)	0.32 (0.17)	0.18 (0.11)
Amsyk 3	0.30 (0.16)	0.17 (0.11)	0.31 (0.16)	0.17 (0.10)
Amsyk 4	0.31 (0.17)	0.23 (0.15)	0.31 (0.17)	0.25 (0.16)
Amsyk 5	0.27 (0.15)	0.39 (0.25)	0.28 (0.15)	0.40 (0.25)
Amsyk 6	0.04 (0.02)	0.03 (0.02)	0.04 (0.02)	0.04 (0.03)
Amsyk 7	0.13 (0.07)	0.13 (0.05)	0.13 (0.07)	0.16 (0.10)
Amsyk 8	0.13 (0.07)	0.19 (0.12)	0.13 (0.07)	0.19 (0.12)
Amsyk 9	0.31 (0.17)	0.18 (0.11)	0.31 (0.17)	0.19 (0.12)

Notes: The ethnicity variables in 'My CV' are compared with data on the country of birth from the Employment Office. Occupation refers to the occupation(s) where the searchers want to find jobs and is reported both as the total fraction of the searchers that searches for a job in a particular region / occupation and (in parenthesis) the fraction of all searchers that searches for a job in a particular region / occupation (i.e. the second fractions are required to sum to one).

Table A 2 Comparison of the searchers in 'My CV' who have been contacted and the vacancies reported to the Swedish Public Employment Office (in fractions)

Variable	'My CV'	Employment Office
<b>Region:</b>		
Stockholm	0.37 (0.22)	0.25
Uppsala	0.10 (0.06)	0.03
Södermanland	0.09 (0.05)	0.03
Östergötland	0.08 (0.05)	0.04
Jönköping	0.05 (0.03)	0.04
Kronoberg	0.04 (0.02)	0.02
Kalmar	0.04 (0.02)	0.02
Gotland	0.02 (0.01)	0.01
Blekinge	0.04 (0.02)	0.02
Skåne	0.14 (0.09)	0.11
Halland	0.08 (0.05)	0.02
Västra Götaland	0.18 (0.11)	0.17
Värmland	0.04 (0.02)	0.04
Örebro	0.06 (0.04)	0.02
Västmanland	0.07 (0.04)	0.03
Dalarna	0.04 (0.02)	0.03
Gävleborg	0.05 (0.03)	0.02
Västernorrland	0.04 (0.02)	0.03
Jämtland	0.03 (0.02)	0.02
Västerbotten	0.04 (0.02)	0.02
Norrbottn	0.04 (0.02)	0.03
<b>Occupation:</b>		
Amsyk 1	0.07 (0.03)	0.02
Amsyk 2	0.37 (0.17)	0.15
Amsyk 3	0.41 (0.19)	0.18
Amsyk 4	0.36 (0.17)	0.07
Amsyk 5	0.32 (0.16)	0.35
Amsyk 6	0.04 (0.02)	0.01
Amsyk 7	0.13 (0.06)	0.05
Amsyk 8	0.13 (0.06)	0.08
Amsyk 9	0.30 (0.14)	0.09

Note: Region and occupation refer to the region / occupation where the searchers want to find jobs. The data from the Employment Office is the inflow of new vacancies in 2004. The data for 'My CV' includes only those searchers that have received at least one contact, and is reported both as the total fraction of the searchers that searches for a job in a particular region / occupation and (in parenthesis) the fraction of all searchers that searches for a job in a particular region / occupation (i.e. the second fractions are required to sum to one).

## Appendix 3: The probability of receiving a contact

Table A 3 Probit estimates of the probability of receiving a contact

	(1)	(2)	(3)	(4)	(5)
Non-Nordic name	-	-	-	-	-0.110*** (0.038)
Nordic name	-0.009 (0.047)	-0.045 (0.048)	-0.004 (0.051)	-0.048 (0.051)	-
African name	0.200 (0.201)	0.239 (0.203)	0.194 (0.198)	0.071 (0.200)	-
Arabic name	-0.075 (0.081)	-0.051 (0.085)	-0.068 (0.087)	-0.138 (0.089)	-
Asian name	-0.193* (0.117)	-0.146 (0.121)	-0.222* (0.124)	-0.296** (0.125)	-
Other foreign name	0.008 (0.038)	-0.071* (0.041)	-0.061 (0.043)	-0.097** (0.044)	-
Female	0.062*** (0.022)	0.124*** (0.024)	0.092*** (0.030)	0.146*** (0.030)	0.146*** (0.030)
Age 26-35	-0.007 (0.028)	-0.186*** (0.036)	-0.033 (0.039)	-0.048 (0.039)	-0.051 (0.039)
Age 36-50	0.044 (0.030)	-0.221*** (0.052)	-0.007 (0.055)	-0.003 (0.056)	-0.007 (0.056)
Age 51-	-0.011 (0.039)	-0.331*** (0.067)	-0.106 (0.072)	-0.124* (0.072)	-0.127* (0.072)
Unemployed	-0.150*** (0.025)	-0.107*** (0.026)	-0.131*** (0.028)	-0.120*** (0.028)	-0.120*** (0.028)
Labor market program	-0.290*** (0.056)	-0.219*** (0.057)	-0.222*** (0.060)	-0.171*** (0.062)	-0.173*** (0.061)
University student	-0.001 (0.053)	-0.004 (0.055)	-0.030 (0.058)	-0.021 (0.059)	-0.023 (0.059)
Other adult training	-0.062 (0.067)	-0.026 (0.068)	-0.107 (0.072)	-0.079 (0.072)	-0.082 (0.071)
High school student	-0.100 (0.134)	-0.063 (0.134)	-0.268* (0.152)	-0.245 (0.154)	-0.250 (0.154)
On parental leave	0.003 (0.107)	0.001 (0.111)	-0.032 (0.112)	0.006 (0.113)	0.009 (0.113)
Other	-0.021 (0.049)	-0.010 (0.050)	-0.080 (0.053)	-0.070 (0.054)	-0.071 (0.054)
Secondary education	-	0.112*** (0.039)	0.091** (0.041)	0.097** (0.042)	0.098** (0.042)
Post-secondary education	-	0.127*** (0.038)	0.136*** (0.040)	0.115*** (0.041)	0.116*** (0.041)
1-2 years experience	-	0.085* (0.045)	0.066 (0.047)	0.070 (0.048)	0.071 (0.048)

	(1)	(2)	(3)	(4)	(5)
2-5 years experience	-	0.193*** (0.041)	0.181*** (0.043)	0.185*** (0.044)	0.186*** (0.044)
5-10 years experience	-	0.267*** (0.048)	0.202*** (0.051)	0.223*** (0.052)	0.225*** (0.052)
10-15 years experience	-	0.271*** (0.056)	0.189*** (0.059)	0.222*** (0.060)	0.223*** (0.060)
15-20 years experience	-	0.216*** (0.063)	0.133** (0.067)	0.168** (0.068)	0.172** (0.068)
20- years experience	-	0.266*** (0.063)	0.160** (0.066)	0.199*** (0.067)	0.204*** (0.067)
Some relevant exp.	-	0.129*** (0.029)	0.073** (0.030)	0.086*** (0.031)	0.086*** (0.031)
Almost only relevant exp.	-	0.134*** (0.031)	0.128*** (0.032)	0.129*** (0.033)	0.129*** (0.033)
Other skills	No	Yes	Yes	Yes	Yes
Dummies for occupation	No	No	Yes	Yes	Yes
Dummies for region	No	No	No	Yes	Yes
Number of observations	18 167	18 167	18 167	18 167	18 167
R <sup>2</sup>	0.19	0.22	0.29	0.32	0.32

Notes: The probability of receiving a contact is estimated using the Probit model. The time vector and a constant are always included. Other skills include the skill variables listed in Table 1. Robust standard errors are in parentheses. \*\*\*, \*\* and \* denote significance at the 1, 5 and 10 percent level. The reference category is a 20-25 years old employed man with a Swedish-sounding name having primary education, and 0-1 year non-relevant experience.



## Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

### Rapporter/Reports

- 2007:1** Lundin Daniela ”Subventionerade anställningar för unga – en uppföljning av allmänt anställningsstöd för 20–24-åringar”
- 2007:2** Lundin Daniela, Eva Mörk & Björn Öckert ”Maxtaxan inom barnomsorgen – påverkar den hur mycket föräldrar arbetar?”
- 2007:3** Bergemann Annette & Gerard van den Berg ”Effekterna av aktiv arbetsmarknadspolitik för kvinnor i Europa – en översikt”
- 2007:4** Junestav Malin ”Socialförsäkringssystemet och arbetsmarknaden – politiska idéer, sociala normer och institutionell förändring – en historik”
- 2007:5** Andersson Christian ”Lärartäthet, lärarkvalitet och arbetsmarknaden för lärare”
- 2007:6** Larsson Laura & Caroline Runeson ”Effekten av sänkt sjukpenning för arbetslösa”
- 2007:7** Stenberg Anders ”Hur påverkar gymnasialt komvux löneinkomster och vidare studier?”
- 2007:8** ”Forslund Anders & Kerstin Johansson ”Lediga jobb, arbetssökande och anställningar – den svenska matchningsfunktionen”
- 2007:9** Kennerberg Louise ”Hur förändras kvinnors och mäns arbetssituation när de får barn?”
- 2007:10** Nordin Martin ”Invandrades avkastning på utbildning i Sverige”
- 2007:11** Johansson Mats & Katarina Katz ”Underutnyttjad utbildning och lönegapet mellan kvinnor och män”
- 2007:12** Gartell Marie, Ann-Christin Jans & Helena Persson ”Utbildningens betydelse för flöden på arbetsmarknaden”
- 2007:13** Grönqvist Hans & Olof Åslund ”Familjestorlekens effekter på barns utbildning och arbetsliv”
- 2007:14** Lindqvist Linus ”Uppföljning av plusjobb”
- 2007:15** Sibbmark Kristina ”Avidentifierade jobbansökningar – erfarenheter från ett försök i Göteborgs stad”
- 2007:16** Hesselius Patrik & Malin Persson ”Incitamenteffekter och Försäkringskassans kostnader av kollektivavtalade sjukförsäkringar”
- 2007:17** Eriksson Stefan & Jonas Lagerström ”Diskriminering i anställningsprocessen: resultat från en Internetbaserad sökkanal”

## Working Papers

- 2007:1** de Luna Xavier & Per Johansson “Matching estimators for the effect of a treatment on survival times”
- 2007:2** Lundin Daniela, Eva Mörk & Björn Öckert “Do reduced child care prices make parents work more?”
- 2007:3** Bergemann Annette & Gerard van den Berg “Active labor market policy effects for women in Europe – a survey”
- 2007:4** Andersson Christian “Teacher density and student achievement in Swedish compulsory schools”
- 2007:5** Andersson Christian & Nina Waldenström “Teacher supply and the market for teachers”
- 2007:6** Andersson Christian & Nina Waldenström “Teacher certification and student achievement in Swedish compulsory schools”
- 2007:7** van den Berg Gerard, Maarten Lindeboom & Marta López ”Inequality in individual mortality and economic conditions earlier in life”
- 2007:8** Larsson Laura & Caroline Runeson “Moral hazard among the sick and unemployed: evidence from a Swedish social insurance reform”
- 2007:9** Stenberg Anders “Does adult education at upper secondary level influence annual wage earnings?”
- 2007:10** van den Berg Gerard “An economic analysis of exclusion restrictions for instrumental variable estimation”
- 2007:11** Forslund Anders & Kerstin Johansson “Random and stock-flow models of labour market matching – Swedish evidence”
- 2007:12** Nordin Martin “Immigrants’ return to schooling in Sweden”
- 2007:13** Johansson Mats & Katarina Katz “Wage differences between women and men in Sweden – the impact of skill mismatch”
- 2007:14** Gartell Marie, Ann-Christin Jans & Helena Persson “The importance of education for the reallocation of labor: evidence from Swedish linked employer-employee data 1986–2002”
- 2007:15** Åslund Olof & Hans Grönqvist “Family size and child outcomes: Is there really no trade-off?”
- 2007:16** Hesselius Patrik & Malin Persson “Incentive and spill-over effects of supplementary sickness compensation”
- 2007:17** Engström Per & Patrik Hesselius “The information method – theory and application”

**2007:18** Engström Per, Patrik Hesselius & Malin Persson “Excess use of Temporary Parental Benefit”

**2007:19** Eriksson Stefan & Jonas Lagerström “Detecting discrimination in the hiring process: evidence from an Internet-based search channel”

### **Dissertation Series**

**2006:1** Hägglund Pathric “Natural and classical experiments in Swedish labour market policy”

**2006:2** Savvidou Eleni “Technology, human capital and labor demand”

**2006:3** Söderström Martin “Evaluating institutional changes in education and wage policy”

**2006:4** Lagerström Jonas “Discrimination, sickness absence, and labor market policy”

**2006:5** Johansson Kerstin “ Empirical essays on labor-force participation, matching, and trade”