IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Discrimination, sickness absence, and labor market policy

Jonas Lagerström

DISSERTATION SERIES 2006:4

Presented at the Department of Economics, Uppsala University

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

This doctoral dissertation was defended for the degree of Doctor in Philosophy at the Department of Economics, Uppsala University, June 14, 2006. The first and third essay of this thesis contain revised versions of research previously published by IFAU as Working Paper 2004:2 and Working paper 2006:4.

ISSN 1651-4149

Doctoral dissertation presented to the Faculty of Social Sciences 2006

Abstract

LAGERSTRÖM, Jonas, 2006. Discrimination, Sickness Absence, and Labor Market Policy. Department of Economics, Uppsala University, *Economic studies* 96, 105 pp., ISBN 91-85519-03-0.

This dissertation consists of an introduction and four self-contained essays:

Essay 1 (with Stefan Eriksson) investigates empirically whether being unemployed per se reduces the probability of getting contacted by a firm. Individuals registered at the Swedish employment offices post their qualifications in a database available to employers over the Internet. Since we have access to exactly the same information as the firms, we can minimize the problems associated with unobserved heterogeneity. Our results show that an unemployed applicant faces a lower contact probability, and receives fewer contacts, than an otherwise identical employed applicant, thus supporting the notion that firms view employment status as a signal for productivity.

Essay 2 evaluates an experiment where employees at randomly chosen establishments received half a day off if they completed a full calendar month without any sick-leave. Using individual panel data, the absence rates of these individuals are compared to the absence rates of individuals at establishments with no such program before, during and after the treatment periods. Overall, the bonus caused a sharp reduction in absenteeism, especially for women, highly educated individuals and part-time workers.

Essay 3 (with Per-Anders Edin) provides evidence on discrimination in the hiring process. We use Internet data generated from a "policy experiment", in which 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 find that women have a 15 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.

Essay 4 examines if and how the personnel at the Swedish Employment Office matter. The analysis shows that caseworkers explain a substantial part of future outcomes in terms of employment status and earnings, but have no significant effect on wages. Caseworkers that send their clients to classroom training or on-the-job training are less successful than caseworkers that provide basic job-search assistance. If caseworkers' preferences in previous years towards treatments are uncorrelated with unobserved present working strategies, these estimates correspond to causal treatment effects.

Keywords: Discrimination, unobserved heterogeneity, absenteeism, social experiment, active labor market policy, program evaluation

Jonas Lagerström, Department of Economics, Uppsala University, Box 513, SE-751 20, Uppsala, Sweden,

CONTENTS

ACKNOWLEDGEMENTS

Several people have contributed to this thesis. First of all, I am very grateful to my advisors Per-Anders Edin and Peter Fredriksson. During my last year as an undergraduate, P-A became the first teacher to talk about important economical issues in a way that I could truly grasp and appreciate. His spooky empirical intuition and his clever ways of looking at data – plus, I confess, a nerve-racking 19 percent unemployment rate for young economists – convinced me to invest years in studying the labor market instead of actually being a part of it myself. Despite some ups and downs during this bumpy journey, that interest is still solid, and occasionally even rising. Our work on discrimination has survived many expiring deadlines, to the frustration of many, but has still taught me much about research in practice and everyday priorities as an economist. Peter has helped me a lot with comments and suggestions. I thank you both.

I would also like to thank friends at the Department in Uppsala: Jovan Zamac, Pär Holmberg, Martin Ågren, Anki Nilsson, Patrik Hesselius, Magnus Gustavsson, Pär Österholm, Fredrik Johansson, Christian Andersson, Erik Post, Iida Häkkinen, Hanna Ågren, and all the rest. In particular, I would like to thank my roommate Martin Söderström. Spending years with you in a tiny office with no air to breath and struggling with frustrating economic mumbo-jumbo none of us was even close to understanding turned out to be… great! I will look back at these years with a smile on my face, we will keep in touch. A special thanks to Stefan Eriksson for being a great co-author on one of the papers in this thesis – and on more work in progress.

In spite of all these excellent people, my presence at the Department has been embarrassingly low (not even the great administrative help from Katarina, Monica, Åke and Eva has kept me in Uppsala). For this, I blame a number of people: Markus Jäntti invited me to work and teach at Åbo Akademi University in Finland, with a huge and cool office surrounded by very nice people. I have learned a lot. Thanks to Ville, Eeva, Tinke and Ralf – and especially Angela, who inspired me to work harder. Besides these people, I also blame some people at the University of California at Berkeley. Spending half a year in the U.S. turned out to be very awarding and very interesting.

Needless to say, I am also very grateful to my family. But the greatest thanks of all goes to Elina. Now, finally, our life together begins for real.

Uppsala, April 2006

Jonas Lagerström

INTRODUCTION

Establishing causal relationships is an important but difficult part of empirical labor economics. The fact that individuals with much schooling tend to have higher earnings does not necessarily imply that an additional year of schooling causes earnings to rise. We would observe the same pattern if higher earnings caused people to acquire more schooling (i.e. reverse causation) or if there were some unobserved variables, such as motivation, causing individuals to have both higher earnings and more schooling (i.e. confounding variables).

In order to assess if schooling per se increases earnings we need to somehow account for all these other factors. In the medical sciences, randomized experiments are often carried out to sidestep such omitted variable problems. By assigning treatment randomly it is unlikely that other variables differ systematically between treated and non-treated individuals, and any difference in outcomes can (given some reasonable assumptions) be interpreted as the causal effect of treatment. In economics, however, experiments are rare and in many cases impossible to conduct. As economists, we therefore have to find other ways to uncover the true relationship between variables.

This thesis consists of four self-contained essays, each with a different strategy to overcome the obstacles of estimating causal effects in empirical labor economics. Every essay deals with a major social problem in Sweden, thoroughly and sometimes divisively discussed in the public debate for many years: Are women and ethnic minorities discriminated against when they apply for jobs? Do employers prefer employed workers to unemployed workers even if they are otherwise identical? Are economic incentives important when people decide whether to call in sick? Do employment officials really make a difference, and if so, why are some personnel more successful than others?

Even though these questions have no clearly unifying theme, the papers have one feature in common. I have tried to overcome the omitted variable bias by finding and collecting previously unexploited data sets, where the key variable X has been manipulated in a way that had no effect on outcome Y other than through the induced change in X. This quasi-experiment approach focuses on understanding the source of variation used to estimate the key parameters (e.g. Meyer, 1995). If we cannot randomize which kid that enrolls in an extra year of schooling, at least we should understand the source of this variation.

I find this approach rewarding for several reasons. First, probing empirical findings with new data sets helps us to judge the fragility of our 'knowledge'. Second, the pitfalls when drawing inferences from non-experimental data are innumerable: omitted variables, measurement errors, reverse causality, and incorrect functional form to mention just a few. Program evaluations attempting to establish causal effects need to address and account for these effects. The alternative if experimental variation is unavailable is to use complicated econometric methods and to employ economic theory as well as statistical modeling to try to 'mimic' the experimental design. However, this approach rests on a number of untestable assumptions. Third, being an economist with severe problems grasping concepts such as 'the joint distribution of u and v', I have found it easier and intuitively more appealing to improve the quality of the underlying data by collecting new data where we can expect, a priori, to find exogenous variation in the variable under study.

Essay I, *Competition between Employed and Unemployed Job Applicants: Swedish Evidence* (written with Stefan Eriksson) investigates empirically whether employers prefer to hire employed applicants rather than unemployed applicants. Previous studies have demonstrated that such a pattern – if existing – can partly explain the persistence of unemployment. The competition for jobs will suffer if firms systematically favor already employed individuals, thereby preventing wages from falling and keeping unemployment rates high. Long-term unemployment has indeed become a major social problem in Sweden, as well as in large parts of Europe. Compared to the U.S., unemployment rates are higher, turnover is lower and the adjustment back to equilibrium is slower. Understanding how employers view a history of unemployment is therefore important not only to the job-seeker deciding when to accept a job offer and when to keep on searching, but to society at large.

To measure empirically whether firms screen against unemployed individuals is complicated. Simply asking them is an obvious and commonly used method in the literature (e.g. Agell and Lundborg, 2003), but employers that do use unemployment as a hiring criterion might be unwilling to admit so in a survey. Other studies have used administrative data. Although register data often contain very large samples and show what people *do* instead of what they claim they *would* do, these data are usually not generated to address a specific research question. Unemployed and employed individuals differ in characteristics – other than employment status – that also have an effect on their labor market outcomes. In order to estimate the causal effect of employment status we need to control for all these confounders (e.g. social skills, motivation, and search intensity).

An alternative way to 'control' for these confounders is to gather new information from a situation where these factors 'by construction' cannot influence the hiring decision. Exploiting data from a large Internet job-search channel, we have access to exactly the same information as the employers have when they decide which job-seeker to contact. Given that we control properly for all this information, we should be able to get an unbiased estimate of how employers screen against unemployed individuals.

We find that an unemployed applicant faces a significantly lower contact probability, and receive fewer contacts, than an otherwise identical employed applicant. For the 'average' searcher, contact probability is about 9 percent lower if she is unemployed instead of employed. Although it is difficult to compare the magnitudes of the effects across different studies, our results are somewhat larger than the estimates from survey studies and smaller than studies that do not control entirely for unobserved heterogeneity.

Essay II, *Economic Incentives, Working Environment and Sickness Absence: Evidence from a Randomized Experiment*, investigates how economic incentives and working environment affect absence behavior. The high level of sickness absence has been one of the most debated issues in the Swedish political debate. About 4 percent of the employees in Sweden report sickness absence exceeding one week – numbers dramatically higher than in most OECD countries. The booming economic and social costs due to this rise in absenteeism have generated an outburst of proposals for strategies to reverse the trend.

Absence data generally show that employees working in a good environment and under strong economic incentives to work have lower sickness absence rates than other employees. However, potential selection effects and reversed causation make it hard to assess whether these factors have actually *caused* absence to fall.

This study attempts to overcome these obstacles by collecting and exploiting data from a unique small-scale experiment. The individuals assigned to the control group were excluded from all treatment, whereas all other individuals were exposed to one of two separate treatments. First, employees could (during 3-12 months) earn half a day off if they completed the full calendar month without any sick-leave. Second, randomly chosen establishments received increased means to improve their working environment.

Using individual panel data, the absence rates of the individuals in the treatment groups are compared to the absence rates of individuals with no such program before, during and after the treatment periods. Although social experiments are no ‘quick fix’, they do provide the most desirable mechanism for manipulating X; this manipulation is indeed likely to change X in a way that does not affect Y except through the changes induced in X.

The results provide strong evidence on a causal link between the cost of being absent and sickness absence. Employees with the possibility to earn half a day off had 1.1 days less absence per calendar month that employees in the control group. The availability of multiple control groups makes it possible to further interpret the program effect and to rule out competing explanations. For example, the estimated effects are larger for employees with stronger economic incentives – supporting the notion that the incentives per se caused absence rates to fall. The effects of the working environment treatment are smaller in magnitude and not statistically significant at conventional levels.

Essay III measures the extent of discrimination facing women and ethnic minorities in the Swedish labor market. In spite of its well known equality of outcomes, the Swedish labor market still produces large differentials in labor market outcomes. For example, women earn about 80 percent of men’s hourly wage and many immigrant groups have dramatically higher unemployment rates and lower earnings than native

Swedes.[1] Discrimination is often considered to be a key factor explaining these inequalities. However, even though some studies (e.g. Goldin and Rouse, 2000) provide some evidence on discrimination, most studies in the literature are sensitive to the criticism of potentially left-out variables. How can we know for certain that an immigrant is not rejected due to, for example, poor language skills instead of his ethnic background? We cannot randomly let person A be an immigrant and person B a Swede and conclude that there is discrimination if these individuals are treated differently. However, nowadays, a substantial part of job-search and the interactions between firms and job-seekers takes place on the Internet. In general, the factors influencing these on-line markets are more limited ('what you see is what you get') than contacts in real life, thereby making the assumptions behind selection on observable models much less extreme (e.g. Ellison and Ellison, 2001; Gottlieb, 2006). In addition, *Blind Dates: Quasi-Experimental Evidence on Discrimination* (written with Per-Anders Edin), exploits a quasi-experiment at the large Internet job-search channel provided by the Swedish public employment offices. Individuals 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 enables applicants to hide part of the information concerning their identities to the potential employers. We exploit this variation in gender and name to investigate if employers discriminate against women and individuals with foreign names in their hiring process.

Our main finding is that women have a 15 percent lower chance than men of getting contacted by employers and that this differential is fully explained by discrimination. Women receive fewer contacts than men do, even when controlling for all qualifications. In addition, women that do not reveal their gender receive as many contacts as men with similar characteristics. Together, these patterns in the data provide clear evidence on gender discrimination in the Swedish labor market. Our results concerning ethnic discrimination are less conclusive, probably due to measurement errors.

Essay IV, *Caseworker Effects and Program Evaluation*, estimates the importance of the administrative personnel at the Swedish public Employment Offices.

[1] According to the labor force survey in 2003, unemployment stood at 3.9 percent among natives, compared to 18.9 percent and 15.7 percent among African and Asian immigrants, respectively.

Traditionally, Sweden has spent a larger share on active labor market policies than most other countries (e.g. Calmfors *et al*., 2001), corresponding in 1996-1999 to about 1.14 percent of GDP. Out of these expenditures, about 21 percent was allocated to public employment services and administration. Given these large public expenditures and the many job-seekers receiving these services, it is important to evaluate the effect of the personnel and the programs they implement: Do participants benefit from these programs?

The major problem when estimating the importance of personnel (e.g. caseworkers, teachers) is the fact that clients (e.g. job-seekers, students) are not randomly distributed within and across units (e.g. offices, schools). By contacting all 244 regular Employment offices, I identify 69 offices where clients are randomly assigned to caseworkers. In most of these cases, randomization was achieved by matching job-seekers to caseworkers based on the job-seekers' date of birth within a month. The results show that the personnel do make a difference: About 2-5 percent of the total variation in job-seekers' future employment status and earnings are explained by which caseworker the job-seekers are randomly assigned. However, there are no evidence on any effect on wages, suggesting that caseworkers affect primarily employment and not productivity.

Examining why some caseworkers are successful while others are not is harder. It is likely that caseworkers influence their clients in a number of ways that we can not observe in the register data. The paper shows, however, that the probability to receive different kinds of treatment differs substantially depending on which caseworker a client is randomly assigned. I exploit this plausibly exogenous variation, which – under some assumptions – provides an alternative method to estimate treatment effects even though the selection process into treatment is essentially a black box.

The analysis shows that caseworkers that send their clients to classroom training or on-the-job training programs are less successful than caseworkers that restrict themselves to provide the basic job-search assistance. Given the assumption that caseworkers' preferences in previous years towards different treatments are uncorrelated with unobserved present working strategies, these estimates correspond to causal treatments effects.

REFERENCES

Agell, J and P Lundborg (2003), "Survey Evidence on Wage Rigidity and Unemployment: Sweden in the 1990s", *Scandinavian Journal of Economics*, Vol. 105, 15-29.

Calmfors, L, A Forslund and M Hemström (2001), "Does active labour market policy work? Lessons from the Swedish experiences", *Swedish Economic Policy Review*, Vol. 8, 61-131.

Ellison, G and S Fisher Ellison (2005), "Lessons About Markets from the Internet", *Journal of Economic Perspectives*, Vol. 19, 139-158.

Freeman, R B (1989), *Labor Markets in Action: Essays in Empirical Economics*, Woodhead Faulkner Publishers, England and Harvard University Press.

Goldin, C and C Rouse (2000), "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians", *American Economic Review*, Vol. 90, 715-741.

Gottlieb, L (2006), "How Do I Love Thee?", *The Atlantic Monthly*, Vol. 297, No. 2, 58-70.

Heckman, J, R LaLonde and J Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", in O Ashenfelter and D Card, editors, *Handbook of Labor Economics*, Vol. 3, North-Holland.

Krueger, A B (2001), "Introduction: Symposium on Econometric Tools", *Journal of Economic Perspectives*, Vol. 15, 3-10.

Leamer, E (1983), "Let's Take the Con out of Econometrics", *American Economic Review*, Vol. 73, 31-43.

Meyer, B D (1995), "Natural and Quasi-Experiments in Economics", *Journal of Business & Economic Statistics*, Vol. 13, 151-161.

ESSAY I

COMPETITION BETWEEN EMPLOYED AND UNEMPLOYED JOB APPLICANTS: SWEDISH EVIDENCE*

with Stefan Eriksson

1. INTRODUCTION

Unemployed workers have to compete for jobs with on-the-job searchers.[2] Thus, a firm with a vacancy will often have several applications, from both unemployed and employed searchers, to choose from. This choice is often difficult and it is likely that firms use easily observable characteristics as sorting criterions. One such characteristic is the employment status of the applicants, and thus an important question is if being unemployed *per se* reduces the chance of getting a job. However, empirically this question is difficult to answer due to problems with unobserved heterogeneity.

The purpose of this paper is to study the relative efficiency of employed versus unemployed search. We do this by empirically investigating whether or not an unemployed searcher has a lower probability of getting contacted by a firm than an otherwise identical employed searcher using a new dataset that allows us to minimize the problems associated with unobserved heterogeneity.

This issue is important for several reasons. First, it may help us understand the evolution of aggregate unemployment. In the literature, the persistence of unemploy-

* This paper has been accepted to be published in the *Scandinavian Journal of Economics*. The authors would like to thank James Albrecht, Michael Burda, Mikael Carlsson, Per-Anders Edin, Nils Gottfries, Bertil Holmlund, Ann-Sofie Kolm, Oskar Nordström Skans, Peter Skogman Thoursie, two anonymous referees and seminar participants at Uppsala University, FIEF, the ESPE Conference in New York and the EEA Annual Congress in Madrid for valuable comments. Thanks also to AMS and Claes-Göran Lock for providing us with the data. Financial support from the Institute for Labour Market Policy Evaluation is gratefully acknowledged.

[2] The importance of on-the-job search is documented in e.g. Pissarides and Wadsworth (1994) and Boeri (1999).

ment has been explained by factors in the wage formation process that prevents wages from falling, thereby, keeping the unemployment rate high (e.g. Bean (1994)). An example is Eriksson and Gottfries (2005), who formulate an efficiency wage model, and show that if firms prefer to hire employed searchers this will lead to higher equilibrium unemployment and a slower adjustment back to equilibrium after a negative shock. Firms are reluctant to lower the wage rapidly fearing a costly rise in turnover. The results indicate that the effects are substantial. Other examples of theoretical models where firms sometimes prefer to hire employed searchers are Kugler and Saint-Paul (2004), Tranæs (2001) and Eriksson (2002).[3] Second, this issue may be important for search theory. If on-the-job search is more efficient, then the best strategy for an unemployed worker is often to accept the first offer received and then continue to search while employed, and it makes less sense for a worker to quit voluntarily into unemployment (e.g. Clark and Summers (1979)).

Intuitively there are both advantages and disadvantages associated with hiring employed rather than unemployed applicants. It is reasonable to expect that unemployed applicants on average are less productive e.g. because firms generally lay off their least productive workers in bad times (selection effects) or because workers lose skills during unemployment (duration effects). Unemployment thus involves a signal of low productivity. The most important advantages of hiring unemployed searchers are that they should be able to start working on a new job sooner and/or accept a lower initial wage.

Anyone wishing to investigate how employment status affects an applicant's probability of getting a job faces a number of difficulties. First, data is needed about the search behavior and search outcomes for both employed and unemployed job applicants. Second, data about all other relevant characteristics of the searchers are needed to isolate the effect from employment status from other factors that may affect the hiring decision. Problems with unobserved heterogeneity are often encountered in studies of discrimination and are usually very difficult to handle because it is rare for the researcher to have access to, and be able to control for, all the information the firms use when they make their hiring decisions. This makes it hard to identify the effects.

[3] A closely related issue is whether the duration of unemployment affects the exit rate from unemployment and/or if firms prefer to hire searchers with short durations (e.g. Vishwanath (1999), Blanchard and Diamond (1994) and Pissarides (1992)).

This paper uses data from the Applicant Database (Sökandebanken), which is kept by the Swedish Public Employment Office. All workers, both employed and unemployed, looking for a new job are invited to submit details about their education, experience and other skills. Employers can then search in this database for applicants that they find interesting and contact them for interviews etc. Most such contacts are registered. The data covers all applicants remaining as active searchers in April 2001 who agreed to participate in this research project.

This dataset has several advantages. First, we have data about the search activities of a large number of employed workers; almost half of all searchers in the sample search on the job. Second, the search intensity is the same for all searchers in the database; to search in this case just means to submit the required information to the database using standardized forms.[4] Third, since the employers only observe what is in the database, we can be certain that we have records of all the information that the employers use. Obviously, we cannot perfectly mimic the way firms use this information in a regression analysis, but if we include properly defined control variables for all other worker characteristics, we should be able to minimize the problems with unobserved heterogeneity and thus obtain more precise estimates than most other studies.

The major disadvantage of using this dataset is that we do not know which applicant the employer finally decides to hire. The hiring process often involves several stages and we capture only the first.[5] However, for an applicant to get hired, he or she must be contacted by an employer. Therefore, if we find that the *contact* probability is lower for unemployed workers, this is a strong indication that the *hiring* probability is also lower.

Our results show that an unemployed searcher faces a lower contact probability than an employed searcher. For an otherwise identical searcher, being unemployed reduces the contact probability by 3.4 percentage points. The relative effect from this

[4] All forms must be completed to register. However, the searcher is asked to write a short personal letter and the quality of the letter may be interpreted as a measure of search intensity (see note 8 and the section on robustness below).

[5] For example, an employer locates a few candidates in the Applicant Database she finds interesting, decides to contact them for interviews and finally hires one of them. This means that the employer may be: (i) less likely to contact unemployed workers and (ii) given that the firm chooses to interview unemployed workers be less likely to hire them. Thus, total discrimination may be a product of these two components. However, it can also be the case that it does not matter, or even is an advantage, to be unemployed in the latter stages of the hiring process for the reasons discussed above.

reduction for a particular searcher depends on his or her other characteristics. For the 'average' searcher having been in the database for the average time (i.e. 35 weeks), this corresponds to a decrease from 45 to 41 percent; i.e. 9 percent.[6] This may seem like a limited effect, but it should be noted that the relative effect from a 3.4 percentage point drop in the contact probability can be bigger – at least 10 to 20 percent – for a low skilled worker searching for unqualified work. We also show that unemployed workers get fewer contacts. All results are statistically significant at conventional levels and are robust to variations of specifications and estimation methods. Thus, the results support the notion that firms view unemployment per se as a signal of negative unobservable characteristics.

Most of the existing literature studying this issue use surveys or interviews, where employers are asked about their hiring procedures. Examples using Swedish data are Agell and Bennmarker (2002), Agell and Lundborg (2003), Klingvall (1998) and Behrentz and Delander (1996). All these studies find that some firms consider unemployment a negative characteristic. Similar results for other countries can be found in Bewley (1999) for the US and Atkinson, Giles and Meager (1996) for the UK. However, even though these studies support the view that labor market status is used as a hiring criterion, it is difficult to draw conclusions about the quantitative importance of such behavior from them. One might also question whether employers that do use unemployment as a hiring criterion are willing to admit to that in an interview or a survey. Such bias might result in an underestimation of the true extent of discrimination. Given these limitations of survey-based data, it is clearly advantageous to use data on what employers *actually do* rather than what they *claim they do*.

There are also some related quantitative studies.[7] Blau and Robins (1990) use data from the Employment Opportunity Pilot Project, a US data set collected in 1979-80. They find a relatively large effect, employed searchers receive 0.24 offers per contact attempt, while unemployed searchers receive just 0.17. However, these results are

[6] The 'average' searcher is a 26-35 year old Swedish man with secondary education and at least five years labor market experience who has a driving license, good computer skills, good language skills in Swedish and English, that searches for technical work (Amsyk 3) in Stockholm.

[7] Andrews et al (2001) find that employed search is slightly more effective for UK youths aged 15-18. In contrast, Holzer (1987) finds that unemployed search is more effective using a sample of US youths. Kugler and Saint-Paul (2004) find that unemployed searchers have a lower probability of finding work in the US and Spain. See also Burgess (1993). The related issue of duration dependence is explored in van den Berg and van Ours (1996).

obtained without controlling for differences in search intensity or other characteristics. Controlling for some of these factors, the difference is reduced but remain statistically significant. Belzil (1996), using a Canadian data set, finds that young male job searchers face a similar arrival rate of job offers irrespective of their employment status, while prime-age male searchers face a much lower arrival rate if they are unemployed; the arrival rates are 0.046 and 0.072 respectably. However, in none of these studies can the authors claim that they have access to *all* the information that firms use when choosing whom to hire. Thus, it is difficult to know if it is unemployment *per se* or some other information obtained by firms in job interviews etc., unobservable to the researcher, that determines whom the firm gives a job offer and/or hires. That is, the issue of unobserved heterogeneity is again crucial.

The rest of the paper is organized as follows. Section 2 presents the Applicant Database. Section 3 discusses identification issues, defines the variables and presents the empirical specification. Section 4 presents the results and a robustness analysis. Section 5 concludes.

2. DATA

The Applicant Database is kept by the Swedish Public Employment Office since 1997. The searchers submit their details using standardized forms, either from home via the Internet or at the Employment Office, and enter information about education, work experience, language skills, computer skills, preferred occupation and region and write a short personal letter.[8] The information is only made visible to employers if all forms have been completed, so there are no missing values. Employers that are registered with the Employment Office can then use the database to locate workers they find interesting and contact them for interviews etc. Most such contacts are registered.[9]

In this section, we discuss the data used in the empirical investigation. We describe how the data was obtained, give summary statistics and discuss selection issues.

[8] In the personal letter, the applicants are free to write whatever they want. Thus, the quality of the letter may be interpreted as a measure of search intensity. It is difficult to find an appropriate method of controlling for the quality of the letter and we do not try to do it in our baseline regressions. Instead, we consider this issue in the robustness analysis.
[9] It is possible that some employers contact applicants directly, e.g. if a worker includes a phone number in the personal letter. However, according to the Employment Office, most contacts are made within the system.

The Characteristics of the Sample

In the spring of 2001, the Applicant Database contained approximately 50,000 searchers with a monthly inflow of around 11,000. All searchers that logged into the system between March 1 and March 12, 2001 were asked if they wanted to participate in a research study investigating the recruitment behavior of firms. Around 50 percent of those asked agreed, giving us a sample of 8,666 individuals. After having excluded youths below 20, we obtained the sample used in this study consisting of 8,043 individuals.[10] These people have been in the database for an average time of approximately 35 weeks. Table 1 gives some descriptive statistics about the searchers and the jobs they hope to find.

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

	All	Employed	Unemployed
Number of applicants	8,043	3,941	3,056
Employment status:			
Employed	0.49		
Unemployed	0.38		
University student	0.08		
In other training	0.04		
On parental leave	0.01		
Highest level of completed education:			
Primary	0.07	0.05	0.12
Secondary	0.49	0.51	0.53
University	0.44	0.44	0.35
Work experience:			
None	0.15	0.05	0.21
Some (less than 5 years)	0.42	0.45	0.40
Long (five years or more)	0.44	0.50	0.39
Other skills:			
Managerial experience	0.34	0.42	0.27
Telecommuting experience	0.12	0.14	0.11
Research experience	0.05	0.06	0.04
Driving license	0.79	0.83	0.73
Good computer skills	0.74	0.76	0.69
Good language skills – Swedish	0.97	0.98	0.96
Good language skills - English	0.56	0.58	0.50
Good language skills – G-F-S	0.20	0.20	0.18

[10] Most of the applicants aged below 20 look for work during the summer break or other temporary work.

Table 1 (*continued*)

	All	Employed	Unemployed
Age:			
Mean (years)	33.8	34.1	34.0
Age 20-25	0.29	0.25	0.28
Age 26-35	0.33	0.36	0.29
Age 36-50	0.28	0.30	0.29
Age 51-	0.10	0.09	0.14
Gender:			
Female	0.49	0.49	0.45
Ethnicity:			
Foreign name	0.13	0.12	0.15
Region:			
Stockholm	0.29	0.31	0.26
Uppsala	0.09	0.09	0.08
Södermanland	0.08	0.07	0.08
Östergötland	0.08	0.08	0.08
Jönköping	0.06	0.06	0.06
Kronoberg	0.05	0.05	0.04
Kalmar	0.05	0.05	0.05
Gotland	0.02	0.02	0.02
Blekinge	0.05	0.05	0.04
Skåne	0.19	0.19	0.19
Halland	0.08	0.08	0.07
Västra Götaland	0.18	0.20	0.15
Värmland	0.05	0.05	0.05
Örebro	0.07	0.06	0.07
Västmanland	0.07	0.07	0.08
Dalarna	0.05	0.05	0.06
Gävleborg	0.06	0.05	0.06
Västernorrland	0.04	0.04	0.05
Jämtland	0.02	0.02	0.02
Västerbotten	0.04	0.04	0.04
Norrbotten	0.03	0.03	0.03
Occupation:			
Legislators, senior officials and managers (Amsyk 1)	0.03	0.04	0.02
Professionals (Amsyk 2)	0.28	0.30	0.22
Technicians and associate professionals (Amsyk 3)	0.29	0.33	0.25
Clerks (Amsyk 4)	0.25	0.27	0.24
Service workers and shop sales workers (Amsyk 5)	0.19	0.20	0.19
Skilled agricultural and fishery workers (Amsyk 6)	0.02	0.02	0.02
Craft and related trades workers (Amsyk 7)	0.12	0.12	0.12
Plant and machine operators and assemblers (Amsyk 8)	0.10	0.10	0.11
Elementary occupations (Amsyk 9)	0.11	0.09	0.13

Notes: Our measure of labor market experience only includes work in those occupations the worker wants to find a job. This explains why some of those who are employed are classified as having no work experience. G-F-S denotes language skills in German, French or Spanish. The column labeled all includes all searchers including students etc. It is possible for the workers to search for jobs in several regions and/or occupations, and this explains why the fractions do not sum to one.

From Table 1 there are several things worth noting. (1) There are more employed than unemployed people in the sample. (2) The people in the sample tend to be quite young and well educated; the average age is just around 34 years and 44 percent have a university degree. (3) There are almost as many women as men in the database and it includes a non-negligible number of workers with foreign names.[11] (4) A substantial fraction of the searchers seek employment in the areas surrounding the three biggest metropolitan areas and they are quite diversified with respect to the types of work they seek. Turning to the number of offers received, the 8,043 searchers in our sample received 7,179 contacts from employers during their time in the database. Table 2 gives some summary statistics about the fraction receiving at least one offer and the number of offers received, both for all workers and for the four employment status subgroups.

Table 2. Descriptive statistics about the contacts received divided into employment status subgroups

Employment status	Fraction receiving at least one contact	Average number of contacts
All	0.34	0.89
Employed	0.41	1.16
Unemployed	0.28	0.64
University student	0.25	0.57
In other training	0.30	0.75

In Table 2, we see that an employed searcher received, on average, almost twice as many contacts as an unemployed searcher. Obviously, we cannot conclude from these numbers that it is unemployment in itself that leads to this outcome, since the groups differ systematically in a number of other dimensions as well. We also see that people currently participating in some sort of education receive quite few contacts.

Selection Issues

There are two potentially important selection issues that we need to consider.

First, for our results to have internal validity for this particular search channel, we must ask whether the searchers that agreed to participate in our study differ from those that did not, and if this may affect our results. Unfortunately, we do not have any

[11] The Applicant Database does not contain information about ethnicity. However, since employers can see the name of the applicant from the information submitted, we might expect some employers to use this as a basis for discrimination. Therefore, all workers in the Applicant Database agreeing to participate in the study were asked whether they believed other people perceive their name as ‘foreign’.

information about the searchers that did not agree to participate. However, this should only affect our results if we fail to include important variables that are correlated with employment status in our regressions, or if our regressions do not fully capture the potentially very complex way the employers use the information. We investigate this possibility in the section on robustness below, and find, for example, no significant interaction effects between employment status and variables such as gender, age, education, region, occupation etc. 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 the Applicant Database, we might wonder whether those that do use it differ from those that do not. This issue is important if we want to generalize our results to the entire economy, i.e. the external validity. To illustrate whether the searchers in our dataset differ from the typical job searcher in Sweden, Table A1 in the Appendix compares the characteristics of our data with data of typical job searchers from the Swedish Public Employment Office. Comparing these two datasets the most noteworthy differences are that our workers have more education, are a bit younger, are less likely to be immigrants and are more likely to search for technical work. As we will see in the section on robustness below, the effect from being unemployed does not seem to differ in those dimensions, thus indicating that our results may be valid for all searchers. 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 A2 in the Appendix compares the preferred occupation/region of the searchers that have received offers in our dataset with the inflow of vacancies to the Swedish Public Employment Office. We see that the most noteworthy differences are that our vacancies are more likely to be for jobs outside of the Stockholm area, more likely to be for jobs as clerks (amsyk4) and less likely to be for jobs as service workers (amsyk5). However, as we will see later, the effect from being unemployed does not seem to differ across occupations or regions.

3. IDENTIFICATION AND ESTIMATION

Our objective is to investigate whether or not the probability of getting contacted by a firm is affected by the current employment status of the applicant. To identify this effect, we need to control for all other factors that may affect the firm's contact decision. In this section, we discuss identification issues, define the variables and present the econometric specification.

Identification

Suppose that an employer has chosen to use the Applicant Database 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, she will consider two types of information. First, all factors that she believes directly will affect the productivity of the applicants; e.g. education and work experience. Second, all factors (e.g. employment status, gender and ethnicity) that the employer believes are signals for other important factors (e.g. ability to co-operate, motivation and other social skills) that are unobservable to her when she makes her choice. A crucial difference between these two types of information is that observable skills will affect the productivity of the applicant directly, while the signals only will affect the productivity because it indicates bad characteristics in some unobservable dimension.

Now, how can we identify the effect of employment status on the probability of getting contacted for a worker in the database? Here it is important to note that the information of the econometrician coincident exactly 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 orthogonal regressors. 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 employment status, and the residual term is therefore orthogonal.

Variables

The variables used correspond to those presented in Table 1. Here, we will try to give some intuition for how we have chosen to construct them.

First, we have observable factors that are directly related to productivity such as education and labor market experience. To control for education, we include dummy variables for the highest level of completed education; primary, secondary and university. To control for experience, we use dummy variables for three lengths of experience; none, some ($0 < t < 5$ years) and long (≥ 5 years).[12] We also use dummy variables for managerial experience, experience of telecommuting, research experience, driving skills, computer skills and language skills in Swedish, English and German/Spanish/French.

Second, we have factors that employers may use as signals. These include age, gender, ethnicity and employment status. To control for age, we divide the searchers into five groups; 20-25, 26-35, 36-50 and 51- years old.[13] For gender and ethnicity, we use naturally defined dummy variables. For employment status, we divide the searchers into five groups; employed, unemployed, university student, in other training and on parental leave.

Third, we include variables for differences across occupational and regional labor markets. Usually, an employer’s choice will be limited to those searchers that have stated that they are interested in a particular occupation at a particular location.[14] Since it is natural to expect that labor market conditions differ both across occupations and regions, we include controls for occupation, based on the nine-group classification system used by the Employment Office, and for location, based on counties.

Fourth, we need to include controls for the length of time applicants have been in the Applicant Database, since applicants 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 section on robustness below).

[12] Only experience in those occupations that the searcher wants to find a job is included.
[13] An alternative would be to include age as a continuous variable, but this does not matter for our results.
[14] Of course, it is possible for firms to ignore such requirements and contact workers anyway. However, in most cases we would expect such action to be pointless.

Estimation

We estimate a model for the probability that a searcher in the Applicant Database receives at least one contact, during his or her time in the database, as a function of the variables introduced above.[15] We use the Probit model and the empirical specification is given by:

$$P^{Offer} = f(S, X, T, Z, t) + \varepsilon, \qquad (1)$$

where S denotes the current employment status of the applicant, X denotes the observable productive characteristics of the applicant, T denotes signals other than employment status, Z denotes the characteristics of the desired job and t is the time vector.

4. RESULTS

This section discusses the results for the probability of receiving a contact, and the number of contacts received, and investigates the robustness of these results.

The Probability of Receiving a Contact

Table 3 below shows the estimates of equation (1) for the probability of receiving a contact. Column 4 of Table 3 presents the Probit estimates for the full specification. However, to make it easier to interpret the results, we discuss the implied marginal effects. The contact probability is approximately 3.4 percentage points lower for an unemployed applicant than for an otherwise identical employed applicant, and this effect is statistically significant at conventional levels. The relative effect from this reduction for a particular searcher depends on his or her other characteristics. To get a feeling for the size of the effect, we can calculate the contact probability for the 'average' searcher who has been in the database for the average time (35 weeks). For such a searcher the contact probability is reduced by around 9 percent; from 45 percent if he is employed, to 41 percent if he is unemployed.

[15] We focus on the contact probability rather than on the number of contacts received because (i) most other related studies follow this approach, (ii) most of our searchers have received zero or one contact and (iii) the econometrics is more straightforward.

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

	(1)	(2)	(3)	(4)
Employment status (S) (ref. employed):				
Unemployed	-0.215***	-0.155***	-0.133***	-0.097***
	(0.034)	(0.035)	(0.036)	(0.037)
University student	-0.129**	-0.148**	-0.169**	-0.203***
	(0.062)	(0.066)	(0.068)	(0.069)
In other training	-0.100	-0.040	-0.061	-0.030
	(0.081)	(0.083)	(0.085)	(0.086)
On parental leave	0.155	0.141	0.167	0.189
	(0.167)	(0.167)	(0.167)	(0.169)
Observable productive characteristics (X):				
Highest level of completed education (ref. primary):				
Secondary		0.369***	0.306***	0.217***
		(0.075)	(0.075)	(0.078)
University		0.521***	0.461***	0.339***
		(0.076)	(0.078)	(0.082)
Work experience (ref. some):				
None		-0.257***	-0.192***	-0.182***
		(0.054)	(0.056)	(0.056)
Long		0.018	0.011	0.104**
		(0.034)	(0.035)	(0.042)
Other skills:				
Managerial experience				0.173***
				(0.039)
Telecommuting experience				0.077
				(0.051)
Research experience				0.013
				(0.073)
Driving license				0.043
				(0.044)
Good computer skills				0.064
				(0.041)
Good language skills – Swedish				0.056
				(0.098)
Good language skills – English				0.110***
				(0.037)
Good language skills – G-F-S				0.104**
				(0.043)
Other signals (T):				
Age (ref. age 20-25):				
Age 26-35				-0.105**
				(0.044)
Age 36-50				-0.278***
				(0.053)
Age 51-				-0.339***
				(0.071)
Ethnicity:				
Foreign name				-0.058
				(0.049)

Table 3 (*continued*)

	(1)	(2)	(3)	(4)
Gender:				
Female				-0.163***
				(0.037)
Dummies for region	No	No	Yes	Yes
Dummies for occupation	No	No	Yes	Yes
Other variables:				
Weeks in the database	0.032***	0.031***	0.031***	0.032***
	(0.001)	(0.001)	(0.001)	(0.001)
(Weeks in the database)^2	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(0.000006)	(0.000006)	(0.000006)	(0.000007)
Constant	-1.121***	-1.514***	-1.690***	-1.674***
	(0.031)	(0.080)	(0.082)	(0.129)
Number of observations	8,043	8,043	8,043	8,043

Notes: Specification (1) estimated using the Probit model. Robust standard errors are in parentheses. ***, ** and * denote significance at the 1, 5 and 10 percent level. The reference category is an employed man with a Swedish sounding name having primary education, some labor market experience and looking for unskilled work in Stockholm.

This may seem small, but it should be noted that the relative effect from a 3.4 percentage point drop in the contact probability can be bigger – at least 10 to 20 percent – for a low skilled worker searching for unqualified work in regions with depressed labor markets. These results support the proposition that firms view unemployment as a signal of undesirable worker characteristics and, ceteris paribus, prefer to contact an employed rather than an unemployed applicant.

Other results worth noting from column 4 are: (1) Searchers currently enrolled in university education also face a lower contact probability than employed searchers. This might reflect the fact that firms want workers that are available for work directly or some other reason. (2) Education and labor market experience have the expected signs. A higher level of completed education, or more labor market experience, has a clear positive effect. (3) Other applicant characteristics functioning as signals, like age and gender, also have quite strong effects. Women and older workers face a significantly lower contact probability.

In Section 2, we saw that employed workers, on average, have a much higher probability of getting contacted by an employer. A large proportion of this difference reflects systematic differences in observable characteristics between employed and unemployed applicants. To get a feeling for what these differences are, it is illuminating to consider columns 1 to 3 in Table 3, where we start with only labor market status variables as regressors and then successively introduce other variables that contain systematic differences between employed and unemployed applicants (the constant and the time vector are included in all regressions).

In column 1, we only include the employment status variables and see that an unemployed worker faces a 7.5 percentage points lower contact probability. In column 2, we include variables corresponding to such observable productive characteristics that are usually included in studies of discrimination. The probability difference now falls to 5.4 percentage points, implying that some of the difference in search outcome is explained by the fact that the unemployed applicants have less education and less labor market experience. In column 3, we introduce the variables for occupations and regions and see that the probability difference now is 4.6 percentage points. This reflects the fact that unemployed applicants seem to search for the 'wrong' kinds of jobs in the 'wrong' regions.

An interesting question is if the disadvantage unemployed searchers face differs between different subgroups of unemployed applicants. To investigate this, we introduce interaction terms between employment status and other variables. We have tried including a large number of such interaction effects (see the section on robustness below), but three particularly interesting ones are based on gender, age and occupation. (1) We may ask whether women face a bigger disadvantage than men do. The result is that the interaction term between unemployment and gender is statistically insignificant. (2) We may ask whether older workers face a bigger disadvantage than younger workers, as in Belzil (1996). The result is that the interaction terms between unemployment and the age groups are statistically insignificant and thus our results do not support what Belzil found. (3) We may ask whether the negative effect of being unemployed differ across occupations. Including interaction terms between occupational variables and the unemployed variable, we do find that the coefficient estimates are bigger for less skilled occupations, but that all these differences are statistically insignificant. Thus, no particular occupational group seems to drive the results.

To summarize the results so far, we can conclude that unemployed searchers have a lower probability of getting contacted by an employer than employed searchers. Some of this difference is explained by the fact that unemployed searchers have less education and less labor market experience and by differences in the type of job they wish to find. However, even after we control for these variables a non-negligible negative effect remains from being unemployed, thus, indicating that unemployment per se is considered a negative signal.

The Number of Contacts Received

We also have information about the number of contacts the applicants have received during their time in the Applicant Database. This means that we can take the analysis a bit further by asking: Do unemployed workers get fewer contacts as well? This is interesting since a searcher may need several offers before getting hired.

To estimate the effect of employment status on the number of contacts received, we estimate a Poisson model using the same explanatory variables as in equation (1), where the dependent variable now is the number of contacts received. The results of the estimation are presented in Table 4, where we only rapport the results for the employment status variables.[16]

Table 4. Poison estimates of the number of contacts received

Employment status, S (ref. employed):	
Unemployed	-0.127***
	(0.049)
University student	-0.300***
	(0.103)
In other training	0.008
	(0.153)
On parental leave	-0.045
	(0.218)
Numbers of observations	8,043

Notes: Specification (1), with the number of contacts received as the dependent variable, estimated using the Poisson model. Only employment status variables are reported. Robust standard errors are in parentheses. ***, ** and * denote significance at the 1, 5 and 10 percent level. The reference category is an employed man with a Swedish sounding name having primary education, some labor market experience and looking for unskilled work in Stockholm.

The results in Table 4 largely confirm the previous results. An unemployed job seeker receives, on average, 0.11 fewer offers than an otherwise identical employed job seeker.

[16] Similar results are obtained if we instead use the Negative Binomial model.

Robustness

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

An important issue to consider is whether or not we have managed to take into account all differences in characteristics across workers. As stated before, we have access to exactly the same information about the applicants as the firms' that use the Applicant Database, but the possibility remains that we have not succeeded in controlling for all such information. Thus, there may remain systematic differences that we have not taken into account. First, in the baseline regressions we have not tried to control for the quality of the personal letter, since any attempt to grade the letters would be highly arbitrary. However, it is possible that the employers use such information when making their choices, even though this becomes a problem only if unemployed searchers, on average, write lower quality letters than employed searchers. One aspect of the letter that may be correlated with quality is its length. If we include a variable in our baseline specification that measures the number of words in the letter, this variable is insignificant, while all other estimates remain unchanged. This is an indication that the letter may not affect our results that much. Second, there may be important interaction effects between variables that should be included in the regression. To investigate whether this is the case, we have run a large number of regressions including interaction terms between the variables. We have especially focused on including interaction terms between the employment status variables and regressors such as gender, age, education, experience, occupation and region. Very few of these interaction effects are significant and none of them affect the main result that unemployed workers face a lower probability of getting contacted.

Another issue that can 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 that has been in the database longer probably is more likely to have received a contact. This is probably the correct way of controlling for these effects if employers look for workers only infrequently because then they are unlikely to remember the workers they have rejected. If they look very frequently, longer time may not be an advantage at all, but rather a liability since it signals low ability. We have tried a number of alternatives to the baseline regression to see whether this matters, among them the following: (1) Including only offers received in March and April 2001 and then running the regression without the time vector; i.e. all searchers have been in the database the whole period considered. (2) Including higher order polynomials of time in the time vector. (3) Dividing the sample into four groups based on a percentile ranking of the searchers' time in the database and then including dummies for the groups as regressors instead of the time vector. (4) Splitting the sample into those that have been in the database longer (≥ 16 weeks) or shorter (< 16 weeks) and then running separate regressions without the time vector. The results of this sensitivity analysis are presented in Table 5.

Table 5. Alternative specifications for time in the database

	Baseline	1	2	3	4a	4b
Employment status, S (ref. employed):						
Unemployed	-0.097***	-0.083**	-0.114***	-0.125***	-0.118**	-0.160***
	(0.037)	(0.040)	(0.037)	(0.037)	(0.058)	(0.047)
Number of observations	8,043	8,043	8,043	8,043	4,136	3,907

Notes: Specification (1) estimated using the Probit model. Only the unemployment variable is reported. Column 1 includes only offers received in March and April 2001 and no time vector, column 2 includes an extended time vector consisting of t, t^2, t^3 and t^4, column 3 includes group dummies based on a percentile ranking of the time in the database where searchers have been divided into four groups instead of the time vector, column 4a is a regression, excluding the time vector, on the subsample of searchers that have been in the database less than 16 weeks, column 4b is a regression, excluding the time vector, on the subsample of searchers that have been in the database at least 16 weeks. Robust standard errors are in parentheses. ***, ** and * denote significance at the 1, 5 and 10 percent level. The reference category is an employed man with a Swedish sounding name having primary education, some labor market experience and looking for unskilled work in Stockholm.

From Table 5, we see that our main result that unemployed searchers face a disadvantage remains in all alternative specifications. The coefficient estimates for unemployment in columns 1 to 3 are very similar to the result in the baseline specification; thus indicating that the way we control for time is not crucial. From the last two columns it is worth noting that the disadvantage of being unemployed seems to be bigger for those that have been in the base longer than for those that have been in the base shorter, even though this difference is not statistically significant. This is hardly surprising since it is natural to expect that employers are more likely to reject long-term unemployed workers. We have tested this further by running a regression including only unemployed workers and only offers from March and April 2001, where we have divided the workers into two groups, those that have been in the database more than twelve months and those that have been in the database for a shorter time. The result is that we do not find a significant difference between the two groups of unemployed workers. However, it should be remembered that the searchers' time in the database are not necessarily equal to the duration of their unemployment.

In summary, the sensitivity analysis indicates that our main result that being unemployed is a negative characteristic is robust to different modeling assumptions.

5. CONCLUDING REMARKS

Firms hiring new workers are often unable to perfectly observe the applicants' productivity. Instead, they try to infer it by using the information they have available. Such information includes easily observable factors, such as employment status, that firms may believe are correlated with unobservable factors that do affect productivity. If employers use employment status as a hiring criterion, an unemployed job seeker should face a lower contact probability than an otherwise identical employed job seeker. The purpose of this paper has been to empirically investigate whether this theoretical implication is valid.

Using Swedish data from the Applicant Database, we find that an unemployed job seeker faces a lower contact probability, and receives fewer contacts, than an employed job seeker. Some of this difference is due to the fact that the unemployed searchers have less education, less work experience etc. and search for work in regions and occupations with high unemployment, but even after controlling for such differences an effect remains. For an otherwise identical searcher, being unemployed reduces the contact probability by 3.4 percentage points. This indicates that, at least, some firms view employment status as a signal of productivity and therefore, ceteris paribus, prefer to contact employed rather than unemployed applicants. Since we do not know which applicant the employer eventually hires, we cannot know whether or not the effects we find over- or underestimate the effect on the *hiring probability*. A low contact probability means that fewer unemployed workers remain in the latter stages of the hiring process, but unemployed searchers may be more likely to accept offers, especially if the wage offered is low. Thus, further research is needed to investigate the aggregate effects. However, we find a smaller difference between employed and unemployed job applicants than previous studies such as Belzil (1996) and Blau and Robins (1990).

An important issue is whether the results in this study are true only for this particular search channel or if they can be generalized to the whole labor market. Obviously, only further empirical analysis can answer such a question. However, *a priori* it is difficult to think of any particular reason why firms using the Applicant Database should be more prone to view unemployment as a negative worker characteristic than employers using other search channels. On the contrary, it seems more likely that the opposite is true; i.e. that firms that really believe unemployment signals low productivity should use other search channels, such as personal contacts, to recruit workers. Moreover, it is also likely that the best employed searchers receive frequent offers from personal contacts implying that they may not find it worthwhile to use the Applicant Database. If this line of reasoning is true, it may be that we also underestimate the true extent of discrimination based on employment status.

A particular strength of this study is that it allows us to, at least to some extent, overcome the problems associated with unobserved heterogeneity. Much of the existing literature on discrimination, e.g. Blau and Robins (1990), is vulnerable to the criticism that unobserved heterogeneity may explain the results rather than the factors that the authors claim. One way of overcoming these problems is to use audit based methods (e.g. Riach and Rich (2002) and Petit (2004)), but audit studies have also been criticized (e.g. Heckman (1998)) and are difficult to perform on a large scale. The use of new datasets, like the Applicant Database, is an alternative way of overcoming such problems and may allow us to better study discrimination in the future.

REFERENCES

Agell, J and H Bennmarker (2002), “Wage Policy and Endogenous Wage Rigidity: A Representative View from the Inside”, Working Paper 2002:12, Institute for Labour Market Policy Evaluation.

Agell, J and P Lundborg (2003), “Survey Evidence on Wage Rigidity and Unemployment: Sweden in the 1990s”, *Scandinavian Journal of Economics*, Vol. 105, 15-29.

Andrews, M J, S Bradley and R Upward (2001), "Estimating the Probability of a Match Using Microeconomic Data for the Youth Labour Market", *Labour Economics*, Vol. 8, 335-357.

Atkinson, J, L Giles and N Meager (1996), "Employers, Recruitment and the Unemployed", Institute for Employment Studies Report 325.

Bean, C R (1994), "European Unemployment: A Survey", *Journal of Economic Literature*, Vol. 32, 573-619.

Behrenz, L and L Delander (1996), "Arbetsgivarnas Rekryteringsbeteende – En Intervjuundersökning" (Employers' Recruitment Behavior – An Interview Based Study), in SOU 1996:34, *Report to the Government Commission on Labour Market Policy*, Fritzes, Stockholm.

Belzil, C (1996), "Relative Efficiencies and Comparative Advantages in Job Search", *Journal of Labor Economics*, Vol. 14, 154-173.

Bewley, T F (1999), *Why Wages Don't Fall During a Recession*, Harvard University Press, Cambridge.

Blanchard, O J and P Diamond (1994), "Ranking, Unemployment Duration and Wages", *Review of Economic Studies*, Vol. 61, 417-434.

Blau, D M and P K Robins (1990), "Job Search Outcomes for the Employed and Unemployed", *Journal of Political Economy*, Vol. 98, 637-655.

Boeri, T (1999), "Enforcement of Employment Security Regulations, On-the-job Search and Unemployment Duration", *European Economic Review*, Vol. 43, 65-89.

Burgess, S M (1993), "A Model of Competition between Unemployed and Employed Job Searchers: An Application to the Unemployment Outflow Rate in Britain", *Economic Journal*, Vol. 103, 1190-1204.

Clark, K B and L H Summers (1979), "Labor Market Dynamics and Unemployment: A Reconsideration", *Brookings Papers on Economic Activity*, Vol. 1, 13-72.

Edin, P A and J Lagerström (2005), "Blind Dates: Quasi-Experimental Evidence on Discrimination", Mimeo, Uppsala University.

Eriksson, S (2002), "Imperfect Information, Wage Formation, and the Employability of the Unemployed", Working Paper 2002:17, Institute for Labour Market Policy Evaluation.

Eriksson, S and N Gottfries (2005), "Ranking of Job Applicants, On-the-job Search, and Persistent Unemployment", *Labour Economics*, Vol. 12, 407-428.

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

Holzer, H J (1987), "Job Search by Employed and Unemployed Youth", *Industrial and Labor Relations Review*, Vol. 40, 601-611.

Klingvall, M (1998), "Arbetsgivarnas Attityder" (Employers' Attitudes), Ura 1998:9, The National Labour Market Board.

Kugler, A D and G Saint-Paul (2004), "How Do Firing Costs Affect Worker Flows in a World with Adverse Selection?", *Journal of Labor Economics*, Vol. 22, 553-584.

Petit, P (2004), "Hiring Discrimination: A Field Experiment in the French Financial Sector", Working Paper, Univ. Paris 1.

Pissarides, C A (1992), "Loss of Skill during Unemployment and the Persistence of Employment Shocks", *Quarterly Journal of Economics*, Vol. 107, 1371-1391.

Pissarides, C A and J Wadsworth (1994), "On-the-job Search – Some Empirical Evidence from Britain", *European Economic Review*, Vol. 38, 385-401.

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

Tranæs, T (2001), "Raiding Opportunities and Unemployment", *Journal of Labor Economics*, Vol. 19, 773-798.

Van den Berg, G J and J C Van Ours (1996), "Unemployment Dynamics and Duration Dependence", *Journal of Labor Economics*, Vol. 14, 100-125.

Vishwahath, T (1989), "Job Search, Stigma Effect, and the Escape Rate from Unemployment", *Journal of Labor Economics*, Vol. 7, 487-502.

APPENDIX: Comparison of the characteristics of the searchers

Table A1. Comparison of the characteristics of the inflows of searchers to the Applicant Database and the Swedish Public Employment Office (in fractions)

Variable	All The Applicant Database	All Employment Office	Unemployed The Applicant Database	Unemployed Employment Office
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
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
Occupation:				
Amsyk 1	0.02	0.03	0.01	0.04
Amsyk 2	0.21	0.15	0.16	0.17
Amsyk 3	0.19	0.08	0.18	0.07
Amsyk 4	0.17	0.12	0.18	0.11
Amsyk 5	0.15	0.26	0.20	0.25
Amsyk 6	0.02	0.02	0.02	0.02

Table A1 (*continued*)

Variable	All The Applicant Database	All Employment Office	Unemployed The Applicant Database	Unemployed Employment Office
Amsyk 7	0.08	0.10	0.11	0.11
Amsyk 8	0.07	0.12	0.10	0.09
Amsyk 9	0.10	0.13	0.14	0.14

Notes: The data is for the inflow 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 the data from the Employment Office. The preferred regions and occupations sum to more than one in the Applicant Database, since it is possible to search for jobs in several regions and/or occupations

Table A2. Comparison of the searchers in the Applicant Database that have been contacted and the vacancies reported to the Swedish Public Employment Office (in fractions)

Variable	Inflow The Applicant Database	Inflow Employment Office
Region:		
Stockholm	0.20	0.32
Uppsala	0.06	0.03
Södermanland	0.06	0.02
Östergötland	0.05	0.04
Jönköping	0.04	0.03
Kronoberg	0.03	0.02
Kalmar	0.03	0.01
Gotland	0.02	0.01
Blekinge	0.03	0.01
Skåne	0.07	0.10
Halland	0.05	0.02
Västra Götaland	0.11	0.19
Värmland	0.03	0.02
Örebro	0.04	0.02
Västmanland	0.05	0.03
Dalarna	0.03	0.02
Gävleborg	0.03	0.02
Västernorrland	0.02	0.02
Jämtland	0.02	0.01
Västerbotten	0.02	0.03
Norrbotten	0.01	0.03
Occupation:		
Amsyk 1	0.03	0.02
Amsyk 2	0.19	0.14
Amsyk 3	0.20	0.18
Amsyk 4	0.17	0.09
Amsyk 5	0.16	0.32
Amsyk 6	0.02	0.02
Amsyk 7	0.09	0.05
Amsyk 8	0.07	0.09
Amsyk 9	0.07	0.09

Notes: The data from the Applicant Database is for those searchers that have been contacted in March 2001. The data from the Employment Office is for the inflow of vacancies in March 2001.

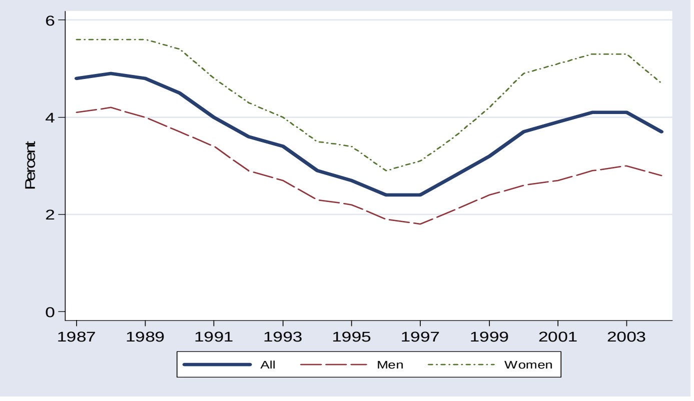
ESSAY II

ECONOMIC INCENTIVES, WORKING ENVIRONMENT AND SICKNESS ABSENCE: EVIDENCE FROM A RANDOMIZED EXPERIMENT*

1. INTRODUCTION

One of the most striking patterns of the Swedish economy may be the high levels of sickness absence. Figure 1, based on data from recent years' labor force surveys, shows that about 4 percent of the employees in Sweden report sickness absence exceeding one week. This is dramatically higher than for most other European countries (e.g. Palmer, 2004).[17] Furthermore, the absence rate varies substantially over time and across seasons, regions, cohorts and types of workers.[18]

Figure 1. Sickness Absence rates in Sweden, 1987-2004



* I am grateful for comments from Peter Skogman Thoursie, Per-Anders Edin, Peter Fredriksson, Katarina Steen Carlsson, Patrik Hesselius, as well as seminar participants at Uppsala University, and the NHESG-conference, August 2005. Thanks to Vello Uibopuu at TietoEnator for providing the data.

[17] Norway and the Netherlands have sickness absence rates similar to Sweden.

[18] For example, from its all time high levels at almost 10 percent in the late 1980's, the total days of absence as a fraction of total working days dropped to 3.8 percent in 1997. However, absenteeism quickly returned to record levels, primary due to a sharp increase in long-term sickness absence. Since 1997, women's yearly sick days have increased from 14 to 27 whereas men's have increased from nine to 15.

The booming economic and social costs due to sickness absence have put the question of how to lower the levels to the political front.[19] Some ideas in the debate focus on the sickness insurance system, arguing that a lack of economic incentives is the key explanation. Another urges employers to improve the working environment for their employees whereas others blame the phenomenon on changes in attitudes towards work.[20]

The purpose of this paper is to estimate how economic incentives and working environment affect sickness absence. To empirically investigate the relevance of these factors is hard for several reasons. Many decision-makers are involved – e.g. the employee, the employer, the medical doctor, the sickness insurance agency – which makes it hard to identify the separate causal effects using observational data. The key challenge is therefore to find plausibly exogenous variation. For example, Larsson (2006) investigates the effect of differences in ceiling levels between the unemployment and the sickness insurance systems and finds that unemployed individuals exploit the possibility of receiving higher benefits by reporting sick. Johansson and Palme (1996) show that the direct cost of being absent has a negative effect on sickness absence. For men, a reduction in the replacement ratio by one percent would decrease the mean level of days absent by about 4.6 percent. In general, however, many studies are open to the criticism that there are systematic differences between treated and non-treated individuals that are not accounted for.[21]

A common method to mitigate these problems is to exploit data from reforms. For example, Henrekson and Persson (2004) conclude that reforms making the Swedish sickness insurance more (less) generous have caused an increase (decrease) in absence.[22] However, reforms are often general, i.e. proper control groups may not be available, and may have been introduced due to high absence levels (i.e. reverse causality). There might also be confounding effects from other changes affecting

[19] The total cost to the state in 2003 was about SEK 110 billion, an increase of almost 50 percent in four years. The government therefore introduced the ambition to reduce the number of sick days between 2002 and 2008 by 50 percent.

[20] For theoretical models on sickness absence, see e.g. Allen (1981) and Shapiro and Stiglitz (1984).

[21] Concerning the importance of job environment, Arai and Skogman Thoursie (2004) find that differences among workers and differences among firms are equally important determinants of Swedish sickness absence. Ichino and Maggi (2004) find a strong environment effect on absenteeism within an Italian bank.

[22] The Swedish sickness insurance system has experienced numerous reforms. Reforms that decreased its generosity in terms of replacement rates and waiting periods before compensation is paid out were implemented in 1992 and 1995. Increases in its generosity occurred in 1963, 1967, 1974, 1987 and 1998.

absence behavior. Controlling for all these factors is hard, raising doubts about a causal interpretation.

Data from randomized experiments have the potential to overcome these difficulties. There are, however, very few studies in social insurance economics based on such strategies (e.g. Krueger and Meyer, 2002). Hesselius *et al*. (2005) provide a rare exception. Using Swedish data from a social experiment, they find strong effects on absence duration when postponing the requirement for a doctor's certificate from day eight to day fifteen. However, most Swedish studies suffer from a lack of new data. From the 1950s until 1991, all sick leaves regardless of length were covered by public sickness insurance. From 1992, the employer became responsible for paying the first 14 days of a sickness absence spell. As a consequence, after 1991 no data on the overall sickness absence rate in Sweden is available.

In this paper, I exploit variation created by an experiment conducted 2001-2002 where individuals were randomly assigned into treatment and control groups. Some individuals faced increased economic incentives during a period of 3 to 12 months. Each month employees received half a day off if they completed the full calendar month without being sick-absent. Other individuals received increased means to improve their working conditions (e.g. work-out opportunities and on-the-job courses in stress reduction). The random allocation into treatment makes it possible, a priori, to estimate the causal effects of the programs by simply comparing absence trends between treated and non-treated individuals. As a robustness check, I also implement a more rigorous empirical strategy using individual panel data from matched income and employment registers.

The results show large and significant effects of the bonus experiment. The bonus reduced the monthly incidence of sickness absence from 22 to 17 percent and the average number of monthly sick days plunged from 2.3 to 1.3. Mainly, the results are driven by the strong response of women and highly educated employees. The results concerning the working environment program are less conclusive; overall absenteeism was reduced by 0.3 days but the effect is not statistically significant at conventional levels. For men, however, there is a relatively large and significant effect.

The rest of the paper is structured as follows. Section 2 takes a closer look at the assignment mechanism, which is crucial for the choice of evaluation strategy. In Section

3, I describe the data collection process and present descriptive statistics. Section 4 contains the empirical analysis with results based on difference-in-differences models and individual fixed effects models. In Section 5, I perform robustness checks and discuss the representativity and the interpretation of the results. Section 6 concludes.

2. IDENTIFICATION STRATEGY

The key problem in all estimation methods is the creation of an accurate control group. If individuals differ in their probability to receive “treatment” for reasons that also affect the outcome variable or factors affecting the outcome variable, we have to control for these differences in order not to attribute them as a treatment effect. With non-experimental data, we need to address two types of potential bias: those due to differences in observables and those due to differences in unobservables (i.e. the selection bias). Several studies have indicated that selection bias can be substantial relative to the actual effect of the treatment (Bell *et al.*, 1995; LaLonde, 1986).[23]

Randomized experiments are, in theory, free from these bias problems and therefore often considered the most robust of the evaluation methods (e.g. Krueger, 2000). Individuals are randomly placed into two groups: one that receive the intervention and one that does not. By randomly assigning treatment, we end up with statistically identical treatment and control groups, given large sample sizes. The groups are therefore likely to respond similarly to underlying trends or shocks.

This paper exploits data from a small-scale experiment in Stockholm. First, I will estimate the treatment effect by simply comparing mean outcomes between treated and non-treated individuals. The availability of a control group is a particular strength of this study, since reforms are often general and control groups therefore not available. However, the experiment is of limited size with respect to establishments. There might therefore be systematic differences in characteristics between treated and non-treated individuals, making it risky to interpret the raw differences in outcomes between the groups as a treatment effect. Second, I will therefore extend the analysis into a multivariate regression, controlling for as many observable differences in characteristics

[23] The intuitive methodology in these studies is to use experimental data and compare the presumably unbiased estimates obtained from these data with estimates obtained by matching treatment groups with non-experimental comparison groups.

as possible. Third, exploiting the fact that we follow individuals over time, I will estimate fixed effect models.

The Experiment

In 2000, a local newspaper published an article about the record levels of sick absence in the municipality of Stockholm, Sweden. Following that report, the municipality hired a program operator with the explicit mission to survey absenteeism and to experiment with methods to reduce it.

The operator randomly sorted the establishments into treatment and control groups. She then contacted the managers at these establishments and offered participation, which everyone accepted. The randomization was conducted within each of the four major sectors in the municipality: child-care, schools, geriatric care and care for the disabled.[24] The main advantage with “site” randomization is that it minimizes the risk that non-treated individuals are affected by the treatment through interactions with individuals in treatment. However, site randomization often results in too few sites for effective randomization to occur. Although the sites have been assigned to treatment and control status by chance, the individuals within the sites have not been randomized. If individuals select into sites, there are differences between sites subject to different treatments.[25]

The experimental design consisted of two separate treatments. In the first, permanent employees at two randomly selected establishments could each month receive half a day off by completing the full calendar month without being sick-absent. Given the monthly average of 15.5 full days of work, the bonus corresponded to a 3.2 percent potential reduction in working time. The possibility to earn the bonus was renewed each calendar month, irrespective of past absence or bonuses. The bonus could not be translated into extra pay but had to be used during the following month.

Prior to the introduction, all employees were informed about the treatment, e.g. the finite time horizon of the experiment. However, in practice, the cancellation of the bonus program was partly sudden and unexpected. Non-permanent workers (i.e. not

[24] The primary reason why the operator conducted establishment randomization instead of randomizing individuals was to avoid the inconvenience of administrating different treatments within the same establishment. In order to get different kinds of establishment, the sampling was conducted within sectors.
[25] Another potential pitfall with site randomization is the fact that non-compliance with the experimental protocol may be higher, since moving out of a site implies dropping out of treatment.

entitled to the treatment) complained to their local labor union. To avoid further complaints among the employees, the program operator aborted the program at this particular establishment. The treatment took place during the period November 2001 to December 2002 but the periods were not identical at the different establishments.

The other program provided four establishments with increased means to improve working conditions. During 3-12 months spells, the employees received on-the-job courses in stress reductions, extended work-out opportunities, discussions on job environment problems and potential solutions.

The Difference-in-differences Approach

The difference-in-differences method compares a treatment and a comparison group (first difference) before and after the intervention (second difference). Hence, I compare the change in absence rates in the treatment group following the introduction of the new policies to the change in absence rates in the control group over the same period. I model individual i's sickness absence in time t as a function of establishment specific effects (δ_e), time specific effects (δ_t) and an interaction term D with value 1 for treated establishments in treatment periods:

$$S_{it} = \delta_e + \delta_t + \alpha \cdot D_{it} + e_{it}, \qquad (1)$$

For the difference-in-differences estimator to produce an unbiased estimate of the program impact we need a common trend between treated and non-treated individuals in the absence of treatment (see e.g. Bertrand and Mullainathan, 2004). Different trends in unobserved variables affecting absence would incorrectly be attributed to the program. The main theoretical advantage with experimental data is that since individuals are randomly assigned into treatment, treatment status is not associated with variables affecting absenteeism independently of the effect through treatment. However, due to the small scale of the experiment there may still be differences, on average, between treated and non-treated individuals. I have two strategies to overcome the potential problem of trends in unobservable characteristics. First, I include in (1) a vector of observable characteristics (X_{it}). Conditioning on these variables improves the credibility of causal inferences. Second, I run individual fixed effects models to look at how an individual responds to the introduction and cancellation of treatment.

3. DATA

The data used in this study consists of all individuals employed at least one day between January 2000 and December 2004 in the part of the municipality of Stockholm where the experiment was conducted. Data on sickness absence is collected from the municipality's internal transaction register, where all absence is recorded on a day-to-day basis. For each day, there is detailed information about the absence, e.g. the employee's stated cause of the absence; how large share of the day the employee was absent; whether or not the day was a work-day (weekend, day off or vacation).[26]

Since the bonus scheme focus on absence per calendar month, I construct the basic outcome variable S, measuring the total number of full workdays that an employee i was absent due to sickness during calendar month t:

$$S_{it} = \text{Sick days} * \overline{\text{Extent}},$$

where 'Sick days' is the total number of days during calendar month t where the employee was (partly) absent due to sickness and 'Extent' is the average share of the day the employee was absent.

By matching this panel to the municipality's employment registers, I include time varying factors (e.g. age, education, profession, position, establishment, establishment sector, monthly wage, contracted hours of work, and an indicator for permanent worker) and a number of time invariant factors (e.g. gender).[27]

There are several advantages with this dataset compared to datasets in related studies. First, data stem from an experiment where we, a priori, expect no systematic differences between the individuals receiving treatment and those who do not. Second, there is a large variation in the key variables. In general, micro data studies are often hampered by a lack of substantial variation. Third, contrary to the Swedish official

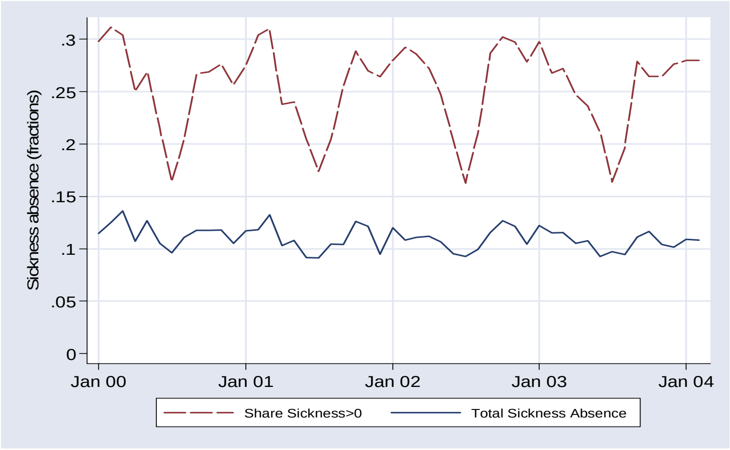
[26] The claimed reasons include, for example, absence due to sickness, rehabilitation, industrial injury, part-time retirement, and disability retirement. In the analysis, I exclude employees on disability pensions.
[27] In the employment registers, there is information about the exact date when any of these variables change. I create my panel by letting a variable take the value it has in the end of the first half of the month.

registers, this dataset includes short-term sickness absence. Forth, with individual panel data we can observe the employees before, during and after the treatments. Intuitively, we expect the incidence of sick absence for treated individuals to fall following the introduction and increase following the cancellation whereas we expect no such pattern in the control group.

Descriptive Statistics

The full sample consists of 3,060 individuals. Figure 2 presents time series data for sickness absence in the sample from 2000 to 2004.

Figure 2. Absence rates, 2000-2004



We see that total sickness absence (i.e. the prevalence) and the share of individuals reported sick at least once during a calendar month (i.e. the incidence) hover around 13 percent and 23 percent, respectively. These numbers are slightly larger than the numbers for the entire Swedish labor market. We also see that the monthly seasonality is very strong, with peaks during December-March. There is no clear sign of any trends over the period.

Table 1 presents some pre-program descriptive statistics about the employees in the sample. The means are presented within each of the five sectors and for treatment groups (T) and control groups (C) separately.

Table 1. Descriptive statistics

	Child-care		Schools		Geriatric		Disability		Other
	T	C	T	C	T	C	T	C	C
# individuals	24	410	23	701	207	1,052	75	357	211
# establishments	1	21	1	28	2	11	2	9	19
Bonus treatment	0	0	0	0	1	0	1	0	0
Sickness absence:									
Prevalence	2.76	2.17	1.96	1.80	1.90	2.31	1.57	1.82	2.75
Incidence	0.38	0.28	0.18	0.25	0.27	0.29	0.21	0.24	0.30
Gender:									
Female	0.95	0.88	0.62	0.76	0.85	0.84	0.72	0.78	0.73
Age:									
Mean (years)	43.1	43.0	42.0	41.8	41.3	42.0	37.9	39.1	46.6
Age 20-35	0.29	0.26	0.22	0.32	0.33	0.31	0.44	0.38	0.11
Age 36-50	0.38	0.45	0.54	0.42	0.47	0.43	0.42	0.46	0.49
Age 51-	0.33	0.29	0.24	0.25	0.20	0.26	0.14	0.15	0.39
Education:									
Primary	0.29	0.36	0.47	0.46	0.67	0.69	0.84	0.76	0.68
Secondary	0.25	0.28	0.35	0.38	0.30	0.25	0.15	0.18	0.13
University	0.46	0.36	0.18	0.16	0.03	0.06	0.01	0.06	0.21
Contract form:									
Wage (t. SEK)	16.3	17.3	18.0	18.0	16.3	16.6	14.9	15.6	20.3
Permanent worker	0.68	0.80	0.78	0.79	0.80	0.83	0.71	0.80	0.88
Weekly hours	36.1	37.3	29.5	18.9	34.6	34.1	33.4	34.2	38.3

Note: The data is for the individuals employed at the municipality in January 2001, when the program operator was hired and started to plan the experiment.

From Table 1 there are several things worth noting. (1) In total 329 - about 11 percent - of the municipality's employees were assigned to treatment and 2,731 were not. The bonus experiment was conducted at one geriatric unit and at one unit for care for the disabled. (2) Compared to the typical Swedish worker, our sample consists to a larger extent of women and elderly workers. This issue is important if we want to generalize our results to the entire economy, i.e. the external validity. (3) The individuals assigned to the treatment group appear to be rather similar to the individuals assigned to the control group. In some cases there are differences though. For instance, the mean wage seems to be slightly lower in the treatment group.

There are several potential reasons for these discrepancies. Most likely, it is due to the relatively small sample sizes within some of the sectors. Establishments might be so different that randomization is unlikely to occur when two establishments are selected.

Even with a large number of establishments a few extreme ones with many workers can tilt the distribution. There may also be incentives for the program administrator to exclude poor individuals or workplaces to improve the outcome of the experiment (see Heckman *et al.*, 1999).

As a more formal test of randomization, I therefore run a regression explaining treatment status with mean characteristics of the establishments (e.g. absence prevalence and incidence, age distribution, gender composition, education level and education field, contract forms). Out of 26 explaining variables, two variables (wage and the share with secondary education) turn up statistically significant at 10 percent level. Since these variables are often correlated with sickness absenteeism it is important to control for these differences in a regression framework.

4. EMPIRICAL RESULTS

Interviews with the program operator as well as managers at the treated establishments indicate that the programs had a positive effect. Overall, the program is described as a success. According to one manager, the major difference was that sickness absence during Mondays and Fridays fell sharply after the introduction of the bonus.

A basic comparison of mean sickness absence between treated and non-treated individuals confirms the conclusions from the interviews. On average, the total number of monthly sick days fell by approximately 1.1 day during the treatment period for individuals subject to the bonus treatment compared to the non-treated individuals. However, there is no clear sign of any effect of the working environment program.

Regression Analysis

This section presents a more rigorous analysis using the difference-in-difference method (conditioning on observables) and the fixed effects model to control for potential differences in observed as well as unobserved factors that may plague small-scale experiments. I use ordinary least squares and the empirical specification is given by:

$$S_{it} = \delta_e + \delta_t + \alpha \cdot D_{it} + X_{it}^{'}\beta + e_{it}, \qquad (2)$$

where S is the individual's total number of sick days during a calendar month, and X is a vector of covariates including for example education, profession, position, gender, age, and wage. Since randomization was conducted within establishment sectors, X also includes controls for sector. Fixed effects for both establishments and months (δ_e and δ_t, respectively) are included. D is an interaction between time and establishments with the value one for treated individuals during treatment. Hence, the parameter of interest is α, measuring the change in monthly sickness absence that individuals experienced during treatment.[28]

The Bonus Program

Table 2 (below) shows the estimates of equation (2) for the sickness absence prevalence in the bonus experiment. In order to check for robustness, the explanatory variables vary by specification. The results confirm the conclusions from the interviews and from the simple comparison of means.

During the treatment period, individuals at the treated establishments had approximately 1.05 day less sickness absence during a calendar month than individuals in the control group. Given the average monthly absence of approximately 2.2 days, this estimated effect corresponds to a 47 percent reduction in total sickness absence. The estimates seem robust to the inclusion of additional control variables and are highly significant irrespective of specification.[29]

Several other results from Table 2 are worth noting: (1) Contract form seems to matter. Absence is significantly higher among permanent employees, consistent with previous research (e.g. Barmby *et al.*, 2004). However, studies based on cross-sectional data are open to the criticism of omitted variables. For instance, permanent workers may have a better health status than temporary workers. Column (d) controls for chronic health status by including individual fixed effects. (2) Females have significantly higher absence rates. (3) Working at an establishment with many employees is associated with significantly higher absence rates.

[28] The standard errors in this paper are clustered with respect to the establishments, since observations can not be considered independent. The results are similar when clustering is conducted on the individuals.
[29] Using the Poisson or the negative binomial model, the estimates are somewhat smaller but still highly significant.

Table 2. Ordinary least squares estimates of monthly sick days

	(a)	(b)	(c)	(d)
Bonus Treatment	-1.083***	-1.030***	-1.053***	-1.298***
	(0.327)	(0.352)	(0.336)	(0.327)
Time dummies	Yes	Yes	Yes	Yes
Establishment dummies	Yes	Yes	Yes	Yes
Field of education dummies	No	Yes	Yes	Yes
Profession and position	No	No	Yes	Yes
Establishment sector	No	No	Yes	Yes
Individual dummies	No	No	No	Yes
Observable individual characteristics (X):				
Gender:				
Female		0.638***	0.807***	
		(0.211)	(0.197)	
Highest level of completed education (ref. primary):				
Secondary		0.097	0.030	
		(0.211)	(0.338)	
University		-0.980***	0.051	
		(0.263)	(0.327)	
Age		0.151*	0.114	
		(0.081)	(0.079)	
Age^2		-0.001	-0.001	
		(0.001)	(0.001)	
Work related factors:				
Wage (t. kr.)			-0.213**	0.095**
			(0.108)	(0.041)
Wage^2			0.001	-0.027***
			(0.002)	(0.001)
Contracted weekly hours of work			-0.005	0.014*
			(0.012)	(0.008)
Permanent worker			0.849***	0.493***
			(0.159)	(0.106)
Establishment size			0.018**	0.038**
			(0.009)	(0.009)
# observations	81,094	81,094	81,094	81,094
R-squared	0.0189	0.0318	0.0651	0.0126

Note: Specification (2) estimated using the ordinary least square model. The treatment effect hence corresponds to α. The dependent variable is the total number of monthly sick days. The treatment consists of a bonus. Column (d) estimated using the fixed effects model. Robust standard errors in parentheses. A constant term was included but not reported. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Another outcome variable is the sick-leave incidence, i.e. the probability of being absent from work due to sickness at all during the month. This is interesting since an employee looses the possibility to earn the bonus in case of *any* sickness absence during the calendar month. To estimate the effect of treatment on the probability of being sick-absent, I estimate models using the same set of explanatory variables as in equation (2) but where the dependent variable now is a dummy variable with the value 1 if the individual has any sickness absence during the month, 0 otherwise. The results of the estimation are presented in Table 3.

Table 3. Ordinary least squares estimates of absence incidence

	(a)	(b)	(c)	(d)
Bonus Treatment	-0.063 (0.063)	-0.058 (0.066)	-0.060 (0.065)	-0.075 (0.050)
Time dummies	Yes	Yes	Yes	Yes
Establishment dummies	Yes	Yes	Yes	Yes
Field of education dummies	No	Yes	Yes	Yes
Profession and position	No	No	Yes	Yes
Establishment sector	No	No	Yes	Yes
Individual dummies	No	No	No	Yes
# observations	81,094	81,094	81,094	81,094
R-squared	0.041	0.052	0.076	0.0128

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 as well as a constant term are included but not reported. The treatment effect hence corresponds to α. The dependent variable is an indicator with value 1 if the individual had *any* sickness absence during the calendar month. The treatment consists of a bonus. Column (d) estimated using the fixed effects model. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Although the effect is not statistically significant, the point estimates show that the probability of having any sickness absence during a month is approximately 7.0 percentage points lower for a treated individual than for an otherwise identical individual. Since the average probability of having a positive amount of sickness absence during a month is about 23 percent, this corresponds to a reduction in the probability by around 30 percent.

To summarize, the bonus treatment caused significant reductions in prevalence - with 45-55 percent depending on econometrical model - whereas the incidence fell by 30 percent. These results suggest that there might have been a reduction in duration as well, a 20 percent reduction from 9.5 days to 7.5 days. A priori, there is one feature in the incentives facing the individuals that would cause duration to decrease. The bonus increases the economic incentives to go back to work earlier at the end of a calendar month since that will give the opportunity to earn the bonus in the upcoming month.[30]

The Working Environment Program

The other part of the experiment consisted of a program to improve the working environment. Table 4 presents the estimated effects of the environment treatment on the number of monthly sick days using the same set of explanatory variables as in Table 2. Only the result for the treatment variable is reported.

Table 4. Ordinary least squares estimates of monthly sick days

	(a)	(b)	(c)	(d)
Environment Treatment	-0.255 (0.674)	-0.283 (0.664)	-0.341 (0.684)	-0.128 (0.287)
Time dummies	Yes	Yes	Yes	Yes
Establishment dummies	Yes	Yes	Yes	Yes
Field of education dummies	No	Yes	Yes	Yes
Profession and position	No	No	Yes	Yes
Establishment sector	No	No	Yes	Yes
Individual dummies	No	No	No	Yes
# observations	85,003	85,003	85,003	85,003
R-squared	0.0199	0.0322	0.0646	0.0128

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 as well as a constant term are included but not reported. The treatment effect hence corresponds to α. The dependent variable is the total number of monthly sick days. The treatment consists of increased means to job environment. Column (d) estimated using the fixed effects model. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Overall, the job environment treatment reduced sickness absence by approximately 0.3 days, corresponding to a 14 percent reduction in average sickness absence (i.e. from 2.2 to 1.9 days). However, the effect is not significant on conventional levels.

The estimated effect on the incidence is shown in Table 5. The probability of being sick absent during a calendar month decreased somewhat, with magnitudes ranging from 0.1 to 1.7 percentage points. This corresponds to a reduction in duration from approximately 9.5 days to 9.0 days. Even though these results are not statistically

[30] Another potential explanation could be Hawthorne effects (i.e. that people respond to the experiment itself rather than the treatment), a topic further discussed in section 5.

significant on conventional levels, they indicate that the main effect of the environment treatment was on spell duration.

Table 5. Ordinary least squares estimates of absence incidence

	(a)	(b)	(c)	(d)
Environment Treatment	-0.009 (0.053)	-0.011 (0.051)	-0.017 (0.049)	-0.001 (0.025)
Time dummies	Yes	Yes	Yes	Yes
Establishment dummies	Yes	Yes	Yes	Yes
Field of education dummies	No	Yes	Yes	Yes
Profession and position	No	No	Yes	Yes
Establishment sector	No	No	Yes	Yes
Individual dummies	No	No	No	Yes
# observations	85,003	85,003	85,003	85,003
R-squared	0.0397	0.0503	0.0751	0.0128

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 as well as a constant term are included but not reported. The treatment effect hence corresponds to α. The dependent variable is an indicator with value 1 if the individual had *any* sickness absence during the calendar month. The treatment consists of increased means to job environment. Column (d) estimated using the fixed effects model. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Heterogeneous treatment effects

An interesting policy aspect is whether changes in economic incentives or job environment have different effects on different subgroups. In the literature, some theoretical models predict differences in response due to for example human capital (e.g. Grossman, 1972, 2000) and household production patterns (e.g. Bolin *et al*., 2002). Tables 6-7 present estimates of heterogeneous treatment effects for the bonus experiment and the job environment experiment, respectively.

From Table 6 (below), we see that there are some heterogeneous effects of economic incentives on sickness absence. Column (i) shows that the effect of the bonus was larger for women than for men. From the last two columns, there is worth noting that: (1) there were no differences across age groups, although the estimate is not significantly different from zero for individuals older than 50, and (2) the point estimates were higher for employees with higher education (i.e. defined as having at least secondary education) than for employees with only lower education.

Table 6. Heterogeneity in Response to Bonus Experiment

	(i)		(ii)		(iii)	
	Men	Women	Age>50	Age≤50	Low educ.	High educ.
Bonus Treatment	-0.436**	-1.135***	-1.057	-1.237***	-0.322	-1.464***
	(0.186)	(0.385)	(0.885)	(0.229)	(0.647)	(0.313)
# observations	12,668	75,237	32,281	55,624	27,475	60,430
R-squared	0.1163	0.0681	0,0892	0,0765	0.1033	0.0688

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 column (c) as well as a constant term are included but not reported. The dependent variable is the total number of monthly sick days. The treatment consists of a bonus. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7. Heterogeneity in Response to Working Environment Experiment

	(i)		(ii)		(iii)	
	Men	Women	Age>50	Age≤50	Low educ.	High educ.
Environment Treatment	-0.864**	-0.307	-0.570	-0.213	-0.880	-0.014
	(0.351)	(0.751)	(0.859)	(0.701)	(0.975)	(0.681)
# observations	12,668	75,237	32,281	55,624	27,475	60,430
R-squared	0.1163	0.0681	0,0892	0,0765	0.1033	0.0688

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 column (c) as well as a constant term are included but not reported. The dependent variable is the total number of monthly sick days. The treatment consists of increased means to job environment. In column (iii) the average monthly sickness absence for the two groups are 3.3 and 0.9, respectively. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7 shows the estimated effect of the working environment treatment across subgroups. There was a significant and fairly strong response by male employees, although the estimated effects across gender are not significantly different. Furthermore, there were no heterogeneous effects across age groups and across education levels.

5. ROBUSTNESS

Even though social experiments are often considered the most robust of evaluation technique, there are a number of pitfalls, see Heckman *et al*. (1999). This section includes more thorough checks in order to enhance the credibility of a causal interpretation of the results and tests of the main assumption behind the difference-in-difference estimator.

Hawthorne Effects

The estimated effects capture the "full" effect of the treatments, i.e. they are reduced-form estimates and hence do not correspond to any "structural" parameters. However, the design of the experiment makes it possible to take the analysis a step forward by asking: By which channels did the treatment affect absenteeism?

One potential causal pathway is the Hawthorne effect. The term refers to a response attributable to the act of experimentation itself rather than the treatment per se. That is, employees may have reduced their absence simply because they were under study. Tables 8 and 9 present some additional heterogeneity in the response to the bonus and environment treatment, respectively, that helps us to interpret the estimated treatment effects.

Table 8. Heterogeneity in Response to Bonus Experiment

	(i)		(ii)		(iii)	
	Permanent workers	Non-permanent workers	High prior sickness absence	Low prior sickness absence	Hours per week <30	Hours per week ≥30
Bonus Treatment	-1.108***	0.137	-1.183**	-0.458***	-2.220***	-0.668
	(0.357)	(0.464)	(0.563)	(0.237)	(0.282)	(0.546)
# observations	80,476	7,429	11,915	67,563	17,817	70,088
R-squared	0.0660	0.0770	0.0985	0.0421	0,1002	0,0702

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 column (c) as well as a constant term are included but not reported. The dependent variable is the total number of monthly sick days. The treatment consists of a bonus. In column (ii) the average monthly sickness absence for the two groups are 3.3 and 0.9, respectively. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Some patterns in the data suggest that Hawthorne effects are not the primary explanation of our results. First, the bonus led to significantly sharper reductions in absenteeism than the environment treatment. Hence, it is unlikely that experimentation *per se* explains the results. Second, groups that we, a priori, expect to respond stronger to the bonus did in fact do so. (1) Only permanent workers were entitled to the bonus. Column (i) in Table 8 shows that we find no effect among non-permanent employees. Hence, people seem to respond to the economic incentive per se and not to the experimentation itself. (2) Column (ii) shows that employees with lower sickness absence prior to the introduction of the bonus experiment responded relatively stronger.

Individuals with an average of one day of sickness absenteeism per calendar month before to the experiment, reduced there sickness absenteeism with approximately 50 percent during the bonus experiment. For individuals with an average prior sickness absenteeism of 3 days, the reduction was only approximately 32 percent. Interestingly, there is also a large effect for individuals with very high prior absenteeism. Employees with an average of 9 monthly sick days prior to the experiment (a subgroup consisting of 8,749 observations) responded to the bonus experiment by reducing their absence by approximately 70 percent.[31]

(3) Part-time workers responded significantly stronger than full-time workers, see column (iii). This may reflect the fact that the incentives from the bonus are stronger for individuals with a smaller amount of working hours per month. Since these individuals spend less time at work, they may have a higher probability to complete the whole month without being sick absent.

Table 9 presents the same robustness checks from the environment treatment.

Table 9. Heterogeneity in Response to Working Environment Experiment

	(i)		(ii)		(iii)	
	Permanent workers	Non-permanent workers	High prior sickness absence	Low prior sickness absence	Hours per week <30	Hours per week ≥30
Environment Treatment	-0.269 (0.768)	-0.530 (0.312)	-4.827 (3.230)	-0.069 (0.535)	-0.321 (1.870)	-0.378 (0.330)
# observations	80,457	7,429	8,743	67,555	17,815	70,071
R-squared	0.066	0.0775	0.3026	0.0420	0.0997	0.0702

Note: Specification (2) estimated using the ordinary least square model. Control variables as in Table 2 column (c) as well as a constant term are included but not reported. The dependent variable is the total number of monthly sick days. The treatment consists of increased means to job environment. In column (iii) the average monthly sickness absence for the two groups are 3.3 and 0.9, respectively. Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

From Table 9, we find no significant differences in the way that these groups responded to the environment treatment.

[31] Some previous Swedish research indicates that individuals on long-term sickness absence may respond strongly to treatment. For example, an internal study at the Insurance Office in Skåne showed that treatment in the form of a phone call to long-term absent employees caused significant reductions in absenteeism.

Falsification Tests

Another question concerns the validity of the common trend assumption in the difference-in-differences estimator. I use three different methods to test this assumption. First, Figure 3 shows the trends of absenteeism before the treatment period for treated and controls separately.

Figure 3. Trends in sickness absence prior to the treatment.

The figure shows that both the levels and the trends in absenteeism prior to the treatment are similar for treated and non-treated. Hence, the assumption that there would be a common trend in the absence of treatment seems reasonable.

Second, I apply the same difference-in-differences estimator to periods before and after the actual treatment period and test whether α in specification (2) is equal to zero. Out of 14 falsification tests, two find estimates of α that are almost significantly (at ten percent level) different from zero (one positive and one negative). The remaining 12 are highly insignificant (five with negative point estimates and seven with positive point estimates). Since we do not find any strong significant treatment effects during periods where we, a priori, have no reason to expect an effect, we have strong indications that the main assumption behind the estimator is valid and that the estimates indeed capture causal effects. Third, there are some differences in average characteristics across establishment. For instance, treated in the disability sector are somewhat younger and might therefore have a different absence trend than the control group. Such a pattern would violate the main assumption behind the estimator. As a robustness check, I interact the time dummies with a number of individual characteristics, e.g. age, gender, education and contract form. The results do not change when these interaction terms are included when estimating equation (2).

Endogenous mobility

When individuals are randomly assigned to treatment, there is – in theory – no association between the treatment status and other variables that might affect the outcome. In practice, however, experiments are often plagued by non-compliance with the experimental protocol (e.g. Angrist and Krueger, 1998).

Individuals assigned to treatment may refuse to participate or drop out during the course of the study. Control groups members may move into treatment or find close substitutes to the treatment. If this non-compliance is non-random, assignment is randomized but actual exposure is not – and the estimated effect of the bonus and environment will capture the impact of the "intention to treat" rather than the treatment itself (e.g. Heckman *et al*., 1999).

A drawback with site randomization is that entering and leaving the establishment implies entering and leaving treatment status. Approximately 20 months elapsed from the moment when employees were first informed about the treatment until the treatment expired. Therefore, the issue of cross-establishment mobility should be investigated and taken into account. Theoretically, the presence of such cross-establishment movements will tend to systematically bias the treatment coefficient towards zero since the true treatment status is measured with error by our proxy that ignores mobility.

There is only weak evidence of non-compliance in the data. Of all employees assigned to the treatment group, approximately 90 percent did indeed receive treatment. This seems to correspond to the average mobility out of establishments during the period. Hence, it is unlikely that the opportunity to receive half a day off or increased spending on working environment led to significant flows into treated establishments. It is possible that the inconvenience and costs associated with a change of workplace compared to a potential gain during a limited number of months, deterred individuals from non-compliance.

Concerning the environment treatment, it is possible that individuals in the control group would find close substitutes for the experimental treatment from other sources. For example, establishments could have changed their environment through e.g.

organizational changes during the same period. It is therefore hard to estimate the treatment effect on those who actually received treatment. For the bonus treatment, however, non-compliance is probably less of a problem since the introduction of the bonus was a unique event in the municipality.

Given the assumption that treatment assignment only influences sickness absence by its effect on actual treatment, I use instrumental variable techniques to estimate the local average treatment effect. Since there seems to be perfect exclusion of the control group from the treatment, it can be shown that this estimator corresponds to the average effect of the treatment for the treated (SATE); see Angrist *et al.*, 1996.

To estimate the impact of treatment on those who actually received treatment, we can therefore inflate the estimate by the fraction of the treatment group receiving treatment (e.g. Krueger, 1999). Hence, compared to the estimate of the “intention to treat”, the estimate of the impact of “treatment on the treated” is approximately 10 percent higher.

External Validity

An important issue is whether the results in this study are true only for these particular individuals or if they can be generalized to other individuals, other places, and other time periods. Extrapolating results from a study to other settings is always associated with uncertainty. First, it should be stressed that the data used in this study is a non-representative sample of the Swedish workforce. It consists to a larger degree of women and elderly, i.e. subgroups with higher sickness absence rates than other subgroups’. Furthermore, individuals employed in the public sector may respond differently to changes in economic incentives and job environment than other groups. In addition, differences in timing imply that other factors may vary from what they were when the experiment was conducted. Hence, extrapolating parameter estimates from this population to other parts of the distribution can be misleading. Still, even if magnitudes differ it is less likely that directions would. Obviously, only further empirical analysis can answer such a question.

6. CONCLUDING REMARKS

A classical and important question in economics is how incentives and moral hazard affect absence behavior. This influences, for instance, the optimal design of an insurance system balancing economic security and incentives to work. The purpose of this paper has been to empirically evaluate how economic incentives and working environment affect absence behavior. Estimating these causal effects is associated with a number of difficulties. For example, replacement ratios and workplace characteristics are most likely correlated with other factors affecting absence. If these factors are not controlled for, the estimated effect will be partly attributed to this unobserved heterogeneity. There is also a lack of short-term absence data, since Swedish employers cover the cost of the first 14 days of a sickness spell.

A social experiment in Stockholm provides an opportunity to circumvent some of the problems plaguing many studies in empirical social insurance economics. In this rare experiment, establishments were randomly assigned into two different kinds of treatment with other establishments serving as controls. By following individuals before, during and after the treatment periods, it is possible to identify the causal effects of the programs on sickness absence.

The results show a large and significant effect of the bonus treatment, whereas the results from the environment treatment are smaller in magnitude and less conclusive. Individuals with the possibility to earn half a day off reduced their sickness absence during treatment by approximately 50 percent. The individuals in this sample have, on average, a sickness absence duration of 8 days. The introduction of the bonus reduces the actual replacement rate for such an individual by approximately 9 percent.[32] Although comparing this magnitude to estimates from previous research is very hard and should not be stressed too far, the effects obtained in this paper are of similar magnitude to previous studies based on Swedish data (e.g. Johansson and Palme, 1996).

This effect is not only statistically significant but also *economically* significant. The program reduced the overall costs of sickness absence, since employees received half a day off but reduced their absence by 1.2 days. In addition, the effect was

[32] In Sweden, there is no compensation paid out for the first day of absence. From day 2 to 8, the compensation rate is 80 percent. Hence, the actual compensation rate without the bonus is 70 percent [i.e. ((0+80*7))/8] and 63.7 percent with the bonus [i.e. ((0+80*7-50))/8].

especially large for highly educated – and well-paid - individuals, making the program even more economically advantageous.

The design of the experiment makes it possible to further assess the interpretation of the results. Although the randomization was conducted on establishments, not all employees at these establishments faced the same economic incentives. The estimations show that, in general, subgroups facing the largest incentives respond stronger, indicating strongly that the incentives per se was the main factor explaining the effect.

There remain, however, potential pathways that we can not observe and, hence, the single parameter estimate has no "structural" interpretation. For example, the increased economic incentive may have encouraged people to go to work in spite of sickness. If so, the long-term effects of the reform may be substantially different from the short-run effects.

REFERENCES

Allen, S G (1981), "An Empirical Model of Work Attendance", *Review of Economics and Statistics*", Vol. 63, 77-87.

Angrist, J and A B Krueger (1998), "Empirical Strategies in Labor Economics", in O Ashenfelter and D Card, editors, *Handbook of Labor Economics*, Vol. 3, North-Holland.

Angrist, J, D Imbens and D Rubin (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, Vol. 91, 444-472.

Arai, M and P Skogman Thoursie (2004), "Sickness Absence: Worker and Establishment effects", *Swedish Economic Policy Review*, Vol. 11, 9-28.

Arai, M and P Skogman Thoursie (2005), "Incentives and Selection in Cyclical Absenteeism", *Labour Economics*, Vol. 12, 269-280.

Barmby, T, M Ercolani and J Treble (2004), "Sickness Absence in the UK: 1984-2002", *Swedish Economic Policy Review*, Vol. 11, 65-88.

Bell, S, L Orr, J Blomquist and G Cain (1995), *Program Applicants as a Comparison Group in Evaluating Training Programs*. W E Upjohn Institute for Employment Research, Michigan.

Bertrand, M, E Duflo and S Mullainathan (2004), “How Much Should We Trust Differences-in-Differences Estimates?”, *Quarterly Journal of Economics*, Vol. 119, 249-275.

Besley, T J and A Case (2000), “Unnatural Experiments? Estimating the Incidence of Endogenous Policies”, *The Economic Journal*, Vol. 110, 672-694.

Bolin, K, L Jacobson and B Lindgren (2002), “Employer Investments in Employee Health. Implications for the Family as Health Producer”, *Journal of Health Economics*, Vol. 21, 563-583.

Broström, G, P Johansson and M Palme (2004), “Economic incentives and gender differences in work absence behavior”, *Swedish Economic Policy Review*, Vol. 11, 33-63.

Grossman, M (1972), “On the Concept of Health Capital and the Demand for Health”, *Journal of Political Economy*, Vol. 80, 223-255.

Grossman, M (2000), “The Human Capital Model”, in A J Culyer and J P Newhouse, editors, *Handbook of Health Economics*, North-Holland.

Heckman, J, R LaLonde and J Smith (1999), “The Economics and Econometrics of Active Labor Market Programs”, in O Ashenfelter and D Card, editors, *Handbook of Labor Economics*, Vol. 3, North-Holland.

Henrekson, M and M Persson (2004), “The Effects on Sick Leave of Changes in the Sickness Insurance System”, *Journal of Labor Economics*, Vol. 22, 87-114.

Hesselius, P, P Johansson and L Larsson (2005), “Monitoring sickness insurance claimants: evidence from a social experiment”, IFAU Working Paper 2005:15, Institute for Labour Market Policy Evaluation.

Ichino, A and R Riphahn (2004), “Absenteeism and employment protection: Three case studies”, *Swedish Economic Policy Review*, Vol. 11, 95-114.

Johansson, P and M Palme (1996), “Do Economic Incentives Affect Work Absence? Empirical Evidence Using Swedish Micro Data”, *Journal of Public Economics*, Vol. 59, 195-218.

Krueger, A B (1999), “Experimental Estimates of Education Production Functions”, *Quarterly Journal of Economics*, Vol. 114, 497-532.

Krueger, A B (2000), “Labor Policy and Labor Research Since the 1960s: Two Ships Sailing in Orthogonal Directions?”, in G Perry and J Tobin, editors, *Economic Events, Ideas and Policies: The 1960s and After*, Brookings Press.

Krueger, A B and B Meyer (2002), “Labor Supply Effects of Social Insurance”, in A Auerbach and M Feldstein, editors, *Handbook of Public Economics*, Vol. 4, North Holland.

LaLonde R J (1986), “Evaluating the Econometric Evaluations of Employment and Training Programs with Experimental Data”, *American Economic Review*, Vol. 76, 604-620.

Larsson, L (2006), “Sick of Being Unemployed? Interactions Between Unemployment and Sickness Insurance”, forthcoming in *Scandinavian Journal of Economics*.

Palmer, E (2004), “Sjukskrivningen i Sverige”, in T Theorell, editor, *Den höga sjukfrånvaron – sanning och konsekvens*, Swedish National Institute for Public Health.

Shapiro, C and J E Stiglitz (1984), “Equilibrium Unemployment as a Worker Discipline Device”, *American Economic Review*, Vol. 74, 433-444.

ESSAY III

BLIND DATES: QUASI-EXPERIMENTAL EVIDENCE ON DISCRIMINATION*

with Per-Anders Edin

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 and 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

* 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.

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 and 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 and 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 and 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 depends 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 and 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, gender and age.

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 and 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 the Employment Office, 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.[33] 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 33 weeks and that a third of the sample has been contacted by an

[33] 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.

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.

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

Table 1 (*continued*)

Variable	Full sample	Blind observations only (name or sex)	LINDA (Händel)
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
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 parental 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 and 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 and 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.[34] The fact that the observable variables determine "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.[35] In the first column we report estimates from our first specification

[34] 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).
[35] Using Logit models we obtain the same qualitative results.

that relies on the assumption of selection on observables. Here we restrict ourselves to the sub-sample of individuals with no concealed 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 labor market experience, has a clear positive effect on the probability to get contacted. Also, employed applicants face significantly higher probabilities of getting a contact.[36]

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.

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

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

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 and Åslund, 2001).

Taking the results in column 1 at face value, we find that employers 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 potential 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 letter.[37] Still, we cannot be sure that we can capture all relevant information in our specification. Therefore, we turn to our second identification strategy.

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

[37] 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.

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 indirectly 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. It is 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 and 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.[38] The male occupations are "Legislators, senior officials and managers", "Craft and related trades workers", and "Plant and machine operators and assemblers". The female 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

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

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

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

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

Goldin, C and C Rouse (2000), "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians", *American Economic Review,* Vol. 90, 715-741.

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

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

Meyerson, E and 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 and J Rich (2002), "Field Experiments of Discrimination in the Market Place", *Economic Journal* 112, F480 – F518.

Riach, P A and 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 A1. 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
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

Table A1 (*continued*)

Variable	All The Applicant Database	All LINDA (Händel)	Unemployed The Applicant Database	Unemployed LINDA (Händel)
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 A2. 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)

Table A2 (*continued*)

	Full sample	Men	Women
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	-.022	-.016	-.028
occupation	(.007)	(.009)	(.010)
University student	-.047	-.080	-.021
	(.013)	(.019)	(.017)
In other training	-.062	-.059	-.068
	(.016)	(.022)	(.022)
On parental leave	-.003	.043	-.003
	(.032)	(.154)	(.034)
Drivers license	.004	-.003	.009
	(.008)	(.012)	(.011)
# 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, as well as a constant.

APPENDIX 3: Baseline models

Table A3. Linear probability models of employer contact

	Non-Blind Sample	Full Sample	Female dominated branches	Male dominated branches
Duration in the database	.011	.010	.009	.010
(weeks)	(.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)

Table A3 (*continued*)

	Non-Blind Sample	Full Sample	Female dominated branches	Male dominated branches
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)
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	.027	.032	.055	.018
occupation	(.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 parental leave	.069	.059	.153	.188
	(.057)	(.050)	(.080)	(.090)
# 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, as well as a constant. 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 A4. 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)
Managerial experience	.037	.051	.060	-.032
	(.012)	(.011)	(.027)	(.042)
Telecommuting experience	.026	.026	-.068	.073
	(.017)	(.015)	(.044)	(.059)
Research experience	.014	.004	.168	-.027
	(.024)	(.022)	(.128)	(.159)
≥ 5 years work experience	.035	.024	-.004	.078
	(.013)	(.012)	(.025)	(.040)
No work experience	-.017	-.030	-.028	-.063
	(.015)	(.013)	(.032)	(.046)

Table A4 (*continued*)

	Non-Blind Sample	Full Sample	Female dominated branches	Male dominated branches
Labor market status:				
Employed in preferred occupation	.027	.032	.055	.022
	(.011)	(.010)	(.021)	(.033)
University student	-.032	-.025	-.038	-.179
	(.020)	(.018)	(.053)	(.060)
In other training	.024	.024	.060	-.004
	(.025)	(.023)	(.049)	(.073)
On parental leave	.069	.059	.152	.196
	(.057)	(.050)	(.079)	(.099)
Drivers license	.011	.004	.016	.078
	(.013)	(.012)	(.022)	(.042)
Private e-mail included	-.014	-.023	-.078	-.110
	(.023)	(.021)	(.050)	(.082)
Length of the private letter (# letters/100)	.00007	.0003	.00001	.0007
	(.0006)	(.0006)	(.001)	(.003)
Unknown words in the private letter	.005	.006	.005	.007
	(.003)	(.003)	(.005)	(.006)
Constant	-.027	-.010	-.150	.066
	(.033)	(.031)	(.051)	(.112)
# observations	6,657	8,043	1,837	703
R^2	0.2784	0.2819	0.2331	0.3298

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.

ESSAY IV

CASEWORKER EFFECTS AND PROGRAM EVALUATION*

1. INTRODUCTION

Already in the 1930's, the Swedish government provided public employment exchanges as an instrument to combat high and persistent unemployment (e.g. Calmfors *et al.*, 2001). Active labor market policies, i.e. measures to raise employment directly targeted at the unemployed, have since then become important ingredients of the Swedish welfare state, as in most other OECD countries.[39] In practice, these policies are often administrated and implemented at the local Employment offices. Here, unemployed job-seekers have to register in order to become entitled to unemployment benefits. At the office, they are matched to a caseworker with the mission to provide assistance throughout the unemployment spell.[40]

The purpose of this paper is to estimate if and how caseworkers matter. To empirically measure these effects is problematic due to the likely sorting of clients to caseworkers. Using only offices that practice random within-office distribution of job-seekers to caseworkers, the results show that caseworkers explain a substantial part of the variation in future outcomes: Approximately 2-5 percent of the total variation in job-seekers' future employment status and earnings are explained by which caseworker they are randomly assigned.

---

* I am grateful for comments from Peter Skogman Thoursie, Per-Anders Edin, Peter Fredriksson, Patrik Hesselius, as well as seminar participants at Uppsala University. Thanks to IFAU and Louise Kennerberg for providing me with the data.

[39] Sweden spends approximately 21 percent of its active labor market policy spending on public employment services, a level close to the OECD average (Cahuc and Zylberberg (2004)).

[40] Caseworkers have a key role when selecting individuals into active labor market programs. They can also support their clients in the job search process with initial interviews, training in resume preparation and interviewing, in-depth counseling, help in job finding and direct job placement. Furthermore, they have a monitoring function, e.g. by imposing economic sanctions if the client does not actively apply for jobs.

Explaining *why* some caseworkers are more successful than others is hard, since there may be pathways which we can not control for, raising doubts about a causal interpretation. Exploiting plausibly exogenous variation in past working strategies across caseworker, the results show that caseworkers that send their clients to classroom training or on-the-job training are less successful than caseworkers that provide basic job-search assistance.

This issue is important for several reasons. First, despite the fact that Sweden spends about 0.75 percent of GDP on public employment services, the overall importance of caseworkers has never been investigated empirically using Swedish data. Previous research has, however, investigated the effects of parts of what caseworkers do, e.g. job-search assistance and allocation of services.

Job-search assistance is one of few policies where international studies often find positive effects on the probability of finding a job (e.g. Martin and Grubb, 2001). The few Swedish studies on the subject find positive effects as well. Delander (1978) shows that a group of randomly selected individuals who received intensified job-search assistance ended up with significantly higher employment rates and wages than individuals in the control group. Engström *et al*. (1988) show that unemployed individuals receiving extra job-search assistance after a large firm closure did significantly better in terms of future employment status. Behrenz (1993) finds that a lower client/caseworker ratio increases the probability to find a job.[41] There is also a large literature on allocation of government services (e.g. Berger *et al*., 2001). For example, Lechner and Smith (2003) find that random assignment into active labor market programs would do about as well as the caseworker allocation, which in turn would be substantially outperformed by a deliberate allocation based on predicted impacts.

Second, every evaluation of active labor market programs has to address that selection into programs is usually associated with the potential outcomes that individuals would attain with and without the program. Several studies have documented that the probability to receive treatment is largely affected by individual

[41] There are a number of studies of U.S. social experiments, where individuals were randomly assigned into either a group subject to special treatment or a group receiving the existing services. Meyer (1995) concludes that, in general, these treatments led to substantial reductions in unemployment insurance receipts.

characteristics that are normally not observed nor accounted for, e.g. motivation and perceived intelligence (Bell and Orr, 2002; Heckman *et al.*, 1999).

This paper exploits the fact that Swedish caseworkers differ substantially in the way they assign clients into treatments in an instrumental variable approach. Assuming that differences in caseworkers' preferences in previous years towards different treatments are uncorrelated with unobserved present working strategies, I estimate the causal link by comparing outcomes for individuals whose caseworkers' preferences to treatment differ. Although these assumptions are fairly strong, this method provides an alternative way to estimate treatment effects even though the selection process into treatment is essentially a black box.

The rest of this paper is structured as follows. Section 2 provides a description of the data upon which I base the empirical analysis. Section 3 documents the importance of caseworkers. In section 4, I show that there is a large variation in the way caseworkers work and I exploit this variation in section 5 to estimate the effects of working strategies on the caseworker fixed effects. Section 6 concludes.

2. DATA

The analysis is based on data from the IFAU database, which contains information on the entire adult Swedish population from 1991 onwards. The database is specifically designed for evaluation and research on labor and education topics and is collected from databases maintained by the Swedish National Labor Market Board (AMS) and Statistics Sweden (SCB). The AMS data contains event history data on all individuals registered at the Employment Office at least once since 1991 with detailed information on the job search, e.g. caseworker assignment, search behavior, and participation in active labor market programs. The SCB database keeps records of annual earnings (based on firms' tax reports to the tax authorities) and employment status for all individuals residing in Sweden. There is also survey data on monthly wages (full-time) for approximately half of the private sector employees (aged 18-64) and for all public sector employees.[42]

[42] For the private sector, only companies with more than 500 employees are surveyed. The survey takes place in September and October.

A key issue that needs to be addressed when estimating the importance of caseworkers is the fact that job-seeker characteristics may differ systematically between caseworkers' clients within-office due to sorting. By e-mail, I contacted all 244 regular offices[43] and asked them about their matching of job-seekers to caseworkers.[44] After call-backs to nonrespondents, I obtained answers from 211 offices, corresponding to a response rate of about 86 percent.

At approximately 70 percent of the offices, job-seekers are non-randomly distributed to caseworkers. This sorting is conducted with respect to a number of factors (e.g. profession, prior unemployment, age, and skills). There are, however, 69 offices that apply randomized allocation. According to the survey, the main motive behind the randomization was that it generates an equal and fair distribution of the number of clients per caseworker within the office. In some cases, the randomization was achieved by drawing job-seekers' names from a list. The most common way of randomization (51 offices), however, was matching based on the clients' date of birth within a month.[45] Often, however, disabled individuals, long-term unemployed, youths under 25 years of age, and individuals classified upon registration as 'not job-ready' (e.g. in need of extra job-search assistance or career counseling) were matched to caseworkers working solely with these groups.[46]

I employ two different methods to investigate empirically whether offices claiming to conduct random allocation did indeed do so in practice. First, including information on caseworkers does not add explanatory power to a model predicting client characteristics (e.g. ethnic origin, family status, gender, and age). Second, there is only a small within-office variation in the number of clients per caseworker. Together, these patterns support the view that these offices did indeed practice random allocation.

The final sample consists of the subsample of the population of Swedish job-seekers since 1991 that results from imposing a number of exclusion criteria. I include only job-seekers that, according to the survey, are randomly distributed to caseworkers.

[43] In total, there are 301 Employment Offices in Sweden. Some offices support specific groups in the labor market, e.g. youths; long-term unemployed; individuals with a university degree; shipping personnel; artists; disabled persons or job-seekers in rehabilitation. These offices are excluded.

[44] The specific question was "In 2002, were job-seekers matched randomly to the caseworkers at your office? If not, how was the matching conducted?"

[45] That is, in an office with two caseworkers, one of the caseworkers takes all job-seekers born in the first 15 days within a month.

[46] In addition, there were sometimes exceptions from randomization, e.g. if the client wanted to change caseworker.

In particular, I keep individuals who registered for the first time (since 1991) in 2002 at an Employment Office practicing caseworker random allocation and have the following characteristics: age between 25 and 65, classified as 'job-ready' upon registration, not disabled. Column (i) in Table 1 presents descriptive statistics about these job-seekers.

Table 1. Descriptive statistics about the characteristics of the job-seekers

	(i) Random	(ii) All offices	(iii) Full population
Number of job-seekers	11,836	33,385	675,056
Number of job-seekers per caseworker	12.0	15.5	313.5
Number of job-seekers per office	171.5	136.8	2777.6
Number of offices	69	244	244
Personal characteristics:			
Age(years)	42.7	43.0	35.6
Female	0.50	0.49	0.53
Married/Living with a partner	0.59	0.60	0.39
Swedish Citizenship	0.79	0.79	0.90
Highest level of completed education:			
Primary	0.24	0.23	0.19
Secondary	0.40	0.41	0.60
University	0.36	0.35	0.21
Field of education:			
General	0.31	0.31	0.33
Aesthetic, classical	0.05	0.05	0.04
Pedagogic	0.03	0.04	0.05
Administration, trade	0.18	0.18	0.14
Industrial, handicraft	0.04	0.04	0.02
Transport, communication	0.20	0.21	0.19
Social and health care	0.01	0.01	0.02
Agriculture, woods, fishing	0.08	0.08	0.10
Service, civil guard, military	0.06	0.05	0.07
Missing, non-assignable	0.03	0.03	0.04
Outcomes:			
Employed Nov. 2003	0.54	0.54	0.61
De-registered from the Employment Service prior to November 2005 due to regular employment or lost contact	0.64	0.63	0.69
Earnings 2003, (SEK)	111,398	111,480	97,953
Reasons for de-registration:			
Employed	0.23	0.22	0.17
Employed on short-term contract	0.14	0.14	0.16
Reemployed by former employer	0.10	0.10	0.14
Other known reason	0.22	0.22	0.19
Lost contact	0.22	0.21	0.22
Education outside active labor market programs	0.09	0.10	0.11

Note: The data is for the inflow to the Employment Office in 2002. Column (i) includes only offices that practice random assignment of caseworkers to job-seekers. Column (ii) includes all unemployed individuals, given the exclusion restrictions. Column (iii) shows the characteristics of the population of all job-seekers at these Employment Offices during 2002.

Column (ii) includes all individuals with the same exclusionary restrictions at *all* offices (i.e. including offices with non-random allocation of caseworkers) Column (iii) shows the characteristics of the entire population (i.e. the stock) of job-seekers registered at these Employment offices during 2002.

There are several issues worth noting from Table 1. First, offices with random allocation are larger in terms of number of job-seekers. They also have fewer job-seekers per caseworker. Second, there are no distinctive differences in characteristics and outcomes between individuals at offices with random allocation and individuals at other offices. Third, our sample is not representative for the entire population of unemployed individuals. Column (iii) shows the characteristics of the total population of job-seekers registered in Sweden during the period. There are several differences between our sample and the generic population of unemployed. Due to our exclusion of youths and individuals with a history of unemployment, the full population is younger and less educated than our sample.[47] Therefore, it should also be noted that the total number of clients per caseworker is substantially larger in the full population.

3. CASEWORKER EFFECTS

The major problem when estimating the importance of personnel (e.g. caseworkers, teachers) is the fact that clients (e.g. job-seekers, students) are not randomly distributed within and across units (e.g. offices, schools). If, for example, there are caseworkers working with one type of job-seekers only (e.g. long-term unemployed, youths, blue-collar workers), we need to take these selection issues into account in order to estimate the true importance of caseworkers.

Caseworker effects: Do clerks matter?

This paper will estimate the importance of caseworkers in explaining job-seekers' future labor market outcomes. We do it by examining the caseworkers' explanatory power in job probability, earnings and wage equations. *If* job-seekers are non-randomly distributed to caseworkers *and* we can not control for these sorting effects, it is likely

[47] Sweden experienced a steep increase in unemployment in the early 1990s. For example, in 1994, more than 5 percent of the labor force participated in active labor market programs (e.g. Calmfors *et al.*, 2001).

that the caseworkers' explanatory power in the outcome equations partly captures job-seeker unobserved characteristic.

This analysis attempts to overcome these selection issues by using only clients that, according to the survey, are within office randomly distributed to caseworkers. Specifically, I use the ordinary least squares model and the empirical specification is given by:

$$Y_{ico} = \alpha_0 + X_i \beta + \tau_{co} + \varepsilon_{ico}, \qquad (1)$$

where X is a vector of individual characteristics (e.g. age, gender, ethnic origin, family status, county, office, length of unemployment, level of education, and education field), and τ is a vector of individual caseworker fixed effects.[48] Comparing estimations from equation (1) with and without caseworker dummies - using only offices with random distribution of caseworkers – provides our estimate of the importance of caseworkers.

In theory, there are no reasons to control for individual characteristic in equation (1) – given that job-seekers are within office randomly distributed to caseworkers and the sample is large. However, some caseworkers have relatively few job-seekers. Therefore, there might be systematic differences in caseworkers' client characteristics within an office. If these differences are not accounted for, they will falsely be attributed to the caseworkers. Including controls for job-seekers characteristics in (1) reduces the risk of such a bias.[49] The key assumption for the identification strategy to produce an unbiased estimate of the caseworker effects is therefore that job-seekers are randomly assigned to caseworkers, conditional on office and a vector of observable characteristics.

I use four different outcomes variables.[50] First, Y takes the value 1 if the individual is employed in November 2003 and 0 otherwise. Second, Y takes the value 1 if the individual has de-registrated from the Employment Service prior to November 2005 due to a regular job. Third, Y denotes log earnings in 2003 (annual sums). Finally, Y denotes log wages in September 2003. Estimation of specification (1) yields the results presented in Table 2.

[48] Since clients with the same caseworker are not perfectly independent observations, I apply throughout this paper clustered standard errors with respect to caseworkers. Clustering with respect to office does not change the conclusions.
[49] The results do not differ much when specification (1) is estimated with and without controls.
[50] It is not obvious which outcome variable that best captures client future outcome. For example, a faster exit from unemployment may result in a worse job placement.

Table 2. Labor market outcomes, offices with random allocation of caseworkers to client only

	Employment		Job		Earnings		Wages	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Caseworker dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	11,836	11,836	11,836	11,836	8,270	8,270	2,954	2,954
R-squared adjusted	0.1678	0.2056	0.1344	0.1861	0.0957	0.1337	0.3280	0.3244
F test (caseworkers)		1.590		2.030		1.397		0.974
p-value		0.000		0.000		0.000		0.655

Note: Specification (1) estimated using the ordinary least square model. Controls for offices, regions of residence, education, and a number of personal characteristics described in Section 2 are included. No controls for programs. Employment denotes employment status in November 2003. Job denotes whether the individual has de-registered from the Employment Service prior to November 2005 due to a regular job. Earnings and wages, respectively, denote log earnings during 2003 (only for individuals with earnings>0) and log wages in September 2003 (only for individuals with wages>0).

In Table 2, we see that caseworkers have a significant influence on job-seekers' outcomes. Columns (i) and (ii) show the effects on employment status in November 2003. Including caseworker fixed effects increases the explanatory power of specification (1) by about 4 percent. The caseworker dummies contribute even more when the outcome variable consists of a dummy indicating whether the job-seeker has de-registered from the Employment Service prior to November 2005 due to a regular job (columns (iii) and (iv)). This may reflect the fact that the caseworkers themselves administer this variable. Another reason may be that this variable captures long-term outcomes. Columns (v)-(viii) show that caseworkers affect log earnings in 2003, but there is no significant effect of caseworkers on monthly wages. It should be stressed, though, that this may partly be a result of the smaller sample size in the wage data.

In summary, these results imply that caseworkers affect their job-seekers' future employment and earnings. Even though the effect is not very large, it should be noted that since caseworkers within these offices are randomly assigned with respect to untreated outcome, the influence they have has to come through differential impacts by caseworkers.

Another question concerns the representativity of these results. Table 3 shows the results from estimation of equation (1) using all offices irrespective of how caseworkers are distributed to job-seekers. The results therefore rest on the assumption of selection of observables, i.e. that specification (1) controls for all sorting on caseworkers.

Table 3. Labor market outcomes, all offices

	Employment		Job		Earnings		Wages	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Caseworker dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	33,385	33,385	33,385	33,385	23,482	23,482	8,325	8,325
R-squared adjusted	0.1634	0.1887	0.1390	0.1788	0.1074	0.1321	0.3343	0.3345
F test (caseworkers)		1.491		1.805		1.379		1.000
p-value		0.000		0.000		0.000		0.492

Note: Specification (1) estimated using the ordinary least square model. Controls for offices, regions of residence, education, and a number of personal characteristics described in Section 2 are included. No controls for programs. Employment denotes employment status in November 2003. Job denotes whether the individual has de-registered from the Employment Service prior to November 2005 due to a regular job. Earnings and wages, respectively, denote log earnings during 2003 (only for individuals with earnings>0) and log wages in September 2003 (only for individuals with wages>0).

Table 3 shows that caseworkers matter with respect to their clients' future outcomes. Due to there being more individuals, the model explains slightly less of the total variation in outcomes than in Table 2. On the contrary, however, the gain in explanatory power from including caseworker fixed effects is slightly smaller. Most likely, this is due to the fact that the average number of clients per caseworker is larger. When trimming the samples, e.g. by restricting the analysis to caseworkers with 11-19 clients, the additional gain by including caseworker dummies is smaller in the sample with offices using random allocation.[51]

To evaluate the sensitivity and assess the validity of these results, I have performed a number of robustness checks. Table A1 in the Appendix presents regression estimates for three variations of equation (1). Row A includes only caseworkers with more than 40 job-seekers. A potential problem with previous estimates is that the average number of clients within some offices is fairly low. There may therefore be remaining differences between caseworkers' clients within an office that we do not control for, yielding

[51] In the sample with offices using random allocation, including caseworker dummies in employment and job equations yields F-values of 1.342 and 1.605, respectively, whereas using all offices F-values are 1.525 and 1.685.

inconsistent estimates of the caseworker effects (e.g. Rockoff, 2004). Estimating equation (1) using only caseworkers with many clients reduces the explanatory power of caseworkers slightly, but the effect is still large and highly significant. Row B adds a number of interaction terms (e.g. between gender and education and between age and education and length of unemployment). Introducing this more flexible form of equation (1) does not affect the main result that caseworkers explain a significant part of future outcomes.

These robustness checks indicate that the main results that caseworkers matter have internal validity, i.e. the results are not driven by incorrect model assumptions. Assessing the external validity is harder. Column (3) estimates the effect for the previous year (2001), with similar results. Hence, even though the results are not necessarily representative for other business cycles, they are not driven by any specific event occurring during 2002. Concerning our ability to extrapolate these results to other settings, it should be noted that these results are not necessarily true for other types of job-seekers (e.g. long-term unemployed, disabled, youths). However, *a priori* it is hard to think of any particular reason why caseworkers should be less important for other types of job-seekers than the job-seekers in our sample. On the contrary, it seems likely that the opposite is true. Caseworkers should, according to guidelines, focus primarily on “weak groups”, and it is therefore likely that caseworkers affect these groups at least as much as they affect the job-seekers in our sample.

4. WHY DO CASEWORKERS MATTER?

The previous section has shown that caseworker fixed effects explain a substantial part of the variation in job-seekers’ future outcomes. We can take the analysis one step further by asking: Why are some caseworkers more successful than others? However, there are many potential ways by which the caseworker may affect the job-seekers. To the extent that these unmeasured factors are correlated with the observed working strategies, our estimates of the impact of strategies will be biased. This section will first present evidence on what Swedish caseworker do. Then, I will estimate the effects of these strategies on the caseworker fixed effects, exploiting plausibly more exogenous variation across caseworkers in their working strategies in *previous* years as an attempt to overcome the confounding effects from non-observed factors.

What do Swedish caseworkers do?

The data contains detailed information on how the caseworkers work with their job-seekers. For example, there is information on participation in active labor market programs, the number of direct job placements and whether the job-seeker has to apply for jobs in a larger geographical area. Table 4 documents differences in working strategies among caseworkers in our sample.

Table 4. Descriptive statistics about working strategies among caseworkers (fractions)

	(i) All	(ii) Random
Program assignments:		
Classroom Training	0.124	0.136
	(0.231)	(0.247)
Employment Training and Preparatory Training	0.080	0.086
Courses (AMU)	(0.181)	(0.194)
Computer/Activity centers (CAC)	0.032	0.034
	(0.114)	(0.127)
Activities within Counseling Guidance and	0.037	0.043
Placement Service	(0.139)	(0.154)
On-the-Job Training	0.072	0.075
	(0.172)	(0.183)
Employability Rehabilitation Program	0.011	0.013
	(0.069)	(0.084)
Work Experience	0.061	0.061
	(0.156)	(0.165)
Projects with Employment Policy Orientation	0.009	0.009
	(0.069)	(0.063)
Job and Wage Subsidy Policies	0.017	0.020
	(0.083)	(0.091)
Start-up Grants	0.017	0.020
	(0.083)	(0.091)
Job Assignments:		
Applying for work outside the close vicinity of home	0.091	0.080
	(0.188)	(0.173)
Number of job assignments	0.520	0.498
	(1.190)	(1.259)
Personal characteristics:		
# caseworkers	2,155	990
# clients	33,385	11,836
# offices	244	69

Note: Column (i) includes all caseworkers employed at Swedish Employment Offices in 2002 and their working strategies. Column (ii) includes only the subset of the caseworkers employed at offices practicing random allocation of clients to caseworkers.

Following Heckman *et al.* (1999), I sort the programs into three overall categories: classroom training, subsidized on-the-job training, and job and wage subsidy policies. Classroom training typically lasts about four months, costs approximately $7,500 and are sometimes provided by firms.[52] On-the-job training programs provide job-relevant skills to disadvantaged workers.[53] Job and wage policies encompass one program where job-seekers are provided start-up grants.[54] Caseworkers also provide placement service, i.e. find available job slots that the job-seeker can apply for.

From Table 4 there are several things worth noting. First, there are no clear differences in working strategies depending on how the office distributes clients to caseworkers. Second, there is a striking variation in how caseworkers work. For example, on average, caseworkers assign 13 percent of the job-seekers in our sample to classroom training, but 25 percent of the caseworkers assign more than a quarter of their clients and 25 percent assign none.

These differences in treatment assignments between caseworkers are partly due to the lack of clear guidelines. In May and June 2003, I interviewed caseworkers at 30 Employment offices in order to learn more about how job-seekers are assigned into active labor market programs. Although informal, some conclusions can be drawn based on these interviews.

Although there are in same cases explicit criteria for program eligibility (e.g. age and unemployment duration restrictions for youth programs and programs for the long-term unemployed) or economic constraints, the caseworkers claim that they have significant discretion to determine the treatment status of their clients.[55] National guidelines are often seen as unclear and conflicting.[56]

---

[52] The purpose of the AMU program is to reduce bottlenecks in the labor market by providing skills necessary for particular jobs in excess demand. The CAC program aims at increasing the basic computer knowledge in the labor force. Training is usually combined with other activities.

[53] The Work Experience Program aims at helping the participant to keep in touch with working life. The participant should only carry out work that would otherwise not have been done.

[54] The Start-up Grants program provides the worker with an opportunity to start up a new business. AMU may precede Start-up Grants with the purpose to help the worker prepare himself for self-employment.

[55] One caseworker explains that in order for his clients to attend a program, they will have to show a paper proving they will get a job *before* the program actually begins.

[56] According to a national guideline, 70 percent of the participants in the most common classroom training program (AMU) should be employed 3 months after the end of the program. At the same time, caseworkers should primarily focus on ‘weak’ job-seekers. A caseworker assigning high-skilled job-seekers to the program would hence follow the first guideline, whereas a caseworker assigning low-skilled job-seekers would follow the second guideline.

This lack of rules has been documented in several studies. Richardson and van den Berg (2001) show that Swedish caseworkers assign individuals according to motivation and subjectively estimated unemployment duration and argue that an empirical evaluation must take this correlation into account. Eriksson (1997) finds that the heterogeneity of the caseworker is more important for determining program participation than the heterogeneity of the individuals.[57] Heckman *et al*. (1999) argue that there is a conflict between social objectives and caseworkers' incentives, generating this variation across caseworkers. Caseworkers concerned solely with their site's performance relative to the standard should admit into program applicants who are likely to be employed (the "cream") regardless of whether they would benefit from the program.

To summarize the results so far, there exist large differences in caseworkers' working strategies. The lack of guidelines has given substantial discretion for the caseworkers to choose how to best help their job-seekers. Even though the question of why job-seekers participate in active labor market programs is crucial for the choice of evaluation strategy, in practice this selection process is often a black box.

5. EXPLAINING THE CASEWORKER FIXED EFFECTS

The two previous sections have presented evidence that i) caseworkers explain a relatively large part of the variation in job-seeker outcomes, and ii) caseworkers differ substantially in the way they work. In this section, I will estimate the effects of these working strategies on the caseworker fixed effects to understand why some caseworkers are more successful than others. That is, we would like to estimate the following relationship:

$$\hat{\tau}_c = \mathrm{f}(\mathrm{S_c}) + \varepsilon_c, \qquad (2)$$

where $\hat{\tau}$ denotes caseworker c's predicted fixed effect from (1), S denotes her observed working strategies (e.g. program and job assignments) and ε is an error term capturing non-observed ways in which the caseworker affect the job-seekers' outcomes.

[57] Other studies have reported a large variation across caseworkers and offices in the monitoring of unemployed individuals; in the use of sanctions; and in forcing individuals to apply for jobs outside the vicinity of their homes (e.g. Olsson, 2004).

There are several potential pathways in which the caseworker may affect the client, i.e. the caseworker fixed effects have no clear 'structural' interpretation. Some of these pathways are observable in the data, but there may also be pathways that we can not control for (e.g. caseworker motivation or skills). If these left-out 'strategies' are correlated with the observable strategies, estimation of (2) will result in biased estimates of caseworker strategies.

One potential way to alleviate the endogeneity problems in (2) is to use other variables that produce variation in working strategies but have no direct effect on outcomes. Exploiting plausibly exogenous variation across caseworkers in their past program assigning behavior, I utilize an alternative strategy to evaluate active labor market programs. Frölich and Lechner (2004) exploit a similar kind of variation by using variation across jurisdictions in participant rates as an instrument to evaluate active labor market programs. Sweetman *et al*. (2003) use a related instrumental variables strategy to investigate the effect of receiving disability insurance in Canada.

It is important to understand under which assumptions this identification strategy captures the causal effects. We can model the observable (s) and unobservable (m) pathways through which the caseworker affects the job-seeker as follows:

$$s_{ij} = g_j + u_{ij}^p \quad (3) \qquad\qquad m_{ij} = f_j + u_{ij}^m , \quad (4)$$

where s and m denote strategies and motivation, respectively, that caseworker j uses with client i; g and f are caseworker specific time-invariant factors (e.g. preferences towards different strategies, and constant motivation); and the u's are error terms reflecting how caseworker j's strategies and motivation vary depending on client i.

For this identification strategy to produce unbiased causal effects of the strategies, one of the following assumptions has to be fulfilled:

$$\mathrm{cov}(y_i , f_j) = 0 \quad (5) \qquad\qquad \mathrm{cov}(f_j , g_j) = 0 , \quad (6)$$

That is, in order for past caseworker behavior to serve as an instrument to estimate the strategy effects, it must be the case that unobservable time-invariant caseworkers factors are validly excluded from the job-seekers outcome function *or* that these are uncorrelated with the constant working strategies. Although these assumptions are fairly strong – implying that the results should not be pushed too far - this method exploits a different source of variation for estimating treatment effects.

To estimate caseworkers' preferences towards different strategies, I construct the data in exactly the same way – but use individuals who registered at an Employment Office (applying random distribution of clients to caseworkers) for the first time one year *earlier*, in 2001. We can then estimate the following specification:

$$S_{ico,t-1} = \lambda_0 + X_{,t-1i}\theta + \pi_{co,t-1} + \upsilon_{ico,t-1}, \qquad (7)$$

where S is a vector of strategies (i.e. programs and job assignments), X is the same full vector of individual characteristics used in specification (1), and π is a vector of caseworker dummies. The purpose of this estimation is that these caseworker dummies should capture the caseworkers' preference to different strategies, i.e. how does the client's probability to participate in active labor market programs or receive placement service depend on the caseworker that she was randomly distributed. Controlling for job-seekers' observable characteristics X increases the likelihood that these effects are not attributed to differences in job-seeker characteristics between caseworkers within an office. Table 5 shows to what extent the caseworker vector in equation (7) can predict programs and job assignments.

Table 5. Linear probability models of working strategies

	F-test (caseworkers)	p-value
Program assignments:		
Classroom Training	3.310	0.00
On-the-Job Training	2.511	0.00
Job and Wage Subsidy Policies	3.901	0.00
Job assignments:		
Applying for work outside the close vicinity of home	1.224	0.00
Number of job assignments	1.220	0.00
# caseworkers	1,173	
# clients	10,187	
# offices	69	

Note: Controls for regions of residence, education, occupation and a number of personal characteristics described in Section 2 are included. Registered for the first time at the Employment Office in 2001.

Table 5 reveals that caseworkers affect both the number of job assignments and the participation in active labor market programs. Although the effects are statistically significant, they account for only a relatively small fraction of the total variation in treatments. Consequently, past caseworker behavior are fairly weak instruments for identifying the causal effects of strategies (e.g. Stock, 2001). If the instrument is weak, any inconsistency due to (an even small) correlation between the instrument and the error term in the outcome equation will be "blown up", making the estimator biased.

We can now predict a job-seeker's programs and job assignments using only the parameter estimates from the caseworker vector in equation (7) and use it to estimate the following specification:

$$\hat{\tau}_c = \mathrm{f}\left(\hat{\mathrm{S}}_{\mathrm{c}}\right) + \varepsilon_c, \quad (8)$$

where τ are the predicted caseworker fixed effects from (1) and S denotes the predicted strategies using only the parameter estimates from the caseworker vector in (7). Since the first step equation uses individuals in the previous year, only individuals with caseworkers working in both periods are used. Table 6 presents regression estimates for four variations of equation (8).

Table 6. Ordinary Least Squares estimates of the caseworker fixed effects

	Caseworker fixed effects			
	(i) (from employment)	(ii) (from job)	(iii) (from log earnings)	(iv) (from log wages)
Program assignments (ref. no program treatment)				
Classroom Training	-0.181***	-0.177***	-0.657***	-0.021
	(0.038)	(0.036)	(0.113)	(0.025)
On-the-Job Training	-0.044	-0.245***	0.083	0.003
	(0.045)	(0.044)	(0.142)	(0.033)
Job and Wage Subsidy Policies	-0.013	-0.026	0.086	0.031
	(0.079)	(0.078)	(0.240)	(0.059)
Job assignments:				
Applying for work outside the close vicinity of home	-0.075	-0.022	-0.129	0.036
	(0.063)	(0.062)	(0.198)	(0.045)
Number of job assignments	0.001	0.003	-0.007	0.001
	(0.005)	(0.005)	(0.015)	(0.004)
# observations	819	819	761	563
R-squared	0.045	0.102	0.051	0.003

Note: Specification (8) estimated using the ordinary least square model. The dependent variable is the predicted caseworkers fixed effects from (1). All regressors are instrumented using the caseworker vector from equation (7). Robust standard errors in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Columns (i) and (ii) show the effects of strategies on the caseworker fixed effects obtained from using employment status in November 2003 and de-registration from Employment office prior to November 2005, respectively. The results show that caseworkers that send their job-seekers to classroom training or on-the-job training are less successful than caseworkers that do not send their clients to any program at all but provide the ordinary job-search assistance. Hence, both types of programs result in a lower probability of employment. Since the caseworker administer the de-registration variable used to provide the outcome variable in column (ii), it is possible that this variable captures other characteristics of the caseworker and may therefore violate the assumptions needed to interpret these estimates as causal. Column (iii) shows that there is also a negative effect of classroom training on earnings.[58]

6. CONCLUDING REMARKS

This paper has shown that a substantial fraction of the variation in future outcomes of the job-seekers at the Swedish Employment Offices is due to their caseworkers. The pattern is clear with respect to both future employment status and future earnings. Since caseworkers play a crucial role for the success or failure of the job search process, it is likely that a policy focusing on the intermediate work at the Employment Office has the potential to be successful.

Understanding why caseworkers matter is harder and any estimation must rely on fairly strong assumptions. There are many potential explanations for the substantial differential impacts of caseworkers. Interviews show that caseworkers find themselves with a discretionary power to decide how to work with their clients. This is confirmed in the data that show large variation across caseworkers in the way caseworkers provide placement services and assign individuals into active labor market programs.

Several potential pitfalls have to be addressed when estimating the causal effects of active labor market programs on future outcomes. Ideally, we would like to manipulate treatment status in a way that does not affect outcome directly but only through its effect on treatment. Then, any systematic differences in outcomes between treated and non-treated individuals can be attributed to the treatment. The analysis exploits information

[58] Columns (iii) and (iv) excludes caseworkers that have no job-seekers with positive earnings in 2003 and positive wages in September 2003.

specific to the employment office clerks in an instrumental variable framework to identify why caseworkers matter.

The empirical results show negative effects of classroom training on both the probability to find employment and future earnings. The estimated effect of the job and wage subsidy program is close to zero and not statistically significant. In none of the cases are there any significant effects on wages, suggesting that the treatment affect primarily employment status and not productivity. Comparing these results to results from previous studies is hard, since results vary a lot between studies. However, evaluations of classroom training programs usually find insignificant or significantly negative effects of these programs. Calmfors *et al*. (2001) summarize the literature and conclude that classroom training programs in the 1990s have not enhanced the employment probabilities of participants, whereas some forms of subsidized employment seem to have had such effects.

Although this identification strategy rests on fairly strong assumptions, it provides an alternative strategy for addressing the potential problem of omitted variable bias by comparing otherwise similar unemployed individuals who receive different treatment simply because the randomness of caseworker assignment. Here, the potential selection problem stems not from confounding variables affecting both job-seeker outcomes and job-seeker treatment status. Several studies have documented that the probability to receive treatment is largely affected by individual characteristics that are normally not accounted for. Therefore, more focus on these issues is needed both in active labor market policy and in the evaluation literature.

REFERENCES

Angrist, J and A B Krueger (1998), "Empirical Strategies in Labor Economics", in O Ashenfelter and D Card, editors, *Handbook of Labor Economics*, Vol. 3, North-Holland.

Behrenz, L (1993), "Effekt- och effektivitetsanalyser av 1987 års personalförstärkning till arbetsförmedlingen", Licentiate dissertation, Lund University.

Bell, S H and L L Orr (2002), "Screening (and creaming?) applicants to job training programs: the AFDC homemaker-home health aide demonstrations", *Labour Economics*, Vol. 9, 279-301.

Berger, M, D Black and J Smith (2001), "Evaluating Profiling as a Means of Allocating Government Services", in M Lechner and F Pfeiffer, editors, *Econometric Evaluation of Active Labour Market Policies*, Heidelberg, Physica.

Bound, J, D A Jaeger and R Baker (1995), "Problems with Instrumental Variables Estimation when the Correlation between Instruments and the Endogenous Explanatory Variables is weak", *Journal of the American Statistical Association*, Vol. 90, 443-450.

Cahuc, P and A Zylberberg (2004), *Labor Economics*, MIT Press, London.

Calmfors, L, A Forslund and M Hemström (2001), "Does active labour market policy work? Lessons from the Swedish experiences", *Swedish Economic Policy Review*, Vol. 8, 61-131.

Delander, L (1978), "Studier kring den arbetsförmedlande verksamheten", in SOU 1978:60, *Arbetsmarknadspolitik i förändring*, Fritzes, Stockholm.

Engström, L, K-G Löfgren and O Westerlund (1988), "Intensified Employment Services, Unemployment Durations, and Unemployment Risks", Umeå Economic Studies 186, Umeå University.

Eriksson, M (1997), "Placement of unemployed into labour market programs: A quasi-experimental study", Umeå Economic Studies 439, Umeå University.

Friedlander, D, D H Greenberg and P K Robins (1997), "Evaluating Government Training Programs for the Economically Disadvantaged", *Journal of Economic Literature*, Vol. 35, 1809-1855.

Frölich, M and M Lechner (2004), "Regional treatment intensity as an instrument for the evaluation of labour market policies", IZA Discussion paper 1095, IZA.

Heckman, J, R LaLonde and J Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", in O Ashenfelter and D Card, editors, *Handbook of Labor Economics*, Vol. 3, North-Holland.

Lechner, M and J Smith, "What is the Value Added by Caseworkers?" (2006), forthcoming in *Labour Economics*.

Martin, J P and D Grubb (2001), "What works and for whom: A review of OECD countries' experiences with active labour market policies", *Swedish Economic Policy Review*, Vol. 8, 9-56.

Olsson, M (2004), "Lokala arbetsmarknader och tillämpningen av reglerna för geografiskt sökområde i arbetslöshetsförsäkringen", IAF Working Paper, Swedish Unemployment Insurance Board.

Richardson, K and G J van den Berg (2001), “The effect of vocational employment training on the individual transition rate from unemployment to work”, *Swedish Economic Policy Review*, Vol. 8, 175-213.

Rockoff, J (2004), ”The Impact of Individual Teachers on Student Achievement: Evidence from Panel data”, *American Economic Review*, Vol. 94, 247-252.

Stock, J H (2001), “Instrumental Variables in Economics and Statistics”, entry in *International Encyclopedia of the Social and Behavioral Sciences*, Amsterdam: Elsevier, 7577-7582.

Sweetman, A, W Warburton, L DeBenedictis and R Warburton (2003), “Disability Benefits and Impacts on Welfare Dependence and Health”, Mimeo, Queen’s University.

APPENDIX: Robustness checks

Table A1. Robustness checks

		Employment		Job		Earnings	
		(i)	(ii)	(i)	(ii)	(i)	(ii)
A. Caseworkers with >40 job-seekers	Caseworker dummies	No	Yes	No	Yes	No	Yes
	Observations	5,440	5,440	5,440	5,440	3,780	3,780
	R-squared adjusted	0.1538	0.1710	0.1470	0.1700	0.0840	0.1044
	F test (caseworkers)		2.506		3.116		2.198
	p-value		0.000		0.000		0.000
B. More flexible form	Caseworker dummies	No	Yes	No	Yes	No	Yes
	Observations	11,836	11,836	11,836	11,836	8,270	8,270
	R-squared adjusted	0.1702	0.2088	0.1434	0.1945	0.0994	0.1379
	F test (caseworkers)		1.578		1.752		1.406
	p-value		0.000		0.000		0.000
C. Other time period (2001)	Caseworker dummies	No	Yes	No	Yes	No	Yes
	Observations	10,187	10,187	10,187	10,187	5,648	5,648
	R-squared adjusted	0.1611	0.2202	0.1235	0.1909	0.0920	0.1652
	F test (caseworkers)		1.652		1.717		1.671
	p-value		0.000		0.000		0.000

Note: Specification (1) estimated using the ordinary least square model. Controls for offices, regions of residence, education, and a number of personal characteristics described in Section 2 are included. No controls for programs. Employment denotes employment status in November 2003. Job denotes whether the individual has de-registered from the Employment Service prior to November 2005 due to a regular job. Earnings denotes log earnings during 2003 (only for individuals with earnings>0). In row C, employment and earnings denote employment status in November 2002 and earnings in 2002, respectively.

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

Rapporter/Reports

2006:1 Zenou Yves, Olof Åslund & John Östh ”Hur viktig är närheten till jobb för chanserna på arbetsmarknaden?”

2006:2 Mörk Eva, Linus Lindqvist & Daniela Lundin ”Påverkar maxtaxan inom barnomsorgen hur mycket föräldrar arbetar?”

2006:3 Hägglund Pathric ”Anvisningseffekter” – finns dom? Resultat från tre arbetsmarknadspolitiska experiment”

2006:4 Hägglund Pathric ”A description of three randomised experiments in Swedish labour market policy”

2006:5 Forslund Anders & Oskar Nordström Ska

2006:6 Johansson Per & Olof Åslund ”'Arbetsplatsintroduktion för vissa invandrare' – teori, praktik och effekter”

Working Papers

2006:1 Åslund Olof, John Östh & Yves Zenou “How important is access to jobs? Old question – improved answer”

2006:2 Hägglund Pathric “Are there pre-programme effects of Swedish active labour market policies? Evidence from three randomised experiments”

2006:3 Johansson Per “Using internal replication to establish a treatment effect”

2006:4 Edin Per-Anders & Jonas Lagerström “Blind dates: quasi-experimental evidence on discrimination”

2006:5 Öster Anna “Parental unemployment and children’s school performance”

2006:6 Forslund Anders & Oskar Nordström Skans “Swedish youth labour market policies revisited”

2006:7 Åslund Olof & Per Johansson “Virtues of SIN – effects of an immigrant workplace introduction program”

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"