Hiring subsidies for people with a disability: Helping or hindering? - Evidence from a field experiment

Eva Deuchert, Lukas Kauer*

Center for Disability and Integration, School of Economics and Political Science, University of St.Gallen, Rosenbergstrasse 51, CH-9000 St.Gallen, Switzerland

6/4/2013

Preliminary draft

Please do not quote and do not circulate this draft without permission of one of the authors
Comments are very welcome

Abstract: Many countries provide hiring subsidies aimed at promoting the employment of people with disabilities. The effectiveness of these subsidy schemes remains unclear. The subsidy lowers wages and may thus increase employment, but may also signal lower quality of the applicant (who has to disclose a disability), which deter employers from hiring. This paper evaluates the effectiveness of employer incentives provided by the Swiss Disability Insurance using a small scale social field experiment. Participants write application letters, where it is randomly decided whether the application discloses the subsidy to the potential employer or not. The effectiveness of the hiring subsidy is measured by call-back rates for interviews. The study is conducted in two waves. The first wave focuses on graduates from sheltered Vocational Education & Training Programs. The second wave (ongoing) is implemented in a sample of clients from employment consulting services. Our results reveal that the effectiveness strongly depends on the respective target group. While the subsidy is ineffective or even counterproductive in a group of adolescents who are at the end of their vocational training program, the subsidy is likely to increase call-back rates in a group of clients of job coaching services.

Keywords: hiring subsidies, effectiveness, social field experiment

JEL classification: I38, C93

Acknowledgements: Financial support was provided by the Ernst Göhner Foundation. Karina Schindler provided excellent research assistance. Furthermore, we would like to thank all participants and coordinating persons at the rehabilitation centers and at the local disability insurance office. We are grateful to Michael Lechner and participants of the 4th Biennial Conference of the American Society of Health Economists and of the international conference "Field Experiments in policy evaluation" in Nuremberg as well as of the 3rd Workshop "Arbeitsmarkt und Sozialpolitik" in Dresden for helpful comments.

* Corresponding author. Tel.: +41 (0)71 224 31 93; fax: +41 (0)71 220 32 90, lukas.kauer@unisg.ch

1 Introduction

Hiring subsidies are common in many industrialized countries.[1] These programs are designed to reduce labor costs and to stimulate employment for disadvantaged groups (such as young people, welfare recipients, or people with disabilities). As been discussed in Neumark (2011), the effectiveness and efficiency of these programs crucially depend on the existence of windfall profits (if hiring had taken place even without the subsidy), substitution effects (if employment of the targeted group increases at the expense of declining employment in other non-targeted groups), and signaling effects (if employers perceive hiring subsidies as a signal for lower productivity).

Since hiring subsidies generate considerable costs and can even be harmful, it is necessary to carefully evaluate them. In this paper, we empirically evaluate the effectiveness of a hiring subsidy scheme targeted at people with disabilities. From a policy perspective, this is an extremely relevant topic: The number of people with disabilities is high (on average across the OECD 14% of the working-age population classify themselves as disabled) and this group is particularly disadvantaged in the labor market (OECD, 2010). Many countries seek to improve employment prospects by implementing subsidy schemes targeted at people with disabilities (OECD, 2003). To be effective, potential employers need to be informed about the hiring subsidy (particularly if not all persons with a disability are eligible). Broaching the possibility of hiring subsidies within the application process, however, inevitably means for the individual to disclose the disability. The signaling effect may thus be likely particularly if employers associate a disability with lower productivity. Hence, there is the risk that these subsidy schemes may be ineffective or even counterproductive.

[1] See Marx (2001) or Katz (1996) for a review of different programs in the US and elsewhere.

To the best of our knowledge, there is only a single study which empirically evaluates the effectiveness of hiring subsidies targeted at people with disabilities (Gupta and Larsen, 2010). Different versions of this paper, however, also demonstrate how difficult it is to find a suitable empirical strategy for isolating the effect. The key problem is to identify an appropriate control group that is not targeted by hiring subsidies.[2] The empirical literature on hiring subsidies for unemployed and welfare recipients proposes three different methods to solve this problem: First, the gold standard is to ex-ante randomize eligibility for a subsidy. This is typically done in voucher experiments, where a randomized proportion of the target population receives a voucher for a wage subsidy, which can be handed over to an employer to cash it from a government agency (e.g. Bell and Orr, 1994; Burtless, 1985; Dubin and Rivers, 1993; Galasso *et al.*, 2004; Woodbury and Spiegelman, 1987). Second, in situations where field experiments are not feasible, natural experimental designs are used for ex-post evaluations (e.g. Hamersma, 2008; Huttunen *et al.*, 2010; Schünemann *et al.*, 2011). The results of these two methods are mixed. While some document (small) positive employment effects, others find that hiring subsidies are ineffective or even counterproductive. In many applications, however, a suitable control group cannot be identified. As a third method, researchers then compare participants in a subsidy program with a control group that either has an unsubsidized job or is still looking for a job (e.g. Carling and Richardson, 2004; Neubäumer, 2010; Sianesi, 2008). Here, estimates are mostly positive and usually much larger, particularly if program participants (who have a job at least for some time by construction) are compared to a general population (where many have no jobs).

[2] Gupta and Larsen evaluate the introduction of the Danish Flexjob scheme using a difference-in-difference framework. To be eligible for Flexjob, individuals must have a long-term disability (at least three years) and a reduction in working capacity. An earlier version (Gupta and Larsen, 2008) uses the general population as the control group, finding only modest employment effects. In a later version (2010), long-term disabled without reduction in working capacity and short-term disabled individuals (based on self-classification) are used as control groups, finding a substantial positive employment effect.

In this paper we propose a novel approach to evaluate a subsidy program when a suitable control group is not available. We conduct a field experiment among participants who are all eligible for a hiring subsidy program. Participants write several applications, where it is randomly decided whether the application discloses the subsidy to a potential employer or not. The effectiveness of the hiring subsidy is measured by call-back rates for interviews. We implement our field experiment in two different samples of people with disabilities, i.e. young people at the end of a sheltered dual track vocational education and training program, as well as clients of job coaching services. All study participants search for employment in the regular labor market and are targeted to hiring subsidies financed by the Swiss disability insurance. Our results document effect heterogeneities. The subsidy is ineffective or even counterproductive in a group of adolescents who are at the end of their vocational training program but is likely to increase call-back rates in a group of clients of job coaching services.

The rest of the paper is structured as follows: The following section provides background information on hiring subsidies targeted at people with disabilities and on the particular evaluated policy in Switzerland. It also discusses the theoretical effects of a subsidy. Section 3 describes the study design, and section 4 the two different samples of study participants and the number of applications sent. Results are presented in section 5. The final section concludes.

2 Background

Hiring subsidies targeted at people with disabilities have been implemented in many countries. The size of these programs differs considerably, ranging from 0.1% of the working-age population in Portugal to almost 11% in Sweden (OECD, 2003). Hiring subsidies can include direct wage subsidies and/or the compensation for (potentially higher) social security contributions, and differ in their temporary activation: Some countries (such as Belgium, Denmark

or France) implemented permanent subsidy schemes. The key idea here is that people with a disability are less productive, and that this permanently lower productivity needs to be compensated. Other countries (such as Austria, Norway, or Sweden for example) implemented temporary subsidy schemes that phase-out after some time. The motivation for these subsidy schemes is that on-the-job training of people with disabilities takes more time, but the productivity gap can be closed. A further motivation for temporary subsidies arises from the substantial uncertainty regarding the productivity of an applicant with a disability. Hiring subsidies can help to overcome these information asymmetries since the trial phase imposes lower costs to employers.

In this paper we evaluate the effectiveness of the subsidy scheme *Einarbeitungszuschuss* (adjustment grant) paid by the Swiss Disability Insurance (DI). This scheme was implemented with the fifth revision of the Swiss Disability Insurance Act (effective since January 2008) to provide employers with an incentive to hire workers with disabilities. Subsidies can be paid for a maximum of 180 days, when the initial performance does not correspond to the agreed wage equivalent.[3] The maximum amount of the grant is 80% of the previous wage and cannot exceed the current salary including social security contributions. Note that not all people with a disability are eligible for adjustment grants. The grant is generally agreed between the DI and a potential employer. Often, the DI guarantees a grant for an applicant already within the application process to support the process. Up to December 2010, the subsidy scheme was hardly used (less than 1'300 grants were permitted).[4]

From a theoretical point of view, the effect of a hiring scheme is unclear. This is demonstrated in Table 1. Hiring subsidies are implemented because policy makers assume that many

[3] In case the productivity is permanently lower, a reduced wage can be negotiated while the person may be eligible for additional DI pension to compensate for the wage loss.

[4] Personal information provided by the Swiss Federal Social Insurance Office.

firms employ a person with a disability only if a subsidy is paid but not if it is unavailable (*compliers*). The policy maker hopes that the benefits (e.g. reduced DI payments or higher tax revenue) will outweigh the costs of the subsidy. If firms do not employ people with disabilities with or without the subsidy (*never takers*), the subsidy is ineffective in increasing employment. However, at least no further costs are caused. Yet, there may be two other groups jeopardizing the (cost-)effectiveness of the subsidy: Firms may employ people with a disability irrespective of the regime (*always takers*). For these firms, the subsidy is also ineffective. Windfall gains, however, may occur if these firms apply for subsidies and the DI cannot distinguish whether the hiring would or would not have taken place without the subsidy. This affects the efficiency of the subsidy scheme. Finally, it may be the case that firms would hire people with disabilities without the subsidy, but refrain from doing so when a subsidy is paid (*defiers*). This situation is particularly likely when an applicant discloses his eligibility for the subsidy to the potential employer. With this information transfer, the employer also receives the information on the disability status and may interpret this as a signal for lower productivity.

One should further mention general equilibrium or substitution effects, i.e. if the employment effect among people with a disability is at the expense of declining employment in other non-targeted groups. This substitution effect can have two important consequences: (1) The policy causes additional costs (for example in form of higher benefits for non-targeted groups). This paper focuses on effectiveness rather than on efficiency and we therefore do not further consider this problem. (2) If non-eligible people are used as a control group, *SUTVA*[5]

TABLE 1: Possible effects of hiring subsidies

[5] *SUTVA* stands for stable unit treatment value assumption, which states that the observation on one unit should be unaffected by the particular assignment of treatments to the other units.

		Hiring subsidy	
		No hiring	*Hiring*
No hiring subsidy	*No hiring*	**Never taker** No employment effect (ineffective)	**Complier** Positive employment effect (incentive)
	Hiring	**Defier** Negative employment effect (signaling)	**Always taker** No employment effect (windfall)

violations are likely to bias the estimates for effectiveness. In our case this is unlikely to be the case since each person is used as their own control group (see section 3). Thus, as long as not all applications are sent to the same potential employers, *SUTVA* violations will not affect effectiveness estimates.

3 Study design

Since the policy has already been put in place, it would be difficult to implement a standard field experiment where the eligibility for a subsidy is randomized. We therefore chose an alternative design, where people with a disability who are both looking for sustainable employment and who are eligible for the hiring subsidy write several applications which either enclose the eligibility for the subsidy or not. It is randomly decided which application type is sent to a potential employer. In this design the same person acts as her own treatment and control group (within-estimator). We test whether the notification of the subsidy results in higher call-back rates for interviews.

Comparison to other evaluation methods

Our design is similar to other evaluation methods known in the literature but avoids some relevant pitfalls. Like in standard social experiments, the randomization provides a very powerful and creditable tool to identify causal treatment effects. Our design, however, avoids several limitations of standard social experiments (see e.g. Bell and Peck, 2012): First, we do not

need to withhold the treatment from a control group, which is often argued to be unethical. This problem might be particularly relevant in settings where people have a legal claim to the treatment and randomizing eligibility would not be possible. Second, the within-estimator has the advantage that we are able to estimate the impact of the treatment on the treated and not an intention to treat effect since we are not plagued by non-compliance (where the treatment group does not take the treatment) or substitution (where the control group receives a similar treatment that is available on the market). And finally, the design is relatively easy and cheap to implement and does not rely on the goodwill of the public administration. In contrast to standard field experiments, our design, however, comes with the disadvantage that we cannot evaluate whether the subsidy results in more job offers or in higher employment rates since it would not be possible to control whether the eligibility for a subsidy was disclosed to an employer during the job interview. We thus use call-backs for job interviews as our prime outcome variable. Furthermore, we are restricted to analyze the effect of the subsidy on the likelihood of call-backs in a single job search method (i.e. written application process). We cannot evaluate whether a subsidy impacts the success of other job search methods, such as informal job search via family and friends or the placement through DI job coaches.

Our design is similar to Falk *et al.* (2005) who evaluate the effectiveness of a training program comparing call-back rates to applications which were written short before and after participants attended the program. Note, however, that their approach relies on the before/after assumption, which means no other (uncontrolled) confounding factors change over time (Heckman *et al.*, 1999). Our approach, in contrast, randomizes contemporaneous applications into treatment (disclosure of the subsidy) and control (no disclosure). The implicit assumption is that employers would not be aware of the eligibility for hiring subsidies without the disclosure. This is likely in our current setting since caseworkers individually decide on eligibility,

which may cause substantial information costs for employers to discover whether an applicant is eligible or not (Bishop and Kang, 1991).

Finally, our design is similar to correspondence tests which are used to discover discrimination in the labor market (e.g. Bertrand and Mullainathan, 2004; Oreopoulos, 2011). These tests are often based on manipulated resumes that randomly vary by a specific characteristic (such as gender, ethnicity, etc.). One of the problems of correspondence tests is that interviews cannot take place in case the employer is interested in the applicant (usually, these tests rely on faked resumes). This raises ethical concerns against correspondence tests because the respective group may be marginalized even more, and because employers are burdened with costs for the screening process (Riach and Rich, 2004). Another problem is that these designs generate invalid results if correspondence tests are used too often (employers may thus be aware of the experiment and behave in a social desirable way). The prime reason that speaks against correspondence tests for evaluating policies is, however, that researchers may not able to realistically mimic applications from the respective group. In our setting, for example, many applications from study participants are flawed, not only grammatically but also formally. We probably would not have put together applications in such a way. In addition, with heterogeneous treatment effects, correspondence tests can only provide treatment effects for the chosen applicant profile (typically limited to very few profiles). With our design we avoid these problems and have furthermore the possibility to study effectiveness in different target groups.

Implementation

The experiment was implemented in two different waves. The main reason for this was to study effect heterogeneities with respect to different target groups. The first wave was conducted in a group of adolescents at the end of a sheltered dual track vocational education and

training (VET) program, who search for employment in the competitive labor market. This group is very often targeted by these subsidy schemes. The experiment was part of an application workshop which took part in four different rehabilitation centers in which students completed their VET program. Almost all students who participated in the workshop also took part in our experiment. In the second wave, the group consists of clients of job coaching services provided by the local cantonal DI office. The recruitment of potential participants was made by the DI case workers. Interested clients contacted researchers by email or telephone to make an appointment for an individual application workshop at the offices of the University of St.Gallen. In this group, selection into the experiment is thus much more likely: Not all eligible clients may have been recruited by DI case workers, and not all recruited clients may have taken active steps to participate in our experiment.

The experimental design is not identical in the two waves (for main differences, see Table 2). A very important difference between the two waves is the way the eligibility for a subsidy was notified. Participants in wave 1 had only an oral confirmation that a job would be subsidized. They thus wrote two cover letters for each selected job application. The letters were identical except for one feature: One letter added a paragraph in which the subsidy is mentioned.[6] Participants in wave 2 had an official letter from the local DI offering general support from the DI office, which also discussed the possibility to provide an adjustment grant including detailed information about maximum length and amount. The notification of the subsidy was therefore more homogenous and likely to be more credible in wave 2.

To isolate the effect of the subsidy from the effect of the offered support, an alternative letter offered only support from the local DI office (identical wording, but without the paragraph

[6] About a third of the participants had only a vague confirmation from their local DI office. In these cases the phrasing of the paragraph was adapted accordingly.

discussing the adjustment grant).[7] In wave 2, we thus have three different treatments: (1) applications without any letter from the DI office; (2) applications with a letter offering both, general support and the subsidy; and (3) applications with a letter which offers general support only. The first and second treatment provided similar information to the potential employer as the treatments in wave 1.[8]

TABLE 2: Study design

	Wave 1	Wave 2
Sample population	People at the end of a sheltered dual track vocational education and training program	Clients of job coaching services
Notification of the subsidy	Within the application letter	As a separate official letter from the DI office
Application disclosing health impairment but not the subsidy	No	Yes
Application types	Applications to advertised vacancies only or unsolicited applications only or both	Unsolicited applications as well as applications to advertised vacancies
Location of application workshops	Rehabilitation center	University of St. Gallen
Time frame	May-July 2011	Ongoing since January 2012
Geographical region	Appenzell, Lucerne, St.Gallen, Thurgau	St.Gallen

When analyzing the results of the first wave (which was conducted roughly 6-12 months before wave 2), we realized effect heterogeneities according to different application types:

[7] The letter offering general support to an employer was the standard procedure in this local DI office to back-up applications.

[8] In group 1, participants would have had to disclose their health impairment within the application letter to allow performing a similar analysis. We discussed this possibility with the case workers at the rehabilitation centers as well as with potential participants. Most people did not feel comfortable to disclose a disability within the letter without mentioning the subsidy or any other support. We thus decided to refrain from such a proceeding in wave 1.

applications to advertised job offers and unsolicited applications (see section 5). In this group, however, many participants could not identify a sufficient number of suitable advertised vacancies and therefore relied on unsolicited applications (a quite typical form of application in Switzerland). The origin of this heterogeneity is thus unclear: either participants who wrote unsolicited applications were different from other participants or the subsidy had a different effect in unsolicited applications. To analyze the origin of this heterogeneity, we asked every participant in wave 2 to write both types of applications (unsolicited and to advertised vacancies).

There were furthermore minor differences with respect to time frame and geographic coverage. We will acknowledge for these differences by providing pooled as well as stratified results (see section 5).

4 Data and descriptive statistics

So far, 51 individuals participated in our experiment, 39 in wave 1 and 13 in wave 2.[9] In total, our experiment includes 384 applications, 233 from wave 1 (about six on average per person) and 151 (about twelve on average per person) from wave 2 (see Table 3). The number of applications is not balanced between the treatments in wave 1 since many participants wrote an odd number of applications. By chance, the last application was more often assigned to include the notification. A notification of general support was included in applications of wave 2 only (see section 3 for details).

All participants received an answer sheet to document the calendar date of a company reaction (invitation for a job interview or for other reasons). These sheets were returned seven weeks (six weeks in wave 2) after the workshop. Although the results are right-censored, we

[9] Note that wave 2 is still ongoing. In total, we expect 20 participants in wave 2.

believe it is highly unlikely that a company would react in any way (especially positively) to an application after more than six weeks. In fact, our results show participants received hardly any response after five weeks (see the survival functions in Figure A1 in the appendix). There is no significant difference in the timing of response between the treatments.

TABLE 3: Number of applications sent by wave, treatment and type of application

	Number of applications by type	Without notification of subsidy	With notification of subsidy	With notification of general support	Total
Total	Regular	106 (8)	104 (10)	33 (1)	243 (19)
	Unsolicited	57 (4)	68 (1)	16 (0)	141 (5)
	Total	163 (12)	172 (11)	49 (1)	384 (24)
Wave 1	Regular	72 (6)	68 (6)	-	140 (12)
	Unsolicited	41 (4)	52 (0)	-	93 (4)
	Total	113 (10)	120 (6)	-	233 (16)
Wave 2	Regular	34 (2)	36 (4)	33 (1)	103 (7)
	Unsolicited	16 (0)	16 (1)	16 (0)	48 (1)
	Total	50 (2)	52 (5)	49 (1)	151 (8)

Notes: Number of successful applications in parentheses.

From a total of 384 applications, 24 (6.25%) resulted in a call-back for a job interview (see Table 3). There is a slightly lower share of successful applications when the subsidy is notified (6.4%) compared to the situation where the subsidy is not notified (7.4%). This seems to be primarily driven by wave 1, whereas the reverse is true in wave 2. One should notice, however, that wave 2 is based on a relatively low number of applications and that the experiment is still ongoing. As expected, the success rate is lowest for applications that include a letter with the general support from the DI only (2%).

Background information on the participants is available from the application dossiers (which consists of a cover letter, the CV and diplomas from previous work experience and education), and from a written questionnaire filled out at the beginning of the application workshop. Information from the application dossiers is exactly the same which the employer

receives. The questionnaire includes questions on the values participants have for future work opportunities, as well as questions about several psychological concepts like self-esteem (Rosenberg, 1979), self-efficacy (Schwarzer, 2000; Schwarzer and Jerusalem, 1999), social support (Schulz and Schwarzer, 2003), positive attitude towards life (Grob *et al.*, 1991), and the Big Five Inventory of which we included a short version (Rammstedt and John, 2007). This information may be determinants of how an application is written and may thus impact the success of an application as well as the effectiveness of our different treatments.

Descriptive statistics of this background information are provided in the appendix (Table A1). This table compares participants from wave 1 and 2 to better understand if differences in the treatment effects could be grounded in differences of these background characteristics. Participants of wave 2 are older (and consequently have more work experience), more likely to be married and less often Swiss and German native speaker. They are far more likely to have completed compulsory school in regular time and are less likely to have visited any kind of special school. This indicates that they have acquired their disability after they completed their formal education. Vocational education is lower for participants of wave 2 by construction, as only two thirds have completed an apprenticeship. Their CV is also more likely to have gaps, especially unexplained ones.[10] Regarding the importance of job characteristics, the only major difference is that participants of wave 2 are more likely to search for part-time positions. Personality traits are fairly similar across the two waves, with major differences only among positive attitude towards life, self-responsibility, and conscientiousness. We also asked participants on their perceptions of their own disability or health status. Surprisingly, only every third participant of wave 1 perceives herself as being disabled or health impaired, while more than 92% of the participants of wave 2 declared a disability. In wave 1, we did not

[10] A gap in a CV is defined as at least a month between two different jobs or educations, i.e. absence from employment or education. A gap is explained if it is mentioned in the CV what has been done in this period (e.g. looking for work, taking care of children).

ask about the type of disability but from our observation at the workshops, most were either mentally or learning disabled. In wave 2, most (92%) participants were physically disabled.

One of the prime concerns may be the low number of participants and whether or not they are representative for the target group. In the first wave, selection is more relevant on the level of rehabilitation centers with four centers participating in our experiment but less on the individual level. Between the two biggest centers (13 and 19 participants),[11] we find some variation regarding marital status and gender, but relatively similar attitudes towards the future job and personality traits. Significant differences can only be found with respect to self-esteem and positive attitude towards life (Table A2 in the appendix). We furthermore compare our sample with a representative sample of adolescents who completed their vocational training in the regular labor market (column 3 of Table A1).[12] Even though some differences in basic demographic variables exist (participants of wave 1 are older and completed a shorter VET program than their peers), both groups are remarkably similar in their wishes and hopes with respect to future employment as well as in their personality traits. The only major difference is that our participants have a somewhat less positive attitude towards life.

Selection into our experiment seems to be more of an issue in the second wave. It is not possible to compare our sample with a general sample of people who are eligible for subsidies since only disbursement and not eligibility would be coded in any data set. To get at least some idea on representativeness, we compare our sample with a general sample of the working-age population with a chronic illness or a long-term health problem (column 4 of Table

[11] In the other two centers, only 3 and 4 adolescents participated in our study, respectively.

[12] TREE (TRansition from Education to Employment 2008) surveys the post-compulsory educational and labor market pathways of a school leavers' cohort in Switzerland. We selected only those individuals who are in their last year of apprenticeship independent of the wave and age.

A1).[13] Here we find some major differences with respect to personality traits, particularly with respect to extraversion and neuroticism (or emotional stability). One should therefore keep in mind that our results are causal for our sample but may not be representative for the target population.

5 Results

To evaluate the effectiveness of the subsidy on call-backs for job interview invitations, we test on differences between the two sample means of call-back rates (Table 4). We compute the standard errors by a regression of the call-back rate on the treatment dummy (=1 if notification of subsidy is included; =0 if any notification is missing) with clustering on the individual. Overall, we do not find significant differences between applications that were sent with or without the notification of a subsidy. Stratifying our results by wave, however, we find opposite, yet not significant, effects indicating that the subsidy reduces call-back rates in wave 1 and increases them in wave 2 (p-value of a test for whether the difference between the two waves is zero: 0.12). The negative effect in wave 1 is predominantly driven by unsolicited applications (see Table A3 and A4 in the appendix). As been discussed in section 3, some participants could not find suitable offers and therefore wrote only unsolicited applications. Considering only participants who wrote both types of applications (N=13) yields an even higher effect (-0.176, p=0.102).

For wave 2 we do not find similar results. The treatment effects for regular and unsolicited applications are both positive, yet not significant (see Table A3 and A4 in the appendix).

[13] The Swiss Household Panel (SHP) is a representative data set of the Swiss population which aims to observe dynamics of changing living conditions in Switzerland. We selected the working-age population with a chronic illness or a long-term health problem from its latest wave (2009).

TABLE 4: Treatment effects

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	-0.010	-0.038	0.056
No notification	(0.033)	(0.040)	(0.048)
Constant	0.074***	0.088**	0.040
	(0.026)	(0.036)	(0.025)
R^2	<0.000	0.006	0.012
N	335	233	102
B			
Notification of subsidy vs.			0.076*
general support			(0.038)
Constant			0.020
			(0.020)
R^2			0.026
N			101
C			
Notification of general support vs.			-0.020
No notification			(0.020)
Constant			0.040
			(0.025)
R^2			0.003
N			99

Notes: Standard errors in parentheses clustered at the individual. ***,**,* indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Given the relatively low number of observations, this could be simply a power problem.[14] In addition, the design of wave 2 allows isolating the signaling effect from the incentive effect of the subsidy. The latter can be estimated by comparing call-back rates for applications that include the letter from the DI office offering the subsidy with those offering general support only (see panel B of Table 4). The effect is 7.6 percentage points and significant at the 10%

[14] A simple back of the envelope estimation for a two-sample comparison of proportions with n1 = n2 and power equal to 0.9 yields a sample size of 380 for each treatment to find a treatment effect of 0.065 to be significant at the 5% level. This estimation, however, does not consider clustered standard errors at the individual.

level. Assuming that offering general support has only a minor impact on firms' decision to hire a person, the comparison of average call-back rates for applications with the letter providing general support with call-back rates for applications without any letter provides the lower bound for the signaling effect. This effect is relatively small and insignificant (see panel C of Table 4).

When regressing the call-back rate on the treatment dummy we do not take into account how many applications a participant wrote. This number varies considerably between three and 24. With every application the probability of a positive call-back rate increases. We therefore rerun the same regressions as in Table 4 but weighting each observation by the inverse of the product of the number of applications a participant wrote and the number of participants. In this way, an observation receives less weight the more applications a participant wrote. As Table 5 (as well as Table A5 and A6 in the appendix) illustrates, the sizes of the effect do not change qualitatively. They are generally smaller in absolute terms except for regular applications where the effect is much more positive in wave 1.

Given the small sample size of participants, looking for effect heterogeneities with respect to various background characteristics may not be a fruitful endeavor. However, looking at employers characteristics may be more successful since there is more variation with 384 different employers. In fact, the effect of the hiring subsidy seems to be particularly negative for firms in the industry sector (-0.100, clustered SE: 0.052) compared to no effect in companies in the service sector (0.010, clustered SE: 0.036). The p-value of a χ^2-test whether these two effects are the same is 0.058. Thus, if a job requires physical strain, which is usually common in the industry sector, an employer might be particularly deterred by a notification of the disability.

TABLE 5: Treatment effects with weighted observations

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	-0.006	-0.023	0.042
No notification	(0.029)	(0.037)	(0.033)
Constant	0.058***	0.070**	0.023
	(0.021)	(0.028)	(0.016)
R^2	<0.000	0.002	0.010
N	335	233	102
B			
Notification of subsidy vs.			0.053*
general support			(0.029)
Constant			0.011
			(0.011)
R^2			0.019
N			101
C			
Notification of general support vs.			-0.012
No notification			(0.013)
Constant			0.025
			(0.017)
R^2			0.002
N			99

Notes: Standard errors in parentheses clustered at the individual. ***,** indicate statistical significance at the 1 and 5 percent level, respectively.

6 Conclusion

We conducted a small-scale social field experiment to evaluate the effectiveness of a subsidy scheme aimed at increasing employment of people with a disability. Our participants, who were all eligible for this subsidy, wrote several applications where it was randomly decided whether or not the subsidy was disclosed. We measure effectiveness by comparing call-back rates. Our results reveal that the effectiveness strongly depends on the respective target group. While the subsidy is ineffective or even counterproductive in a group of adolescents

who are at the end of their vocational training program, the subsidy is likely to increase call-back rates in a group of clients of job coaching services.

The different treatment effects may be grounded in the difference between the two groups: The first group consists predominantly of young people who acquired their disability before they completed their education and who are just at the end of their vocational training program. For this group, the adjustment grant may not be needed. Disclosing the disability, however, raises the signaling effect. The second group consists of people who typically acquired their disability after they completed their formal education. Due to their health limitation, many of them are forced to change profession. For them, an adjustment grant may be necessary. On the other hand, the signaling effect may be lower since disclosing the disability provides an explanation for gaps in the employment history.

Our study fails to make a general judgment on the overall effectiveness of the subsidy for two reasons: First, we analyze the effectiveness in a single application method only and it is unclear how the subsidy affects success rates of other application methods. There is, however, a considerable risk that in informal applications as well as job coaching activities substantial windfall profits are generated if the firms would be anyway willing to accept a candidate with a handicap. Second, and far more relevant, our sample may not be representative for the general population eligible for adjustment grants. Our results are thus causal for the chosen groups but may not correspond to the average treatment effect.

Particularly in the light of the strong effect heterogeneities we urge policy makers in Switzerland and elsewhere to carefully revise their existing subsidy schemes. Larger-scale studies would be needed to assess the average impact as well as to isolate groups that would particularly benefit/harm from those subsidy schemes.

References

Baumgärtner, M., S. Böhm and D. Dwertmann (2011) *Call Yachol*, mimeo, University of St.Gallen.

Bell, S. H. and L. L. Orr (1994). 'Is Subsidized Employment Cost Effective for Welfare Recipients? Experimental Evidence from Seven State Demonstrations', *The Journal of Human Resources*, vol. **29**(**1**), pp. 42-61.

Bell, S. H. and L. R. Peck (2012) *Obstacles to and Limitations of Social Experiments: 15 False Alarms*, Abt Thought Leadership Paper.

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**(**4**), pp. 991-1013.

Bishop, J. H. and S. Kang (1991). 'Applying for entitlements: Employers and the targeted jobs tax credit', *Journal of Policy Analysis and Management*, vol. **10**(**1**), pp. 24-45.

Burtless, G. (1985). 'Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment', *Industrial and Labor Relations Review*, vol. **39**(**1**), pp. 105-114.

Carling, K. and K. Richardson (2004). 'The relative efficiency of labor market programs: Swedish experience from the 1990s', *Labour Economics*, vol. **11**(**3**), pp. 335-354.

Dubin, J. A. and D. Rivers (1993). 'Experimental estimates of the impact of wage subsidies', *Journal of Econometrics*, vol. **56**(**1-2**), pp. 219-242.

Falk, A., R. Lalive and J. Zweimüller (2005). 'The success of job applications: a new approach to program evaluation', *Labour Economics*, vol. **12**(**6**), pp. 739-748.

Galasso, E., M. Ravallion and A. Salvia (2004). 'Assisting the Transition from Workfare to Work: A Randomized Experiment', *Industrial and Labor Relations Review*, vol. **58**(**1**), pp. 128-142.

Grob, A., R. Lüthi, F. G. Kaiser, A. Flammer, A. Mackinnon and A. J. Wearing (1991). 'Berner Fragebogen zum Wohlbefinden Jugendlicher (BFW)', *Diagnostica*, vol. **37**(**1**), pp. 66-75.

Gupta, N. D. and M. Larsen (2008) *Evaluating Employment Effects of Wage Subsidies for the Disabled – the Danish Flexjobs Scheme*, The Danish National Centre for Social Research.

Gupta, N. D. and M. Larsen (2010) *Evaluating Labour Market Effects of Wage Subsidies for the Disabled – the Danish Flexjob Scheme*, The Danish National Centre for Social Research, Research Department of Employment and Integration, Working Paper 07:2010.

Hamersma, S. (2008). 'The effects of an employer subsidy on employment outcomes: A study of the work opportunity and welfare-to-work tax credits', *Journal of Policy Analysis and Management*, vol. **27**(**3**), pp. 498-520.

Heckman, J. J., R. J. LaLonde and J. A. Smith (1999). 'The Economics and Econometrics of Active Labor Market Programs', in (Ashenfelter, O. and D. Card Eds.), *Handbook of Labour Economics*, pp. 1865–2097, Amsterdam: North-Holland.

Huttunen, K., J. Pirttilä and R. Uusitalo (2010). 'The Employment Effects of Low-Wage Subsidies, ' *CESifo Group Munich, CESifo Working Paper No. 3043*.

Katz, L. (1996). 'Wage Subsidies for the Disadvantaged', *National Bureau of Economic Research Working Paper Series*(**No. 5679**).

Marx, I. (2001). 'Job subsidies and cuts in employers' social security contributions: The verdict of empirical evaluation studies', *International Labour Review*, vol. **140**(**1**), pp. 69-83.

Neubäumer, R. (2010). 'Can Training Programs or Rather Wage Subsidies Bring the Unemployed Back to Work? A Theoretical and Empirical Investigation for Germany', *IZA Discussion Paper Nr. 4864*.

Neumark, D. (2011). 'Policies to Encourage Job Creation: Hiring Credits vs. Worker Subsidies', *National Bureau of Economic Research Working Paper Series*(**No. 16866**).

OECD (2003) *Transforming Disability into Ability*, Paris OECD Publishing.

OECD (2010) *Sickness, Disability and Work: Breaking the Barriers - A Synthesis of Findings across OECD countries*, Paris OECD Publishing.

Oreopoulos, P. (2011). 'Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes', *American Economic Journal: Economic Policy*, vol. **3**(**4**), pp. 148-171.

Rammstedt, B. and O. P. John (2007). 'Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German', *Journal of Research in Personality*, vol. **41**(**1**), pp. 203-212.

Riach, P. A. and J. Rich (2004). 'Deceptive Field Experiments of Discrimination: Are they Ethical?', *Kyklos*, vol. **57**(**3**), pp. 457-470.

Rosenberg, M. (1979). *Conceiving the self,* New York: Basic Books.

Schulz, U. and R. Schwarzer (2003). 'Soziale Unterstützung bei der Krankheitsbewältigung: Die Berliner Social Support Skalen (BSSS)', *Diagnostica*, vol. **49**(**2**), pp. 73-82.

Schünemann, B., M. Lechner and C. Wunsch (2011). 'Do Long-term Unemployed Workers Benefit from Targeted Wage Subsidies?', *Discussion Paper 2011-26, University of St. Gallen, Department of Economics*.

Schwarzer, R. (2000). 'General Perceived Self-Efficacy in 14 Cultures', in (Editor Ed.)^Eds.), *Book General Perceived Self-Efficacy in 14 Cultures*, City.

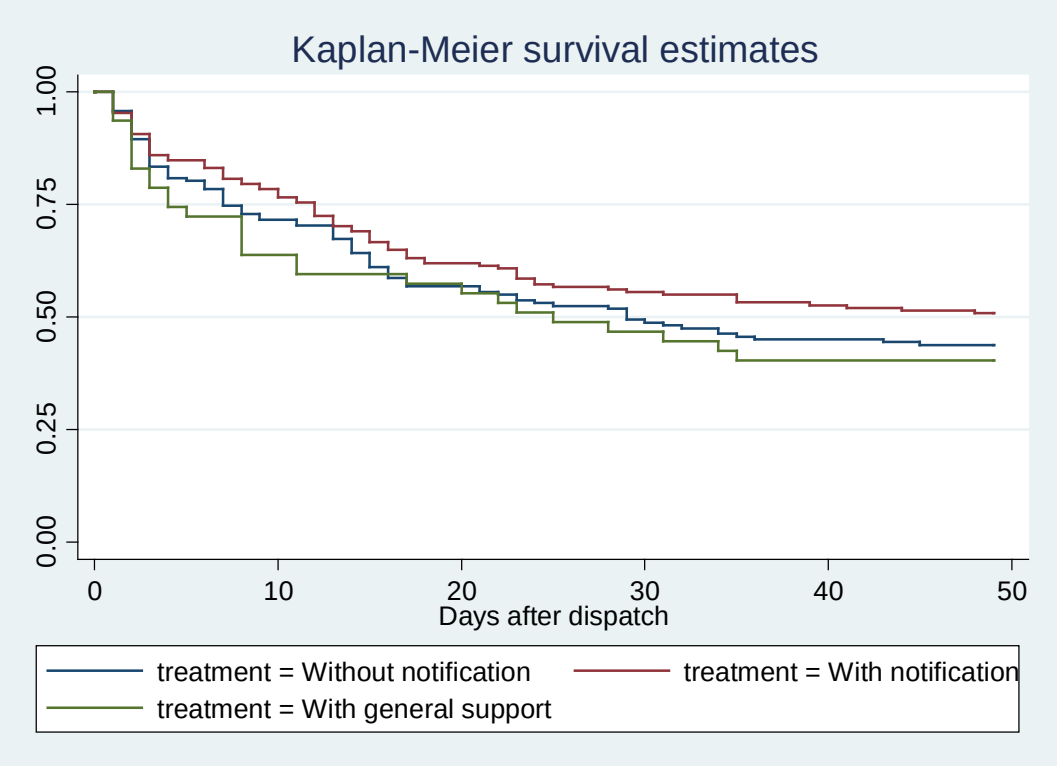
Schwarzer, R. and M. Jerusalem (eds.) (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*, Berlin.

Sianesi, B. (2008). 'Differential effects of active labour market programs for the unemployed', *Labour Economics*, vol. **15**(**3**), pp. 370-399.

Woodbury, S. A. and R. G. Spiegelman (1987). 'Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois', *The American Economic Review*, vol. **77**(**4**), pp. 513-530.

Appendix

FIGURE A1: Kaplan-Meier survival functions



Notes: Failure is equal to receiving a rejection letter. The p-value of the log-rank test, a nonparametric test for whether the survival functions are the same, is 0.247.

TABLE A1: Descriptive statistics

	Wave 1		Wave 2		TREE (N=1750)	SHP (N=1810)
Demographics observable from CV	N	Mean	N	Mean	Mean	Mean
Male	39	0.49	13	0.53	0.48	0.43
Age	39	21.53	13	44.90***	19.55**	46.78
Married	32	0.06	13	0.38***	0.01***	0.58
Swiss	34	0.88	12	0.58**	0.81	0.93†††
German language	32	0.88	13	0.69	0.50***	0.74
Non-german family name	39	0.33	13	0.46		
Work & education history						
Finished apprenticeship	39	0.97	13	0.69***		
Duration of apprenticeship in years	38	2.13	9	3.11***	3.14***	
Finished school in regular time	39	0.52	10	1***		
Has ever visited special school	39	0.21	13	0*		
Years of work experience in sustainable employment	39	1.30	13	20.96***		
Gaps in CV	39	0.36	13	0.85***		
Unexplained gaps in CV (in months)	39	5.59	13	9.85		
Explained gaps in CV (in months)	39	1.51	13	32***		
Learned profession	39		13			
Industrial sector		0.67		0.69		
Service sector		0.33		0.31		
Importance of job characteristics						
Short commute to work	39	2.64	13	2.69	2.62	
Working part-time	38	2.16	13	2.85*	2.08	
Job security	39	3.69	13	3.61	3.62	
Able to learn something new	38	3.34	13	3.15	3.36	
Personality traits						
Self-efficacy	37	2.88	13	2.96	3.00	
Social Support	39	3.41	13	3.31	3.60	
Self esteem	38	4.02	13	3.90	4.10	
Positive attitude towards life	37	4.41	13	3.92**	4.66**	
Self-responsibility/locus of control	37	4.89	13	4.31**		
Big Five Personality traits						
Extraversion	36	0.33	13	0.23		1.32††
Agreeableness	37	1.22	13	1.23		1.37
Conscientiousness	38	1.53	13	2.31*		1.94
Neuroticism	38	0.08	13	0.23		-1.04††
Openness	36	1.28	13	0.77		1.12
Perception about disability						
Self-perceived disability or health impairment	34	0.35	13	0.92***		

Type of disability	
Physical	0.92
Mental	0.33
Intellectual	0

Source: own calculations, TREE (2008) and SHP (2009).

Notes: Individuals in the TREE sample are in their last year of their apprenticeship independent of the wave and age. Numbers of observations vary since some questions were not included in all waves. The SHP sample includes the disability insured population (age 18-65). Variables of importance of job characteristics are measured on a 4-point Likert scale, ranging from 1=totally subordinate to 4=very important. All personal traits are average values of agreements to statements (except for social support, which is composed of only one statement). Self-efficacy and positive attitude towards life have three components, while self-esteem has four. Self-efficacy and social support are measured on a 4-point Likert scale ranging from 1=not at all true to 4=exactly true. Self-esteem is measured on a 5-point Likert scale ranging from 1=not at all accurate to 5=very accurate. Positive attitude towards life is measured on a 6-point Likert scale ranging from 1=totally wrong to 6=totally right. The score of every Big 5 personality trait is the sum of two statements where one is positively phrased and the other negatively. The agreement on the statement is scaled on a 5-point Likert scale, ranging from 1=completely disagree to 5=completely agree. The score of the negatively phrased statement is reversed, so the total score can lie in the spectrum of [-4, 4]. Self-responsibility/locus of control is measured on a 6-point Likert scale ranging from 1=totally wrong to 6=totally right and is the average of two questions (see Baumgärtner *et al.*, 2011). Tests for significant differences are conducted with Wilcoxon rank-sum tests. ***,**,* statistically different from wave 1 at 1, 5 and 10 percent, respectively.†††,††: statistically different from wave 2 at 1 and 5 percent, respectively.

TABLE A2: Comparison of descriptive statistics from two largest participating rehabilitation centers

Characteristic	Center 1	Center 2
Demographics observable from CV		
Male	0.684	0.231**
Age	21.106	22.467
Married	0	0.286**
Swiss	0.824	1
German language	0.778	1
Duration of apprenticeship in years	2.105	2.083
Finished school in regular time	0.474	0.385
Has ever visited special school	0.158	0.308
Years of work experience in sustainable employment	0.843	2.282
Importance of job characteristics (4-point scale)		
Short commute to work	2.474	2.769
Working part-time	2.158	1.917
Job security	3.579	3.769
Able to learn something new	3.316	3.500
Personality traits		
Self-efficacy	2.947	2.861
Social Support	3.421	3.308
Self esteem	4.184	3.646*
Positive attitude towards life	4.702	3.939**
Self-responsibility/locus of control	4.861	4.792
Big Five Personality traits		
Extraversion	0.105	0.500
Agreeableness	1.684	1.077
Conscientiousness	1.579	1.154
Neuroticism	0.211	-0.154
Openness	1.474	1.692
Perception about disability		
Self-perceived disability or health impairment	0.333	0.273
N	19	13

Notes: See Table A1 for details.
***,**,* statistically different from wave 1 at 1, 5 and 10 percent, respectively.
††,†: statistically different from wave 2 at 5 and 10 percent, respectively.

TABLE A3: Treatment effects for regular applications

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	0.021	0.005	0.052
No notification	(0.045)	(0.058)	(0.065)
Constant	0.075**	0.083**	0.059
	(0.030)	(0.042)	(0.037)
R^2	0.001	<0.000	0.009
N	210	140	70
B			
Notification of subsidy vs.			0.081
general support			(0.049)
Constant			0.030
			(0.029)
R^2			0.024
N			69
C			
Notification of general support vs.			-0.029
No notification			(0.029)
Constant			0.059
			(0.037)
R^2			0.005
N			67

Notes: Standard errors in parentheses clustered at the individual. ** indicate statistical significance at the 5 percent level.

TABLE A4: Treatment effects for unsolicited applications

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	-0.056	-0.098*	0.063
No notification	(0.043)	(0.054)	(0.060)
Constant	0.070*	0.098*	0
	(0.040)	(0.054)	-
R^2	0.020	0.057	0.032
N	125	93	32
B			
Notification of subsidy vs.			0.063
general support			(0.061)
Constant			0
			-
R^2			0.032
N			32
C			
Notification of general support vs.			0
No notification			-
Constant			0
			-
R^2			-
N			32

Notes: Standard errors in parentheses clustered at the individual. * indicates statistical significance at the 10 percent level.

TABLE A5: Treatment effects for regular applications with weighted observations

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	0.045	0.055	0.025
No notification	(0.040)	(0.055)	(0.036)
Constant	0.059*	0.070*	0.030
	(0.030)	(0.041)	(0.021)
R^2	0.007	0.009	0.004
N	209	139	70
B			
Notification of subsidy vs.			0.037
general support			(0.026)
Constant			0.014
			(0.014)
R^2			0.010
N			69
C			
Notification of general support vs.			-0.017
No notification			(0.019)
Constant			0.033
			(0.023)
R^2			0.003
N			67

Notes: Standard errors in parentheses clustered at the individual. * indicates statistical significance at the 10 percent level.

TABLE A6: Treatment effects for unsolicited applications with weighted observations

Dependent variable: Call back rate	Full sample	Wave 1	Wave 2
A			
Notification of subsidy vs.	-0.034	-0.096	0.038
No notification	(0.035)	(0.056)	(0.039)
Constant	0.051	0.096	0
	(0.030)	(0.056)	-
R^2	0.009	0.054	0.021
N	119	87	32
B			
Notification of subsidy vs.			0.038
general support			(0.039)
Constant			0
			-
R^2			0.020
N			32
C			
Notification of general support vs.			0
No notification			-
Constant			0
			-
R^2			-
N			32

Notes: Standard errors in parentheses clustered at the individual.