IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Blind dates:
quasi-experimental evidence on discrimination

Per-Anders Edin
Jonas Lagerström

WORKING PAPER 2006:4

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Industry, Employment and Communications, 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 and evaluations of the labour market effects of measures within the educational system. 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 traditional board, consisting of a chairman, the Director-General and eight other members. The tasks of the board are, among other things, to make decisions about external grants and give its views on the activities at IFAU. 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
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.

ISSN 1651-1166

Blind dates: quasi-experimental evidence on discrimination*

by

Per-Anders Edin[a] and Jonas Lagerström[b]

May 17, 2006

Abstract

This paper provides evidence on discrimination in the hiring process. We use data generated from a "policy experiment" conducted at the Swedish public employment offices. Individuals registered at these offices can post their qualifications in a database available to employers over the Internet. Potential employers are free to search this database for job candidates and contacts between employers and candidates are recorded. We use two complementary identification strategies. First, since our data contain all information available to employers, we argue that selection on observables is viable. Second, we utilize the fact that individuals can choose not to reveal their name and gender to potential employers. Our main finding is that women have a 20 percent lower chance than men of getting contacted by employers and that this differential is fully explained by discrimination. Our results concerning ethnic discrimination are less conclusive, probably due to measurement errors.

* We gratefully acknowledge comments from Nils Gottfries, Peter Fredriksson, Björn Öckert as well as seminar participants at IFAU, Uppsala University and at the CEPR conference on Discrimination and Unequal Outcomes held in Le Mans, France, 2002. We also thank AMS and Claes-Göran Lock for providing us with the data.

[a] Department of Economics, Uppsala University, and Institute for Labour Market Policy Evaluation (IFAU). E-mail: Per-Anders.Edin@nek.uu.se.

[b] Department of Economics and Statistics, Åbo Akademi University, and Institute for Labour Market Policy Evaluation (IFAU), E-mail: Jonas.Lagerstrom@nek.uu.se

Table of contents

1 Introduction

Like in many other Western economies, discrimination in the labor market is a major issue in the Swedish policy debate. In spite of its well known equality of outcomes, the Swedish labor market still produces large differentials in labor market outcomes. The two groups that are most often mentioned in the Swedish debate are immigrants and women. The key question, which is very hard to answer, is how important labor market discrimination is to explain these differences. This paper analyzes discrimination in the hiring process. There is ample evidence that observed differentials are mainly driven by differences in hiring and promotion, rather than by differences in wages within jobs.

Immigrants in the Swedish labor market earn substantially less than native Swedes and have actually been loosing ground over the last decade. In 1998, the average non-OECD immigrant earned about 45 percent of what a native Swede with similar observed characteristics earned per year (Edin & Åslund, 2001). Roughly a quarter of this difference was due to differences in hourly wages. Another quarter was due to less working hours among those employed. The remaining half of the earnings difference was due to lower employment rates among immigrants.

Even though Swedish women are relatively high paid, compared to in most other Western economies, they still earn only about 80 percent of men's hourly wage. A large share of the earnings gap is driven by occupational segregation. Controlling for standard "human capital variables", reduces the wage gap by about half, e.g. le Grand (1997) and Albrecht *et al*. (2003). Most of the remaining gap, though, is eliminated if detailed controls for occupations are introduced (Meyerson & Petersen, 1997). Both these examples illustrate that the sorting of workers to jobs, through hiring and promotion, is crucial for generating the observed differences in outcomes across groups in the labor market. Consequently, we need to get a better understanding for how this sorting occurs to get a grip of the role of discrimination in the labor market.

The standard approach to analyzing discrimination, building on the seminal work by Becker (1957), has been to estimate various outcome equations in the spirit of Blinder-Oaxaca. Even though these analyses are informative, they require very strong assumptions to infer anything about discrimination. For instance, we have to assume that the unobservables are not systematically different across groups.

One approach that tries to deal with this issue in the hiring process is the "Audit method", surveyed by Riach & Rich (2002). Here, observably similar individuals from different groups, e.g. sex or ethnicity, apply for jobs at the same firms. A recent example is Bertrand & Mullainathan (2004) who found that résumés carrying distinctively Black names are less likely to receive job interviews. This approach seems to be a step forward, but also has it's limitations as discussed by Heckman (1998). He shows that the Audit studies may actually be worse than regular observational studies under some assumptions. For example, a man and a woman who share the same personal characteristic may send a different signal in terms of anticipated productivity which the researcher cannot control for. Also, Heckman argues that the findings considering discrimination depend on differences in the variance of uncontrolled characteristics between groups and/or the qualifications needed for the applied job. In addition, of course, there are ethical issues: in these experiments the firms cannot choose whether to participate and they get an extra cost of recruiting applicants who have no intention of accepting a job offer.

The most compelling evidence of discrimination in the recruitment process using observational data has been produced in an analysis of what we refer to as a natural experiment. Goldin & Rouse (2000) use the introduction of blind auditions in U.S. symphony orchestras to analyze discrimination of women in hiring. In a differences-in-differences analysis, they find that the introduction of blind auditions increased the probability that a woman will be hired by a substantial amount. The probability that a woman would be advanced out of a preliminary round was increased by 50 percent, and her likelihood of winning the final round increased by 30 percent when blind auditions were introduced.

Our paper is mainly concerned with gender discrimination. We use data from the Swedish public employment offices. Individuals registered at these offices can post their qualifications in a database available to employers over the Internet. Potential employers are free to search this database for job candidates and contacts between employers and candidates are recorded. An important feature of this system is that individuals can choose to "censor" some of the information available to potential employers. In particular, individuals can choose not to reveal their name and gender.

We use two complementary empirical strategies for identification. The first strategy is closely related to the audit method in that it relies on selection on observables. We argue that our data, that contain all information observed to

employers, provides a good setting for identifying discrimination. The second approach is heavily inspired by the Goldin & Rouse (2000) paper in that we make use of a "quasi-experiment". By comparing the "contact rate" of censored and non-censored women and minorities, we are able to investigate how employers use gender and "foreign names" as a screening device in their hiring process.

The rest of the paper is outlined as follows. In Section 2 we describe the institutional features of the internet search service and the "experiment" we are using. We then turn to describing the data collection procedure and our sample in Section 3. Section 4 contains our estimation strategy and the empirical estimates of discrimination. In Section 5 we conclude by discussing the implications of our results for outcomes in the labor market.

2 The Internet Applicant database

Sweden has a long history of publicly provided employment exchanges. Already in the 1930's, there were public (municipal) employment offices whose main objective was to improve the matching process in the labor market. Nowadays, the employment offices are run by the National Labor Market Board (AMS), who also administer the large supply of various active labor market policies.

In the fall of 1997 AMS started up a new internet based search database to further promote efficiency in the matching of job searchers and employers. This database, called the Applicant database ("Sökandebanken"), provides the data for our study. The basic idea with this tool is that all job applicants (employed or not) can post their resumes on the applicant database free of charge. Furthermore, there is no requirement to register at the employment office before entering the database. Job searchers can present their job histories and qualifications, as well as list their preferred occupations and other aspirations. They are also required to write a more personal letter about themselves. All this can be done either at one of the employment offices or through internet. The software also provides examples of how to put up a CV and similar practical issues. By the spring of 2001, when our sample was drawn, about 50,000 individuals were registered in the Applicant database. This corresponded to about 30 percent of the number of unemployed according

to the Labor Force Survey. The monthly inflow of new individuals in the database was about 11,000 individuals.

The Applicant database is open for employers who are recruiting, provided that they are registered employers in the public registers and in AMS's internal customer register. If an employer finds a potential candidate in the pool of job searchers in the database, she is free to contact the candidate. In some cases the contact can occur outside the system, e.g. by an e-mail to the job searchers private address, and the contacts are not registered. According to AMS, however, the most usual way of contacting is by e-mail to the job searcher's mailbox within the Applicant database. These contacts are registered in the database.

The most important feature of the Applicant database, for our purposes, is that the individual job searcher can choose not to disclose all personal information. This option allows individuals to censor information on their name, sex and age. In practice, since there is no separate entry for ethnicity, this means that individuals can choose to censor information on age, sex and ethnicity. This option was primarily introduced as a service to employed job searchers, who did not want their employers to find out that they were looking for other jobs. The presence of "blind" observations concerning some key variables is the cornerstone of our identification strategy further discussed below. A second important feature of the data is that we observe all the information that the employers observe.

3 The data

The Applicant database has not been readily available for research purposes. In order to get access to the data we had to obtain permission from each individual job searcher. This was achieved, in cooperation with AMS, by adding an introductory page to the Applicant database. This page contained a question about whether the job searchers were willing to permit that the data was used for research purposes. All individuals that were or became users of the applicant database got this question the first time they logged in to the database from March 1, 2001. If they then agreed to "participate", they got two additional questions directly motivated by our research topic:

1. Are you a male or a female?

2. Do you think that employers in general perceive your name as Swedish or foreign?

The answers to these questions were needed to get information on sex and "ethnicity" for individuals who had exercised their option to censor these entries in the applicant database.

The primary data used in this paper was collected in March 2001. It consists of all individuals who accepted to participate among those who were in the database and logged in to the database between March 1 and March 12. Approximately 50 percent of those who logged in during this time period accepted to participate, resulting in a sample of 8,666 individuals. Because we did not want to include youth in secondary school in the sample, we excluded all individuals aged below 20.[1] That gives us the sample used in this study consisting of 8,043 individuals.

The sample characteristics are reported in Table 1. The first column refers to the entire sample, while the second column refers to individuals who have censored information on gender and/or name. In the full sample we note that the average duration in the database is over 34 weeks and that a third of the sample has been contacted by an employer at least once during their "spell". We also see that half the sample is female and that 13 percent consider themselves having a foreign name.

The number of individuals that have concealed their gender or name (in column 2) was 922, corresponding to roughly 11 percent of the full sample. There are at least three differences between the sample with blind observations and the full sample worth mentioning: i) they have shorter duration in the database, ii) they have not received as many employer contacts, and iii) they are to a larger extent low educated.

In most other respects, the two samples look pretty similar. In particular, it's worth noting that the share of females and foreign names are fairly similar across samples.

[1] Most of the applicants aged below 20 look for work during the summer break or temporary work on school holidays etc. Therefore, it seems natural to exclude them in our empirical investigation.

Table 1 Descriptive statistics, means

Variable	Full sample	Blind observations only (name or sex)	LINDA (Händel)
Contacted	0.341	0.293	-
Duration (weeks)	34.5	25.7	58,7
Education:			
Primary	0.079	0.172	0.228
Secondary (gymnasium)	0.489	0.372	0.616
University	0.439	0.456	0.156
Good language skills:			
Swedish	0.969	0.966	-
English	0.561	0.498	-
French, Spanish or German	0.197	0.192	-
Good computer skills	0.738	0.629	-
Managerial experience	0.343	0.344	-
Telecommuting experience	0.124	0.124	-
Research experience	0.054	0.057	-
≥ 5 years work experience	0.421	0.393	0.298
Drivers license	0.788	0.772	-
Region:			
Stockholm	0.293	0.304	0.089
Uppsala	0.089	0.087	0.023
Södermanland	0.078	0.066	0.033
Östergötland	0.080	0.073	0.053
Jönköping	0.059	0.047	0.038
Kronoberg	0.046	0.036	0.021
Kalmar	0.049	0.047	0.031
Gotland	0.020	0.013	0.008
Blekinge	0.046	0.034	0.020
Skåne	0.187	0.149	0.131
Halland	0.075	0.044	0.041
Västra Götaland	0.182	0.144	0.190
Värmland	0.049	0.042	0.042
Örebro	0.066	0.061	0.034
Västmanland	0.074	0.060	0.033
Dalarna	0.052	0.039	0.043
Gävleborg	0.055	0.042	0.045
Västernorrland	0.042	0.023	0.037
Jämtland	0.021	0.021	0.021
Västerbotten	0.041	0.030	0.028
Norrbotten	0.031	0.017	0.041
Preferred occupations:			
Elementary occupations (Amsyk 9)	0.105	0.064	0.103
Legislators, senior officials and managers (Amsyk 1)	0.030	0.030	0.014

Table 1 (*continued*)

Variable	Full sample	Blind observations only (name or sex)	LINDA (Händel)
Professionals (Amsyk 2)	0.279	0.280	0.090
Technicians and associate professionals (Amsyk 3)	0.290	0.253	0.104
Clerks (Amsyk 4)	0.248	0.178	0.143
Service workers and shop sales workers (Amsyk 5)	0.190	0.134	0.309
Skilled agricultural and fishery workers (Amsyk 6)	0.021	0.011	0.026
Craft and related trades workers (Amsyk 7)	0.116	0.085	0.102
Plant and machine operators and assemblers (Amsyk 8)	0.100	0.062	0.102
Foreign name	0.134	0.152	0.206
Female	0.487	0.474	0.584
Age	33.8	34.5	41.0
Age 20-25	0.289	0.279	0.091
Age 26-35	0.331	0.316	0.259
Age 36-50	0.279	0.287	0.374
Age 50-	0.101	0.118	0.256
Employed	0.490	0.441	0.357
Unemployed	0.385	0.459	0.520
University student	0.081	0.074	0.087
In other training	0.040	0.022	0.054
On child leave	0.009	0.011	0.028
Blind name	0.033		-
Blind gender	0.084		-
Blind age	0.084		-
Blind name * Foreign name	0.007		-
Blind gender * Female	0.041		-
Blind age * Age > 45 years	0.029		-
# Observations	8,043	922	26,532

An issue that arises naturally here is the question of representativity. To what population can we possibly generalize our results? There are several steps in the selection process on which we have very little information. First, both employed and unemployed individuals choose whether to register in the database. This selected sample may well be very different from the typically used samples of unemployed. Second, individuals were free to choose whether

to release their data for research. We have no way of assessing this selection process.

One way of assessing the specificity of our sample is to compare it with a random sample of job searchers. In the third column of Table 1 we report the mean characteristics of the stock of job searchers in 2001 using data from the unemployment register (Händel) in LINDA (Edin & Fredriksson, 2000). There are some distinctive differences between the two groups of job searchers. We find that our sample is younger, more educated, and has more work experience. We also have a smaller share of females and minorities in our sample.

One explanation of these differences is that the individuals in our sample have much shorter job search duration, i.e. we compare high quality individuals in the Applicant database to low quality individuals in LINDA. In Table A1, we account for these effects by comparing inflows instead of stocks. The two first columns show that the difference between the samples decreases if we compare the inflow into the Applicant database to the inflow into LINDA. The similarities are even more striking in the last two columns of Table A1, where we compare the inflows of unemployed into the two bases. This is because an unemployed individual who register at the Employment Office is encouraged by the caseworker to join the Applicant database. Participation is not forced upon the individual but simply recommended; there are no sanctions should the client refuse. However, the vast majority of the people who register also choose to join the base.

Concerning the representativity of our results, this indicates that our results have some external validity to the unemployed population in Sweden. However, there are other selection issues as well. For example, there may be differences in the left-out variables between those who agreed to participate in this study and those who did not. This should be kept in mind when drawing inferences from our study to broader settings.

4 Empirical results

The empirical strategy of this paper is two-folded. In our baseline analysis we rely on the assumption of selection on observables and estimate a simple linear probability model of the form

$$P_i = \alpha + \beta' F_i + \theta' X_i$$

where P is the probability of receiving at least one employer contact, F is a vector of characteristics that we believe may be subject to discrimination (female, foreign name and age), and X is a vector of individual characteristics including information on job preferences and a quadratic in duration in the Applicant database.

Under our maintained assumptions, this simple procedure provides an estimate of β that can be interpreted as a measure of discrimination. However, even if we have access to all information available to employers, we cannot rule out that our empirical specification is not properly specified. In particular, it is very difficult to introduce the information contained in the "personal letter" of the job applicants in a quantitative model. Therefore, we also apply a second empirical strategy.

The second approach is inspired by the work of Goldin & Rouse (2000). We make use of the fact that some individuals have concealed their gender, age and (foreign) name in a "differences-in-differences" framework. We write our estimating equation as

$$P_i = \alpha + \beta' F_i + \gamma' B_i + \delta'(F_i * B_i) + \theta' X_i$$

where B is vector of variables showing what characteristics are concealed. The parameter of interest here is δ, the vector of coefficients on the interactions between F and B. There are three interactions; between female and concealed gender, between foreign name and concealed name, and between age and concealed age. Under some additional assumptions, the coefficients of these interactions measure the change in the probability of receiving an employer contact that e.g. a female experiences by concealing her gender.

The key assumption here is that there are no systematic differences in the selection (on left-out variables) into "blindness" across groups. To get an indication whether this assumption is valid, we have estimated linear probability models of concealed identity (see A2 in the Appendix). The effects of the observable characteristics are similar across sexes; only four of the 55 are significantly different.[2] The fact that the observable variables determine

[2] Formally, including interaction terms of gender with all the other explanatory variables does not make our model significantly better (F-value of 1.28, p-value of 9 percent).

"blindness" in the same way across groups may support the assumption that the effect of potential left-out variables is the same across groups as well.

The vector of coefficients on B, γ, captures the change in contact probability that applicants face by not disclosing different parts of their identity (i.e. name, gender or age). These effects probably consist of several things. For example, they might reflect discrimination; given that discriminating employers understand that a share of "blind" applicants consists of individuals from the group that is discriminated against, these employers will be resistant to contact an applicant who has not revealed his/her identity. In addition, noting that the option of concealing the identity was introduced as a service to employed job searchers who desired anonymity, the effects may partly capture employers' preferences towards employed applicants.

We start our empirical analysis by showing some further descriptive information. In Table 2 we report the share of individuals in four groups that have been contacted at least once by an employer. It turns out that the share of women that have been contacted is about 7 percentage points lower than for men. Similarly, individuals with foreign names have a 3 percentage point lower share than individuals with a Swedish name. The issue in the remainder of this section is to what extent these differences in employer contacts reflect discrimination of women and ethnic groups.

Table 2 Employer contacts by group

Group	Contact	# Observations
Males	0.378	4,127
Females	0.302	3,916
Swedish name	0.346	6,965
Foreign name	0.310	1,078

The main results of our analysis are presented as linear probability models of employer contacts in Table 3.[3] In the first column we report estimates from our first specification that relies on the assumption of selection on observables. Here we restrict ourselves to the sub-sample of individuals with no concealed

[3] Using Logit models we obtain the same qualitative results.

information. The dependent variable is the probability of having been contacted at least once by an employer. The estimates for the control variables show that the contact rate is increasing at a decreasing rate with duration in the database and is increasing with different measures of skills. A higher level of completed education, or more labour market experience, has a clear positive effect on the probability to get contacted. Also, employed applicants face significantly higher probabilities of getting a contact.[4]

Turning to our variables of interest, it is evident that the age of the applicant is strongly related to the contact rate. An applicant above age 50 is 11 percentage points less likely to have been contacted by an employer compared to an applicant age 25 or less. There is also a significant gender difference. Females have a 4.7 percentage points lower contact rate than males. However, we find no strong association between foreign names and the contact rate. Our estimates indicate a 1 percentage point disadvantage for applicants with foreign names, but this estimate is not statistically significant.

The absence of a significant differential in contact rates between applicants with Swedish and foreign names may seem surprising, but we suspect that this is at least partly a result of measurement errors. Our indicator for foreign names does not distinguish between names of different national or ethnic origin. Consequently, labor immigrants from the Nordic countries and Western Europe are lumped together with refugee immigrants from Africa and the Middle East. This aggregation results in a very heterogenous group of "immigrants". The included groups differ greatly in terms of labor market outcomes (see e.g. Edin & Åslund, 2001).

Taking the results in column 1 at face value, we find that employer are using age and gender as a screening device in hiring in a way that clearly indicate discriminatory behavior. However, this interpretation relies crucially on the maintained assumption of selection on observables. Even if we are in the unusually favorable situation of having the same information as the employers, we are still dependent on having a correctly specified model. The most obvious problem is to handle the personal letter written by the job applicant. Our estimates seem robust to the inclusion of various quantitative measures of the

[4] Eriksson and Lagerström (2004) provide an analysis of whether firms view employment status as an important signal for productivity that can explain the persistence of unemployment.

letter.[5] Still, we cannot argue that we can capture all relevant information in our specification. Therefore, we turn to our second identification strategy.

Table 3 Linear probability models of employer contact

	Non-blind sample	Full sample	Female dominated occupations	Male dominated occupations
Foreign name	-.010	-.019	0.019	0.113
	(.015)	(.014)	(.028)	(.056)
Female	-.047	-.051	-.002	-.218
	(.011)	(.011)	(.024)	(.038)
Over 50 years of age	-.113	-.099	-.088	-.123
	(.022)	(.020)	(.041)	(.068)
36-50 years of age	-.079	-.076	-.090	-.073
	(.016)	(.014)	(.031)	(.048)
26-35 years of age	-.032	-.029	-.004	-.056
	(.013)	(.012)	(.026)	(.043)
Blind name	-	.031	.037	.137
		(.033)	(.075)	(.118)
Blind gender	-	-.005	-.064	-.004
		(.020)	(.047)	(.064)
Blind age	-	-.013	-.023	.032
		(.024)	(.049)	(.096)
Blind name * Foreign name	-	.051	-.166	-.039
		(.068)	(.178)	(.280)
Blind gender * Female	-	.057	.145	.185
		(.029)	(.064)	(.167)
Blind age * Over 50 years	-	.042	.102	.100
		(.037)	(.087)	(.157)
# observations	6,657	8,043	1,837	703
R^2	0.2780	0.2819	0.2319	0.3264

Note: Standard errors in parentheses. Controls for other personal characteristics, region of residence and preferred occupations are included (for more detail, see Table A3). Female (male) dominated occupations are defined as the three occupations where women (men) are most likely to apply for jobs, relative to the other sex.

In the second column of Table 3 we use the full sample and utilize the interactions between characteristics and concealed information to identify

[5] In Table A4, we report estimates where we have extended the model with 1) the length of the private letter, 2) the numbers of unknown words/spelling errors (using a spell check), and 3) whether a private e-mail address was included.

potential discrimination. A first observation is that the effects of control variables and the main effect for our variables of interest are very similar to those in column 1. Interestingly, there seems to be no effects of concealing information on the contact rate. None of the main effects (blind name, blind gender, and blind age) is statistically significant and the point estimates are fairly small.

Turning to the parameters of interest, we see that only the interaction effect for women is significant. It indicates that a woman's chance of receiving an employer contact increases by 5 percentage points if she conceals her gender. Thus, women can undo their lower contact rate by concealing their gender.

The estimates of the interaction effect foreign names and those over 50 years of age are similar in magnitude, but not statistically significant. Once again, we need to consider the role of measurement errors. It turns out that this may be a serious problem with the interaction with foreign names, where only about 50 percent of the "blind foreign names" are truly blind. We were able to accurately identify the other half using for example rare language skills or the personal letter in the database. This will of course introduce potentially serious attenuation bias in our estimate of the effect of having a foreign name. Similarly, information on work experience may be a way of identifying older applicants. For the female applicants with "blind gender", the share that is truly blind is higher and the attenuation bias smaller since we have found it harder to identify the gender using for instance working experiences or skills.

In the final two columns of Table 3 we report separate estimates for occupations with different gender composition of applicants. Earlier studies suggest that the degree of gender discrimination may depend on the gender composition of the industry and/or occupation. For example, using data from a field experiment, Riach & Rich (2006) find evidence of discrimination against males in a female occupation (secretary), and females in a male occupation (engineer). In order to investigate this we singled out the three most female and male dominated occupations in our sample.[6] The male occupations are "Legislators, senior officials and managers", "Craft and related trades workers", and "Plant and machine operators and assemblers". The female

[6] We have defined these as the occupations with the largest relative difference across gender in the probability to apply in an occupation.

occupations are “Clerks”, “Service workers and shop sales workers”, and “Elementary occupations”.

Our results lend some support to the hypothesis that discrimination against females is more important in male occupations. The main effect of gender is very large, 22 percentage points lower contact rates, and statistically significant. In female dominated occupations, on the other hand, there is no evidence of discrimination against females. These result are not so clear using the blind observations as an additional “robustness check”. The interactions between concealed gender and females are large and positive, but the standard errors are also large.

5 Concluding remarks

In this paper we use data generated from a “policy experiment” conducted at the Swedish public employment offices. Individuals registered at these offices can post their qualifications in a database available to employers over the Internet. Potential employers are free to search this database for job candidates and contacts between employers and candidates are recorded. We use two complementary identification strategies. First, since our data contain all information available to employers, we argue that selection on observables is viable. Second, we utilize the fact that individuals can choose not to reveal their name and gender to potential employers. By comparing the “contact rate” of censored and non-censored women and minorities we are able to investigate how employers use gender and “foreign” names as a screening device in their hiring process.

Our empirical results show that women receive less job contacts than men do even when controlling for qualifications. We also find that women that do not reveal their gender receive as many job contacts as men with similar characteristics. These results clearly demonstrate that employers use the gender of the applicant as a screening device, and we interpret this as a clear sign of discrimination.

Our empirical findings on discrimination against applicants with foreign names and older applicants are less conclusive. This is probably mainly due to weaknesses in our data concerning these two groups. Our measure of foreign names is a catch all variable that makes it impossible to look closer at this very

heterogenous group. Also, there are major measurement error problems when it comes to concealing foreign names or age.

We have found strong evidence of discrimination against females in the hiring process. Assessing the importance of this discrimination for outcomes in the Swedish labor market using these estimates is a much more difficult task. First, we have no clear "structural" interpretation of our estimate. Second, we only observe the first part of the chain of events that lead to a possible hiring. We have no idea whether the mechanism we observe is reinforced or weakened in later stages of the hiring process.

References

Albrecht, J, A Björklund & S Vroman (2003), "Is there a Glass Ceiling in Sweden?", *Journal of Labor Economics*, Vol. 21 (1), 145-177.

Becker G (1957), *The Economics of Discrimination*, University of Chicago Press.

Bertrand, M & S Mullainathan (2004), "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review*, Vol. 94, 991-1011.

Edin, P-A & O Åslund (2001), "Invandrare på 1990-talets arbetsmarknad", in *Ofärd i välfärden*, SOU 2001:54.

Edin, P-A & P Fredriksson (2000), "LINDA – Longitudinal Individual DAta for Sweden", Working Paper 2000:19, Department of Economics, Uppsala University.

Eriksson, S & J Lagerström (2004), "Competition between Employed and Unemployed Job Applicants: Swedish Evidence", forthcoming in *Scandinavian Journal of Economics*.

Goldin, C & C Rouse (2000), "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians", *American Economic Review*, Vol. 90, No. 4, 715-741.

le Grand, C (1997), "Kön, lön och yrke – yrkessegregering och lönediskriminering mellan män och kvinnor", in I Persson & E Wadensjö (eds) *Kvinnors och mäns löner – varför så olika?*, SOU 1997:136.

Heckman J (1998), "Detecting Discrimination", *Journal of Economic Perspectives* 12, 101 – 116.

Meyerson, E & T Petersen (1997), "Lika lön för lika arbete. En studie av svenska förhållanden i internationell belysning", in *Kvinnors och mäns löner – varför så olika?*, SOU 1997:136.

Riach, P A & J Rich (2002), "Field Experiments of Discrimination in the Market Place", *Economic Journal* 112, F480–F518.

Riach, P A & J Rich (2006), "An Experimental Investigation of Sexual Discrimination in Hiring in the English Labor Market", *Advances in Economic Analysis & Policy* 6, Article 1.

Appendix 1: Comparison of the characteristics of the inflows

Table A 1 Comparison of the characteristics of the inflow of unemployed in the Applicant Database and the inflow of unemployed in Händel (in fractions)

Variable	All The Applicant database	All LINDA (Händel)	Unemployed The Applicant database	Unemployed LINDA (Händel)
Highest level of completed education:				
Primary	0.17	0.34	0.29	0.34
Secondary	0.45	0.41	0.48	0.39
University	0.38	0.25	0.23	0.27
Work experience:				
None	0.30	0.24	0.43	0.36
Some or long	0.70	0.66	0.57	0.64
Age:				
Mean (years)	31.1	35.1	30.5	33.4
Age 20-25	0.39	0.32	0.43	0.38
Age 26-35	0.33	0.21	0.30	0.23
Age 36-50	0.22	0.32	0.21	0.26
Age 51-	0.06	0.15	0.06	0.14
Gender:				
Female	0.49	0.47	0.41	0.43
Ethnicity:				
Foreign name	0.16	0.28	0.19	0.34
Region:				
Stockholm	0.22	0.18	0.18	0.19
Uppsala	0.06	0.03	0.06	0.04
Södermanland	0.04	0.04	0.03	0.05
Östergötland	0.05	0.07	0.05	0.06
Jönköping	0.03	0.04	0.04	0.04
Kronoberg	0.02	0.02	0.02	0.02
Kalmar	0.02	0.04	0.02	0.05
Gotland	0.01	0.00	0.01	0.00
Blekinge	0.02	0.00	0.02	0.00
Skåne	0.11	0.11	0.10	0.11
Halland	0.04	0.03	0.04	0.03
Västra Götaland	0.12	0.18	0.13	0.19
Värmland	0.03	0.03	0.04	0.03
Örebro	0.04	0.03	0.04	0.03
Västmanland	0.06	0.03	0.06	0.03
Dalarna	0.03	0.02	0.03	0.01
Gävleborg	0.03	0.03	0.03	0.03
Västernorrland	0.02	0.03	0.02	0.03

Table A 1 (*continued*)

Variable	All The Applicant database	All LINDA (Händel)	Unemployed The Applicant database	Unemployed LINDA (Händel)
Jämtland	0.01	0.02	0.01	0.01
Västerbotten	0.02	0.03	0.03	0.03
Norrbotten	0.03	0.04	0.03	0.04
Preferred occupations:				
Legislators, senior officials and managers (Amsyk 1)	0.02	0.03	0.01	0.04
Professionals (Amsyk 2)	0.21	0.15	0.16	0.17
Technicians and associate professionals (Amsyk 3)	0.19	0.08	0.18	0.07
Clerks (Amsyk 4)	0.17	0.12	0.18	0.11
Service workers and shop sales workers (Amsyk 5)	0.15	0.26	0.20	0.25
Skilled agricultural and fishery workers (Amsyk 6)	0.02	0.02	0.02	0.02
Craft and related trades workers (Amsyk 7)	0.08	0.10	0.11	0.11
Plant and machine operators and assemblers (Amsyk 8)	0.07	0.12	0.10	0.09
Elementary occupations (Amsyk 9)	0.10	0.13	0.14	0.14
# observations	1,285	797	538	588

Note: The data from the bases is for the inflow into unemployment in March 2001. The variable "foreign name" in the Applicant database is compared to the variable "being born in a country other than Sweden" in Händel. The regions and the preferred occupations sum to more than one in the Applicant Database, since it is possible to apply for several jobs.

Appendix 2: Comparison of the selection into "blindness"

Table A 2 Linear probability models of concealed sex, by sex

	Full sample	Men	Women
Duration in the data base (weeks)	-.002	-.002	-.002
	(.0002)	(.0003)	(.0003)
Duration in the data base^2/100	.000007	.000008	.000007
	(.000001)	(.000002)	(.000002)
Foreign name	-.014	-.014	-.017
	(.009)	(.012)	(.013)
Female	-.002	-	-
	(.007)		
Over 50 years of age	-.003	-.006	-.009
	(.013)	(.018)	(.020)
36-50 years of age	-.006	-.009	.0001
	(.010)	(.014)	(.014)
26-35 years of age	-.012	-.023	-.0004
	(.008)	(.012)	(.012)
Education:			
Secondary (Gymnasium)	-.118	-.120	-.113
	(.013)	(.017)	(.019)
University	-.106	-.116	-.094
	(.014)	(.019)	(.021)
Good language skills:			
Swedish	-.003	.014	-.034
	(.018)	(.022)	(.029)
English	-.019	-.018	-.020
	(.007)	(.010)	(.010)
French, Spanish or German	.001	-.007	.008
	(.008)	(.012)	(.011)
Good computer skills	-.038	-.030	-.042
	(.007)	(.011)	(.010)
Managerial experience	.004	-.013	.023
	(.007)	(.010)	(.011)
Telecommuting experience	.007	.015	-.002
	(.010)	(.012)	(.016)
Research experience	.007	.027	-.023
	(.014)	(.018)	(.022)
≥ 5 years work experience	-.002	.008	-.008
	(.008)	(.011)	(.011)
No work experience	.064	.091	.034
	(.010)	(.013)	(.014)
Labor market status:			
Employed in preferred occupation	-.022	-.016	-.028
	(.007)	(.009)	(.010)

Table A 2 (*continued*)

	Full sample	Men	Women
University student	-.047	-.080	-.021
	(.013)	(.019)	(.017)
In other training	-.062	-.059	-.068
	(.016)	(.022)	(.022)
On child leave	-.003	.043	-.003
	(.032)	(.154)	(.034)
Drivers license	.004	-.003	.009
	(.008)	(.012)	(.011)
Constant	.302	.299	.316
	(.023)	(.029)	(.036)
# observations	8,043	4,127	3,916
R^2	0.071	0.096	0.060

Note: Standard errors in parentheses. Controls for regions of residence and preferred occupations included.

Appendix 3: Baseline models

Table A 3 Linear probability models of employer contact

	Non-blind sample	Full sample	Female dominated branches	Male dominated branches
Duration in the database (weeks)	.011	.010	.009	.010
	(.0003)	(.0003)	(.001)	(.001)
Duration in the data base^2/100	-.004	-.004	-.003	-.003
	(0.0002)	(0.0002)	(0.0005)	(0.0009)
Foreign name	-.010	-.019	0.019	0.113
	(.015)	(.014)	(.028)	(.056)
Female	-.047	-.051	-.002	-.218
	(.011)	(.011)	(.024)	(.038)
Over 50 years of age	-.113	-.099	-.088	-.123
	(.022)	(.020)	(.041)	(.068)
36-50 years of age	-.079	-.076	-.090	-.073
	(.016)	(.014)	(.031)	(.048)
26-35 years of age	-.032	-.029	-.004	-.056
	(.013)	(.012)	(.026)	(.043)
Blind name	-	.031	.037	.137
		(.033)	(.075)	(.118)
Blind gender	-	-.005	-.064	-.004
		(.020)	(.047)	(.064)
Blind age	-	-.013	-.023	.032
		(.024)	(.049)	(.096)
Blind name * Foreign name	-	.051	-.166	-.039
		(.068)	(.178)	(.280)
Blind gender * Female	-	.057	.145	.185
		(.029)	(.064)	(.167)
Blind age * Over 45 years	-	.042	.102	.100
		(.037)	(.087)	(.157)
Education:				
Secondary (Gymnasium)	.014	.022	.012	-.045
	(.019)	(.017)	(.028)	(.048)
University	.045	.053	.089	-.004
	(.021)	(.019)	(.036)	(.062)
Good language skills:				
Swedish	.025	.011	.058	-.017
	(.027)	(.025)	(.041)	(.072)
English	.034	.032	.022	.005
	(.011)	(.010)	(.021)	(.036)
French, Spanish or German	.031	.031	.038	.080
	(.014)	(.013)	(.029)	(.058)
Good computer skills	.013	.012	.028	.011
	(.012)	(.011)	(.022)	(.033)

Table A 3 (*continued*)

	Non-blind sample	Full sample	Female dominated branches	Male dominated branches
Managerial experience	.037	.052	.059	-.034
	(.012)	(.011)	(.027)	(.042)
Telecommuting experience	.026	.025	-.072	.075
	(.017)	(.015)	(.044)	(.059)
Research experience	.015	.005	.174	-.021
	(.024)	(.022)	(.128)	(.161)
≥ 5 years work experience	.034	.024	-.003	.078
	(.013)	(.012)	(.025)	(.039)
No work experience	-.017	-.030	-.026	-.069
	(.015)	(.013)	(.032)	(.046)
Labor market status:				
Employed in preferred occupation	.027	.032	.055	.018
	(.011)	(.010)	(.021)	(.033)
University student	-.032	-.025	-.041	-.176
	(.020)	(.018)	(.053)	(.060)
In other training	.024	.024	.059	-.015
	(.025)	(.023)	(.049)	(.073)
On child leave	.069	.059	.153	.188
	(.057)	(.050)	(.080)	(.090)
Drivers license	.011	.005	.015	.077
	(.013)	(.012)	(.022)	(.041)
Constant	-.020	-.003	-.146	-.146
	(.033)	(.031)	(.050)	(.050)
# observations	6,657	8,043	1,837	703
R^2	0.2780	0.2819	0.2319	0.3264

Note: Standard errors in parentheses. Controls for regions of residence and preferred occupations included. The female (male) dominated branches consist of the three branches where women (men) are most likely to apply for jobs, relative to the other sex.

Appendix 4: Extended models

Table A 4 Linear probability models of employer contact

	Non-blind sample	Full sample	Female dominated branches	Male dominated branches
Duration in the database (weeks)	.011	.010	.009	.010
	(.0004)	(.0003)	(.0008)	(.001)
Duration in the data base^2/100	-.004	-.004	-.003	-.004
	(0.0002)	(0.0002)	(0.0005)	(0.0004)
Foreign name	-.012	-.020	-.017	0.110
	(.015)	(.014)	(.028)	(.056)
Female	-.046	-.051	-.001	-.220
	(.011)	(.011)	(.024)	(.039)
Over 50 years of age	-.113	-.100	-.086	-.119
	(.022)	(.020)	(.041)	(.068)
36-50 years of age	-.079	-.077	-.088	-.072
	(.016)	(.014)	(.031)	(.048)
26-35 years of age	-.032	-.029	-.003	-.054
	(.013)	(.012)	(.026)	(.043)
Blind name	-	.030	.034	.130
		(.033)	(.075)	(.117)
Blind gender	-	-.005	-.063	-.001
		(.020)	(.048)	(.063)
Blind age	-	-.012	-.025	.027
		(.024)	(.049)	(.098)
Blind name * Foreign name	-	.051	-.139	-.033
		(.068)	(.178)	(.283)
Blind gender * Female	-	.057	.146	.192
		(.029)	(.064)	(.168)
Blind age * Over 45 years	-	.042	.101	.101
		(.037)	(.087)	(.158)
Education:				
Secondary (Gymnasium)	.014	.022	.012	-.047
	(.019)	(.017)	(.029)	(.048)
University	.046	.052	.088	-.004
	(.021)	(.019)	(.036)	(0.061)
Good language skills:				
Swedish	.027	.013	.058	-.024
	(.027)	(.025)	(.041)	(.071)
English	.034	.032	.023	.009
	(.011)	(.010)	(.021)	(.036)
French, Spanish or German	.030	.031	.037	.075
	(.014)	(.013)	(.029)	(.059)
Good computer skills	.014	.013	.030	.017
	(.012)	(.011)	(.022)	(.034)