Disability, Education and Work: Empirical Evaluations of Policies to Sustain the Employment and Education of People with a Disability

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Lukas Kauer

from

Dürrenroth (Bern)

Approved on the application of

Prof. Dr. Michael Lechner

and

Prof. Per Johansson, PhD

Dissertation no. 4256

Difo-Druck GmbH, Bamberg

Disability, Education and Work: Empirical Evaluations of Policies to Sustain the Employment and Education of People with a Disability

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Lukas Kauer

from

Dürrenroth (Bern)

Approved on the application of

Prof. Dr. Michael Lechner

and

Prof. Per Johansson, PhD

Dissertation no. 4256

Difo-Druck GmbH, Bamberg

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, November 18, 2013

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

I dedicate this work to my wife, Simone, my parents, Max and Ingrid, and my sister, Barbara, with her family, Andreas, Corrie, and Nathanael. They all deserve special thanks for their sacrifice and support to make my education possible as well as for providing encouragement and desired distraction from this work.

This work would equally not exist without Eva Deuchert, who first believed in my abilities by hiring me. She then inspired much of this work with her invaluable creativity. Through the years I was able to learn and benefit a lot from her. Her guidance improved my work substantially.

I would like to express my gratitude towards my supervisor, Michael Lechner and the other members of my thesis committee, Per Johansson and Reto Föllmi for valuable comments.

Special thanks are due to Helge Liebert, more than my office mate, who was always ready to help, not only on technical issues. Beatrix Eugster provided me with valuable insights and recommendations, which also deserve gratitude. I thank the other colleagues of the Center for Disability and Integration - David Dwertmann, Kirill Bourovoi, Miriam Baumgärtner, Nils Jent, Regula Dietsche, Stephan Böhm, Ursula Würmli - for a relaxing and supportive working atmosphere. Furthermore, I have also greatly benefited from my colleagues at the Swiss Institute for Empirical Economic Research, particularly Andreas Steinmayr, Christina Felfe, Conny Wunsch, Giovanni Mellace, Hans Fricke, Martin Huber, and Petra Thiemann.

I am also indebted to my co-author Flurina Meisen Zannol. The cooperation with her proved to be a success for interdisciplinary research. It has not only rendered the researchers more inspired but also with a sustainable friendship. Carl Wuppermann, my other co-author, has impressed me with his cleverness from which this work has benefited. Sofie Cabus, who shortly shared an office with me, then left to pursue her career at another university, has continued to be a friend for sharing the swayings of academic work. Mark Pletscher has also accompanied my journey and I am grateful for the companionship and discussions.

My gratitude belongs to Joachim Schoss, the founder of the Center of Disability and Integration, who generously supported this work with his funds. Finally, I gratefully acknowledge financial support from the Swiss National Science Foundation, the Emil Zaugg Fund and the Ernst Göhner Foundation.

St. Gallen, November 2013 Lukas Kauer

Contents

List of Figures

List of Tables

Abstract

There are many reasons why people with a disability have both a lower employment rate and lower levels of education than their non-disabled counterparts. This thesis addresses policies which have been designed to sustain the employment and education of people with a disability. A priori, the most obvious reason for these lower rates may be the limitation of the capacity to work or study by a disability. However, part of the gap may also be due to a market failure. Information asymmetries cause frictions in the alignment of interests between the parties involved. Employers may have incomplete information about the work ability of a person with a disability. All four chapters of this dissertation deal in a way with imperfect information. The second and third chapter are concerned with the relationship between the employer and the worker. Chapter 2 analyzes the willingness and barriers of employers to train a person with a mental disorder in a new approach. Chapter 3 evaluates the effectiveness of a hiring subsidy for people with a disability. The fourth chapter deals with the relationship between the insurer and the insurant. It analyzes whether a partial benefit cut affects labor supply of married couples. The last chapter focuses on access to higher education for people with different disabilities and the relationship between the applicant and the student advisor. Note that all chapters are self-contained and can be read independently of each other.

Zusammenfassung

Es gibt viele Gründe warum Menschen mit einer Behinderung sowohl eine tiefere Beschäftigungsquote als auch ein niedrigeres Ausbildungsniveau ausweisen als Menschen ohne Behinderung. Diese Dissertation befasst sich mit Massnahmen, welche entwickelt wurden um die Beschäftigung und Ausbildung von Menschen mit einer Behinderung zu unterstützen. A priori ist wohl der offensichtlichste Grund für diese tieferen Quoten die durch die Behinderung eingeschränkte Fähigkeit zu arbeiten oder zu studieren. Einen Teil des Unterschieds könnte aber auch von Marktversagen herrühren. Informationsasymmetrien verursachen Friktionen in der Abstimmung von Interessen der involvierten Parteien. Arbeitgeber können unvollständige Informationen über die Arbeitsfähigkeit einer Person mit einer Behinderung haben. Alle vier Kapitel dieser Dissertation befassen sich auf gewisse Weise mit unvollständiger Information. Das zweite und dritte Kapitel befassen sich mit der Beziehung zwischen Arbeitgeber und Arbeitnehmer. Kapitel 2 analysiert die Bereitschaft und die Hindernisse von Arbeitgebern Menschen mit einer psychischen Beeinträchtigung in einem neuen Ansatz auszubilden. Kapitel 3 evaluiert die Wirksamkeit eines Einarbeitungszuschusses für Menschen mit Behinderung. Das vierte Kapitel setzt sich mit der Beziehung zwischen Versicherer und Versicherten auseinander. Es wird analysiert ob sich eine teilweise Kürzung der Versicherungsleistungen auf das Arbeitsangebot verheirateter Paare auswirkt. Das letzte Kapitel konzentriert sich auf den Zugang zu höherer Bildung von Menschen mit unterschiedlicher Art der Behinderung und auf die Beziehung zwischen Interessent und Studienberatung. Alle Kapitel sind in sich abgeschlossen und können unabhängig voneinander gelesen werden.

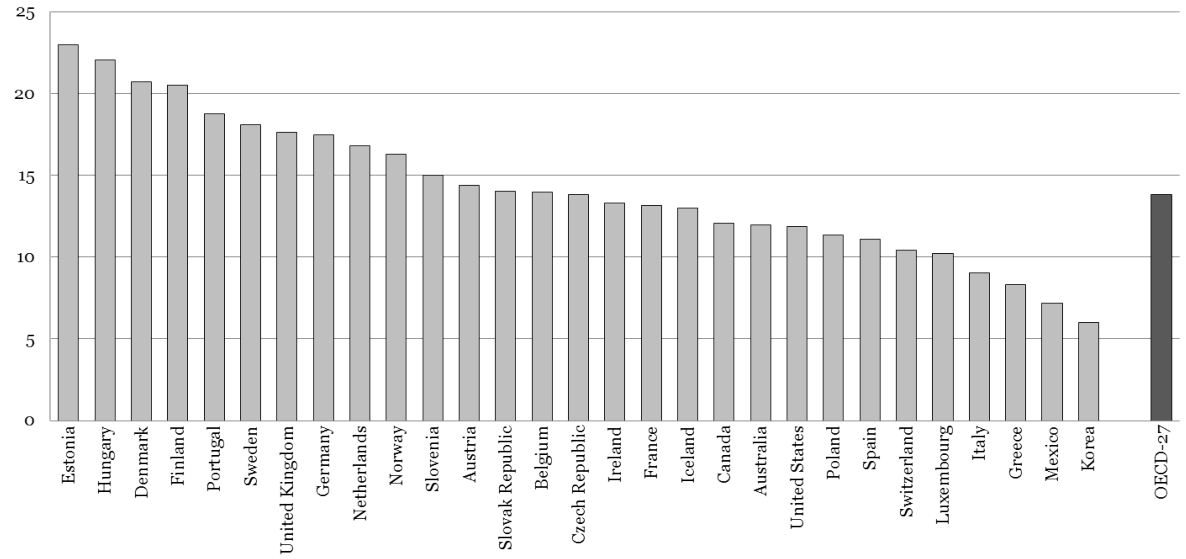
1 Introduction

The United Nations call people with a disability the world's largest minority (UN, 2006). Despite the methodological challenges to measure disability and to compare the measures between different countries (Bound, 1991; Kapteyn et al., 2007), the numbers in Figure 1.1 intend to give a broad outline of the prevalence of disability in various member countries of the Organisation for Economic Co-operation and Development (OECD). While the measure of people who report to have a disability as a percentage of the working-age population varies from 23% in Estonia to 6% in Korea, the OECD average is about 14%.

Looking at employment among people with a disability, Figure 1.2 illustrates their employment rate is on average across the OECD countries just above 40%, which is only slightly more than half of the rate for people without a disability. If people with a disability work, they are more likely to do so part-time. Other employment characteristics generally differ little by disability status. People with a disability tend to be self-employed as frequently as people without a disability; they work to a similar degree in temporary jobs; and they seem to be represented relatively equally in the public sector as well as across industries (OECD, 2010).

The high number of non-working people with a disability has several implications on different levels. On the individual level, income of people with a disability is 15% lower than the national average. This leads to a significantly higher risk of relative income poverty among working-age households with a person with a disability (OECD, 2010). There are also non-monetary losses as life satisfaction and social inclusion are lower when having no job (Kapteyn et al., 2009). On the company level, employers lose valuable human capital since the onset of disability for most people is not at birth or in childhood but at some time during their working age (OECD, 2010). On the societal level, considerable costs are borne because on average 6% of the OECD working-age population was on disability benefits in 2007, a figure of similar magnitude to the average OECD unemployment rate. OECD countries spend 1.2% of GDP on disability benefits, which is almost 2.5 times as much as what is spent on unemployment benefits (OECD, 2009). Despite these economically relevant facts, this topic has not caught much attention in the

Figure 1.1: Self-assessed disability prevalence as a percentage of the working-age population



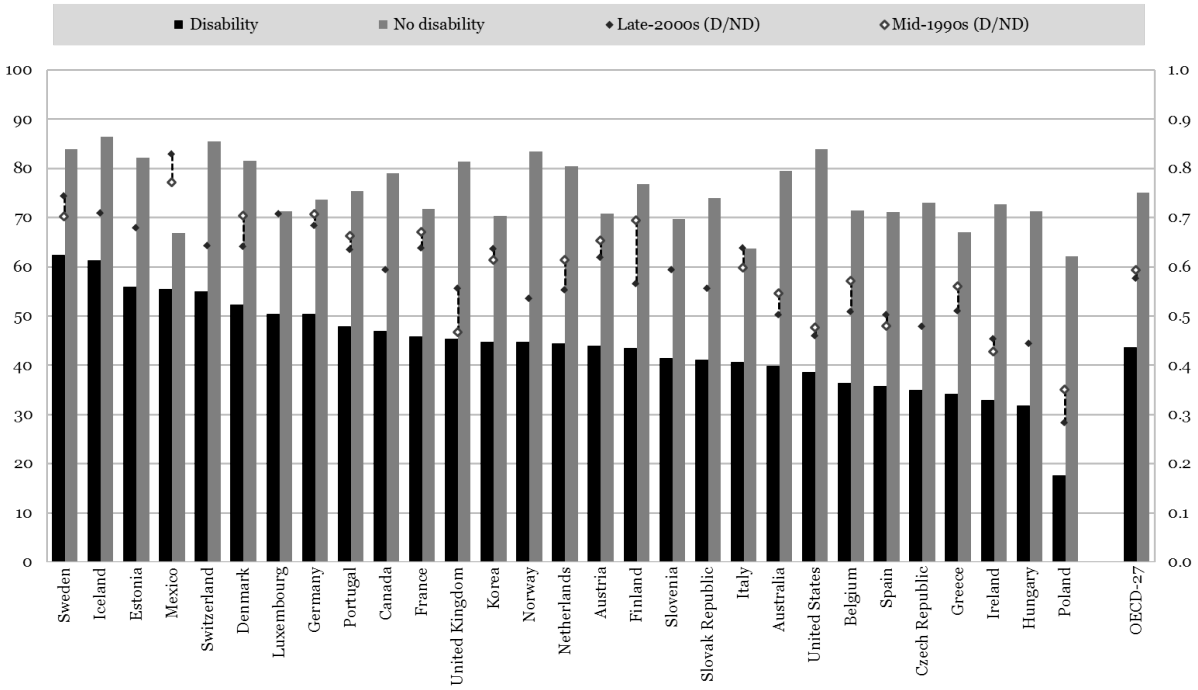
Notes: Working-age population is defined as persons aged 20-64. Data are from national surveys from the late-2000s.
Source: OECD (2010)

economic literature. Using search results in EconLit, the electronic bibliography of the American Economic Association, as an arguably somewhat crude measure of relevance of a topic, and again comparing disability to unemployment, the bibliography returns 2,313 entries for the term *disability* compared to 38,421 entries for the term *unemployment* (date of access: July 12, 2013).

People with a disability not only have a lower employment rate but also significantly lower levels of education than their non-disabled counterparts (OECD, 2010). As Figure 1.3 illustrates the share of persons with a disability with less than upper secondary education is on average across OECD countries almost twice the share of those without a disability. Again the implications are similar to those from a lower employment rate as education and work are obviously intertwined. Higher education enables reaching higher paid jobs and therefore protects from income poverty. On the other side, employers may not be able to hire as much skilled labor as demanded. Finally, the society may again be borne by increased social welfare spendings.

In this dissertation I explore on the reasons why people with a disability have both a lower employment rate and lower levels of education than their non-disabled counterparts. The chapters address policies which have been designed to sustain the employment and

Figure 1.2: Employment rates by disability status

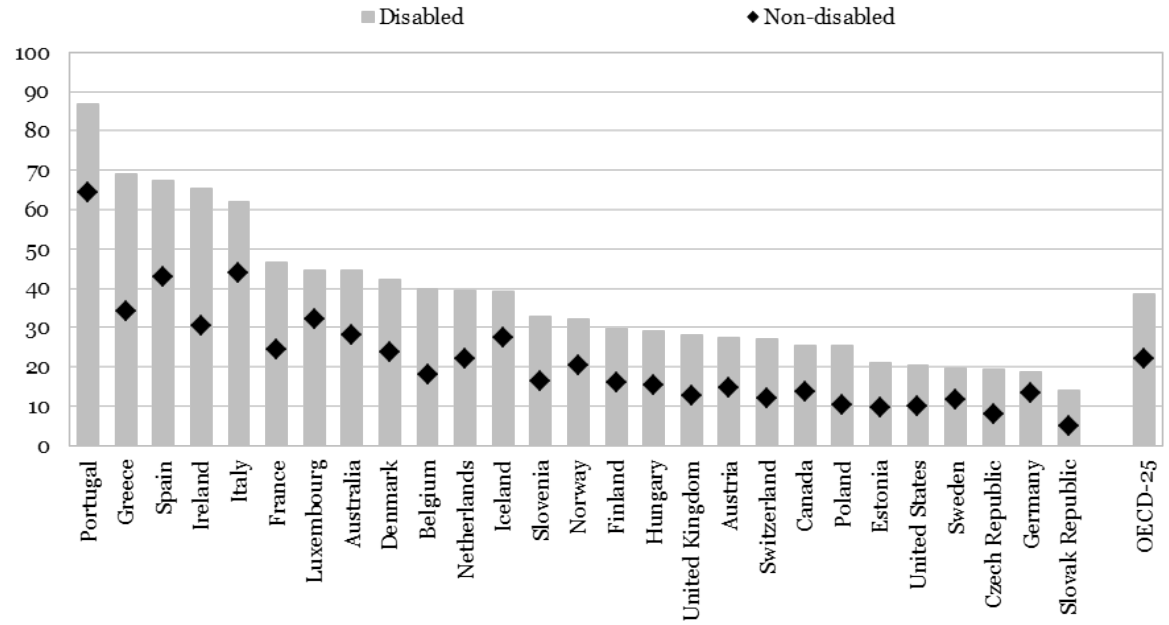


Notes: Employment rates by disability status in the late-2000s (left axis) and trends in relative employment rates since the mid-1990s (people with a disability over those without, right axis). D/ND refers to the employment rate of people with a disability relative to the employment rate of those without a disability.
Source: OECD (2010)

education of people with a disability. A priori, the most obvious reason for these lower rates may be the limitation of the capacity to work or study by a disability. However, part of the gap may also be due to a market failure. Information asymmetries cause frictions in the alignment of interests between the parties involved. Employers may have incomplete information about the work ability of a person with a disability. All four remaining chapters of this dissertation deal in a way with imperfect information. In the second and third chapter I look at the relationship between the employer and the worker. The fourth chapter deals with the relationship between the insurer and the insurant. The last chapter focuses on access to higher education for people with different disabilities and the relationship between the applicant and the student advisor. Note that all chapters are self-contained and can be read independently of each other.

At the beginning of the hiring process, information is incomplete between an employer and an applicant. To avoid hiring inferior workers, employers seek to gain as many

Figure 1.3: Share of the working-age population with low education level by disability status



Notes: A low education level corresponds to an educational attainment of less than upper secondary (ISCED 0-2).
Source: OECD (2010)

signals as possible on the suitability of the applicant. The indication of a disability may be viewed by employers as a sign of inferiority of the applicant. Chapter 2 analyzes this reservation by estimating the willingness of employers to train a person with a mental disorder. The study of interventions trying to increase the employment rate of persons with a mental illness deserves a particular focus as mental illness is nowadays the prime reason for the inflow into disability insurance in many countries. Getting people with a mental illness back to work is therefore necessary in order to ease the pressure off social welfare. The training takes place in a new framework - Supported Vocational Education and Training (S-VET) - which has been proven to be more effective in increasing sustainable employment for people with a mental disorder. Unlike the traditional framework, where persons with a disability are first trained in a sheltered rehabilitation environment and then placed in the competitive labor market, S-VET reverses this process by first placing the client in the competitive market and then helping the client and employer with whatever support is needed. Before scaling up S-VET and

implement it nationwide, it is important to know the willingness to participate in it from the employers' perspective as well as their potential barriers to do so.

All employers in the commercial sector in Eastern Switzerland who offer training in the most popular apprenticeship were invited to participate in an online survey. In the survey, persons responsible for the training and/or hiring are confronted with competing hypothetical, but realistic cases of applicants, which differ in levels of several attributes. They are asked whether they would train this applicant or not. Due to the forced trade-off between attributes, the enforced choice reveals its relevance and its relative importance to the other attributes. This method is known as discrete choice experiment and heavily used in marketing, health economics and environmental economics to value goods. Its application to labor economics is fairly new.

While 22% of the applicants are accepted, it is clear that this number of stated preference is inflated due to the hypothetical situation. Using a well-established method to correct this hypothetical bias by asking a follow-up question how sure they really are, overall acceptance drops to 8.6%. With a survey response rate of 35%, it could be even as low as 3% assuming non-respondents are not interested in hiring people with a mental illness. The main deterrents to train a person with a mental illness are identified to be deficiencies in non-cognitive skills, e.g. adherence to regulation, contact with others. These skills are not revealed in an usual application process but are often limited in persons with a mental disorder. Testing for stigmatization of particular diagnoses by comparing the acceptance of profiles which discloses a diagnosis with profiles without a diagnosis reveals that patients with psychotic disorders are particularly disadvantaged.

In light of this study, S-VET is unlikely to be a magic bullet even though employers do not bear any direct costs. For a broad implementation of these programs additional incentives, in form of subsidies or legal requirements, may need to be considered.

While subsidies might appear as an attractive intervention, they need to be carefully evaluated. Hiring subsidies are in fact already an exceedingly popular intervention in many industrialized countries to encourage the provision of employment to workers with a disability. They have been introduced in hope of many firms employing a person with a disability due to the financial incentive, while this hiring would fall through in the absence of the subsidy. The subsidy should then reduce the information asymmetry so that the employer realizes during the subsidy period that there is in fact no inferiority of the worker with a disability compared to non-disabled co-workers. However, it is important to keep in mind three other possibilities which hamper the effectiveness of the

subsidy. (i) If a company does not employ a worker with a disability even if a subsidy is available, the subsidy is not effective. At least, there are no costs for the government since no subsidy is paid. (ii) The next case are employers who would hire a worker with a disability in any case whether they receive a subsidy or not. This would lead to a deadweight loss as the subsidy is not necessary. (iii) The last case represent employers who would hire a disabled worker without a subsidy but not with one. At first sight, this possibility might seem puzzling. But since the notification of the eligibility to a subsidy also transports the information that the person has a disability might deter the employer from hiring this person. In the absence of this notification, the employer may not receive this signal and hire the person.

Chapter 3 evaluates a hiring subsidy in Switzerland using a small scale field experiment with a special design. Since the policy has already been in place since January 2008, we cannot implement a standard *voucher* experiment assigning eligibility randomly. Instead, it was randomized whether an application included a notification of the subsidy or not. The main advantage of this design is that the treatment is not withheld from any participant. Denying access to a potentially beneficial treatment is often viewed as unethical by policy makers. Furthermore, this design is not plagued by non-compliance (i.e. applicants do not notify their eligibility to the employer) or substitution bias (i.e. non-eligible applicants receive access to the subsidy anyway). On the other hand, the outcome is restricted to invitations to job interviews. Job offers cannot be taken into account since applicants were allowed to notify their eligibility at the job interview.

The study was conducted with two different populations both eligible for the subsidy: Graduates from sheltered Vocational Education and Training Programs in the first wave and clients from employment consulting services, who developed an impairment during their working life, in the second wave. The effect of the subsidy tends to go in different directions for the two groups. While the young graduates receive more invitations to a job interview when the notification of the subsidy is missing in the application compared to applications which include a notification, the subsidy might be effective for the other group. The young graduates are still in employment when writing their application so that a notification of the subsidy raises the signaling effect. Clients from employment counseling services often cannot return to their former profession. For them the subsidy may help to bridge the adjustment period and may serve as an explanation for gaps in the employment history. Policy makers should therefore carefully assess whom to target with a hiring subsidy, although larger-scale studies are needed to substantiate this conclusion.

In contrast to the previous chapters, which mainly address labor demand, Chapter 4 switches to labor supply. Disability insurance (DI) is designed to protect individuals from monetary losses in case of limited work ability. Yet, asymmetric information between insurer and claimant leads to moral hazard. The design of DI benefit schemes may create work disincentives that make it not rewarding for DI beneficiaries to look for work. These work disincentives are difficult to quantify. Policies or policy reforms which apply only to part of the population can be used as exogenous variation in the benefit level or in eligibility criteria. In the previous literature, the assessment of the extent of work disincentives in the DI has mainly been limited to the inflow into DI. At the time of application, substantial work disincentives have been established resulting in a change in labor supply. It remains unclear whether existing beneficiaries react similarly when they lose part of their benefits. This is the research question of Chapter 4, in which spillover effects into spousal labor supply and other social insurances are additionally controlled for.

The Swiss DI granted an additional pension (AP) to the spouse for every new DI beneficiary up to January 2004 unconditional on the spouse's labor supply. Beneficiaries who entered DI after December 2003 did not receive any AP, while beneficiaries who entered DI before January 2004 kept receiving an AP until January 2008 when all remaining AP were canceled. The amount of the AP equaled 30% on top of the DI benefits, which is equal to about 5% of total household income. I compare the change in labor supply of those losing the AP by January 2008 with the change in labor supply of an unaffected group, namely married DI beneficiaries who entered DI after December 2003 so that they never received an AP.

Unlike in the literature which analyzes labor supply of DI applicants, I find existing beneficiaries not changing their labor supply when they lose part of their benefits. One of the main reason is that they could lose even more benefits when increasing their labor supply. This is not the case for the spouse, whose labor supply is completely unrelated to the DI benefits of the partner. Even for the spouse, no change in the labor supply can be detected. If anything, DI beneficiaries seem to increase their participation in means-tested supplementary benefits.

Given the rigidity in the outflow, which results in an exceedingly expensive stock of DI beneficiaries putting pressure on public finances, policy makers are advised to continue to focus on the prevention of inflow into DI rather than spending scarce resources on the promotion of the outflow from DI.

While Chapter 2 has focused on vocational education of people with a disability, the final chapter deals with their university education. The UN Convention on the Rights of Persons with Disabilities calls governments to provide access to tertiary education for people with disabilities. In a randomized field experiment we test whether people with different disabilities are treated differently at the time of their first contact with the higher education institution (HEI) in Germany. Fictitious potential students with a disability request information from student counseling offices of the respective HEI by email on the admission process and on special accommodations to ease studying. The disability types included are physical, mental, and learning. The diagnoses are chosen to have a high prevalence in the students' population (chronic kidney failure, depression, dyslexia). As a control we include a non-disabled student, who needs to take care of a sick family member forcing her to miss classes regularly and therefore also asks for special accommodations. There are two main reasons why student counselors might reply to emails from students with different types of disability in a systematic different way: (1) Stigmatization. For example, people with a learning disability may be perceived as less capable or intelligent compared to students with a physical disability. (2) Marginalization. Some of the described health conditions may not be considered as a disability deserving special assistance even though all chosen impairments are considered to be a disability under German law.

We analyze the written replies using text mining tools as keyword searches and content analysis. Despite the heterogeneity of special needs of students with different disabilities students should receive the same kind of information on the admission process. If there was no marginalization, students should also have equal access to special services for students with disabilities. As a test for stigmatization we analyze whether the replies are different with respect to how often the counselor questions the student's study choice by suggesting a different HEI, a different study program or even to delay studying.

The results indicate students with dyslexia or depression are confronted with marginalization and stigmatization in comparison to students with chronic kidney failure. They receive both less often information on special accommodations and are less often provided access to disability counseling. Furthermore, students with depression see their decision questioned more often. These differences are particularly pronounced in study programs for teaching and technical studies, and in HEIs located in states which do not require appointing a representative for the special needs of students with disabilities.

A policy intervention requiring to appoint a representative for disabled students may therefore reduce discrimination of particular disabilities. If our findings translate into

lower access and a lower probability to graduate, already existing disadvantages on the labor market may be reinforced.

2 Would You Train Me with My Mental Illness? - Evidence From a Discrete Choice Experiment

joint with Eva Deuchert, Flurina Meisen Zannol

This chapter has been published in The Journal of Mental Health Policy and Economics, 2013, Volume 16, 67-80.

Abstract

Background: Mental illness is the prime reason for the inflow into disability insurance in many countries. The integration of persons with a disability into the regular labor market is costly and in the case of mentally ill persons, particularly difficult. Supported Education and Employment - a rehabilitation method that directly places patients in a realistic work environment - has been shown to be effective in increasing competitive employment. However, it has not yet been widely implemented.
Aims of the Study: We evaluate ex-ante the willingness to participate in Supported Education and Employment and the barriers to do so from the employer's perspective.
Methods: We conducted a discrete choice experiment implemented in an online survey. The survey was carried out among all Swiss companies which provide standard dual-track vocational education and training for commercial occupations in Eastern Switzerland. We presented respondents (employees who are responsible for vocational training and/or for the selection of applicants) with a sample of five hypothetical profiles. These profiles vary along different medical diagnoses, different illness related (dys-)functions, and other characteristics that may be associated with a company's willingness to accept the candidate (such as school performance, motivation, and illness related absences). Respondents were asked whether or not they would train this person.
Results: 22% of the profiles are accepted. However, our results demonstrate that the hypothetical bias - which is the difference between individual saying what they *would* do in a hypothetical setting and what they *will* do when they have the opportunity - is severe. Correcting for this bias using follow-up scales ("Are you sure?") reduces the overall acceptance in our sample to 9%. Keeping in mind the response rate to our survey

of 35%, overall acceptance may be as low as 3%. Non-cognitive dysfunctions, which are related to mental disorders, are the main deterrents. Patients with psychotic disorders are particularly disadvantaged.
Implications for Health Policy: Although there are no direct costs to the employer, a wide introduction of Supported Vocational Education and Training is likely to fail at the current stage with the unwillingness of companies to train people with special needs. There may be additional incentives needed, for example in form of subsidies or legal requirements. Even though our experiment has been tailored to the Swiss system, our results may also be relevant in other countries with similar dual-track education systems.

2.1 Introduction

Various studies document the limited employment opportunities of people with disabilities (Bound and Waidmann, 2002; Jones et al., 2006; Kreider and Pepper, 2007). On average across the OECD, the employment rate of people with disabilities is 40%, which is half of the rate for people without a disability (OECD, 2009). People with mental disorders are particularly difficult to integrate into the labor market (Alexandre and French, 2001; Chatterji et al., 2009; Cornwell et al., 2009; Ettner et al., 1997; Savoca and Rosenheck, 2000). Despite the insufficiently low labor market integration of people with disabilities in general, and mentally ill persons in particular, there is relatively little research in the economic literature on employment measures for this particular group.

It is well documented in the psychiatric literature that vocational rehabilitation of people with mental disorders works best with a *Supported Education* or *Supported Employment* strategy. Instead of receiving therapeutic services as well as vocational education and training in a protected work or school environment, patients are directly placed in integrated job or education settings and provided with whatever ongoing support is necessary (Corrigan and McCracken, 2005; Mowbray et al., 2005). Various randomized controlled trials document the effectiveness of Supported Education or Employment in increasing sustainable employment (Clark et al., 1998; Crowther et al., 2001).

Our study focuses on Supported Vocational Education and Training (S-VET), which is a combination of Supported Employment and Supported Education. A pilot study conducted in Germany shows that competitive employment rates after completing S-VET more than doubled compared to students who completed their dual track training in sheltered workplaces - the standard vocational training approach for students with disabilities (Seyd et al., 2007).

It is unclear, however, which factors explain this success: One possibility may be that this program increases employability, either by increasing productivity (students receive

their practical training in a realistic rather than a sheltered work environment) or by reducing discrimination/stigma (a potential disability does not indirectly appear in the CV since the training was not completed in a sheltered workplace). On the other hand, it may be the case that the training takes place in "socially responsible" firms and that the student stays with this employer after completing training. Higher employment rates are then simply the result of the early contact with these employers. If this is the driving factor, it will be difficult to replicate the favorable results from these pilot studies once this program is widely rolled out - particularly if the pilot studies were conducted in cooperation with employers already willing to participate in this rehabilitation method (as it was the case in the German pilot study for example (Seyd et al., 2007)). The effectiveness of these programs then depends also on employers' willingness to participate in this program.

Our paper provides an ex-ante evaluation of the general willingness of the labor market to participate in this rehabilitation program. We conduct a discrete choice experiment (DCE) to study the overall acceptance of Supported Vocational Education and Training as well as the main barriers for participation. Using a DCE to ex-ante evaluate the potential for a labor market program is a quite novel approach. To our knowledge there are only four other studies that use the DCE framework to analyze hiring decisions of companies but these studies neither focus on evaluating the potential of a labor market program nor on people with disabilities (Biesma et al., 2007; Norwood and Henneberry, 2006; van Beek et al., 1997; Eriksson et al., 2012). Using hypothetical scenarios to evaluate the employability of persons with a disability compared to persons without a disability is provided by a single study (Bricout and Bentley, 2000). However, the hypothetical applicants only differ in the disability status, while a DCE allows for a higher dimensionality of characteristics.

Our results document the overall willingness to participate in this rehabilitation method is currently relatively low. Non-cognitive dysfunctions (e.g. non-adherence to regulations, difficulties with contacts with others) that are related to mental disorders are the main deterrents. Consequently, despite the fact that Supported Education/Employment programs may be effective from the patients' point of view and cost-saving from the state's perspective (Cimera, 2007), our results suggest that it will be difficult to widely implement this rehabilitation method.

2.2 Institutional Setting

Formalized vocational training programs are common in many industrialized countries. In Switzerland for example, about two-thirds of people who graduate from compulsory education enroll in a Vocational Education and Training (VET) program (BBT, 2010). Dual-track VET programs are the most common form of vocational education and training. In the dual-track approach, students attend courses at vocational schools on a part time basis. The remaining time is spent at a host company where the students are taught the practical skills needed for their chosen occupation. The period of training is usually two to four years and is completed by a state examination.

Providing people with disabilities with vocational education and training is one of the key vocational rehabilitation strategies: First of all, people with disabilities, like other people too, have far better labor market prospects if they have a good education. Thus, education serves as a buffer to (at least partly) protect against the negative effects of disabilities (Hollenbeck and Kimmel, 2008). Second, outflow from the disability insurance (DI) is rare. In Switzerland for example, less than 1% of DI beneficiaries have left the insurance for reasons other than death or retirement (BSV, 2011). Thus, it is unlikely that young people leave the DI rolls once they entered it. Great effort to integrate young people into the regular labor market is therefore justified to avoid high follow-up costs.

Vocational Education and Training for people with a disability often takes place in sheltered work places provided by specialized rehabilitation centers. Young people receive training in practical skills in special vocational training centers, while they join the regular vocational school. These centers provide support and assistance in form of educational and psychological services. The goal of these VET programs is the integration of young people with a disability into the competitive labor market by the end of the program.

The key problem of these specialized programs lies in the fact that the protected work environment can only mimic the competitive work environment. It therefore offers preparation to the economic reality only to a limited extent. This generates a double hurdle for young professionals after completing their training. Graduates do not only need to find employment in the competitive labor market, but also a path from a protected rehabilitation environment into the competitive economic environment.

Supported Vocational Education and Training (S-VET) programs can ease this pathway. Instead of receiving vocational training within the protected work environment for the full training period, students are placed in companies operating in the competitive labor market (for example in form of long-term internships). The rehabilitation provider offers

job coaching services to the student, as well as advice and support to the employer. During the training period, students join the regular vocational school. A replacement into the protected work environment is possible in case the student experiences a relapse. Usually there are no direct costs for the employer because salaries are paid in form of subsidies or directly by a social security institution. In Switzerland, there are currently only few institutions which place candidates for S-VET. The majority of young people with a mental illness are still educated in a sheltered VET program. It is thus interesting to analyze the potential demand for this rehabilitation method from the employer's perspective before it should be widely spread.

2.3 Method

2.3.1 Discrete Choice Experiment

To estimate the acceptance of candidates for S-VET programs, we conduct a discrete choice experiment (DCE). This technique is also known as contingent valuation or conjoint analysis and is used to analyze stated preferences. It is now widely used in different fields of economics (Hoyos, 2010; Louviere et al., 2000), health economics being one of them (de Bekker-Grob et al., 2012). People are confronted with competing hypothetical, but realistic scenarios, which differ in levels of several attributes. Due to the forced trade-off between attributes, the enforced choice reveals its relevance and its relative importance to the other attributes (Louviere et al., 2000).

In our DCE, we present respondents, who are responsible for standard VET services within companies, five different student profiles one after the other and ask them whether they would train the candidate within a S-VET program. Our DCE is thus of dichotomous choice format. The profiles differ along eleven different attributes; each has three to five different levels. Respective attributes are chosen to include the key factors driving the decision to participate in this program. A survey conducted at Swiss employers in 2004 revealed that possible limited functioning is the main deterrent for hiring people with disabilities (Baumgartner et al., 2004). We are therefore interested to understand which disability related functional deficiencies may hinder the placement of trainees in a S-VET program and which other factors may compensate for these deficiencies. This is particularly relevant since the application process for S-VET is often backed up by a Job Coach, who gives employers further information on the candidate, particularly on illness related functionality.

Measuring deficiencies resulting from mental illnesses is extremely challenging. Mental illnesses can affect thinking, feeling, socializing, functioning, or self-care. While the functional deficiency of a paraplegic may be more obvious, this is not the case for mental illnesses as these illnesses affect cognitive but mainly non-cognitive skills.

Economic research has demonstrated that non-cognitive skills have substantial impact on education and labor market outcomes, as well as on health and social behavior (Chiteji, 2010; Heckman and Rubinstein, 2001; Heckman et al., 2006). To measure these skills, this literature generally employs scales developed by psychologists, which are based on a battery of questions aiming at identifying underlying personality traits (such as locus of control or self-efficacy). To generate candidate profiles, however, these measures cannot be used. These scales do not provide a reasonable basis for an employer to judge the employability of a person. Our respondents do not have a psychological education, and therefore cannot interpret the meaning of these scales. We need a far more descriptive measure for illness related (dys)functions.

The *International Classification of Functioning, Disability and Health* (ICF) provides a standard framework for the description of health and health-related states (WHO, 2002). The advantage of the ICF is that it does not rely only on a medical model, but also includes external environmental factors (for example social attitudes, architectural characteristics, legal and social structures, etc.), as well as internal personal factors. The major disadvantage of the ICF is, however, that this framework is very complex and exhaustive because it describes the functional status of a patient with approximately 1,400 items.

For practical use, core sets were developed to reduce the complexity of the ICF. These core sets present a list of ICF categories that are relevant to most patients with a specific health condition. We base our selection of relevant attributes on the Mini-ICF-Rating for Mental Disorders (Linden et al., 2009). This Mini-ICF-Rating provides 13 different functionalities, such as the adherence to regulations or planning, competency or flexibility. This Mini-ICF-Rating seems to be particularly useful for our purpose since work performance (measured by the Endicott Work Productivity Scale) and ICF items are strongly correlated (Linden et al., 2010). The original description of each attribute is rather long. We therefore shortened the description of these items (provided to the respondent) and rephrased them when necessary.

A somehow similar challenge of choosing the appropriate measure of non-cognitive skills applies to the choice of an appropriate measure of cognitive skills. Again, we cannot rely on IQ test scores or any other cognitive ability test score since our participants are

not familiar with these scores. Instead we use school performance in the form of the Grade Point Average, which is a familiar measure of cognitive skills to every trainer and widely used to select apprentices for VET (Imdorf, 2009; Schmid and Storni, 2004).

Trainers may receive additional information on other characteristics, such as age, previous work experience and health of the applicant. As a measure of health, which should be both easily interpretable and relevant for productivity, we selected days of absence due to illness. It was unclear whether the provision of exact medical diagnoses would be relevant information for employers or whether the diagnosis would be further stigmatizing. To test for a possible stigmatizing effect, we provide one group of respondents with a diagnosis, while we do not show this information to the control group. We selected the four diagnoses (bipolar disorder, schizophrenia, eating disorder, and borderline personality disorder) because they not only have a fairly high prevalence, but also because these diseases may be known to the employers, although this information may be incomplete.

To be sure to have selected the relevant attributes for our experiment, we conducted qualitative interviews with companies that had already participated in a S-VET program, as well as two quantitative pre-tests. The first pre-test was conducted among job coaches from vocational rehabilitation providers. The second pre-test was conducted among company representatives who are responsible for standard vocational education and training services, but which company is not located in our target region. Many pre-testers suggested including motivation as an additional attribute. Therefore, we constructed a description and levels of this attribute similar to the ones taken from the Mini-ICF-Rating. The attributes included, a short description of how each attribute is presented in the profile, and the respective levels for each attribute are given in Table A-1 in the appendix. Unlike in common DCEs used to elicit the willingness-to-pay for a good, we do not include a price as an attribute. Employers neither face any direct costs nor receive any financial compensation if taking part in S-VET. Since implausible combinations are ex-ante hard to identify, we did not rule out any of them. Instead we fully randomize the different attribute levels to the profiles making all possible combinations equally likely.

2.3.2 Hypothetical Bias

The main threat to the validity of DCEs is the hypothetical bias. The experiment describes a hypothetical situation with no real consequences. Under a hypothetical scenario responses are usually more positive than under a real scenario (Murphy et al., 2005). There is a substantial literature showing the hypothetical bias is relevant, particularly

in willingness to pay studies (Blumenschein et al., 2008; Özdemir et al., 2009). In our particular setting, we believe that two factors contribute to the hypothetical bias: *(i)* The topic of vocational integration of people with mental disorders may be prone to a social desirability bias - the tendency to reply in a manner that will be viewed favorably by others. *(ii)* The problem of cognitive dissonance arises if respondents are hypothetically willing to train a person with a disability, while other circumstances preclude the implementation in reality.

The hypothetical bias may cause a substantial bias in our estimates for acceptance. In the following we will provide a formalized framework to illustrate how the hypothetical bias affects the results, and how the well-established method of follow-up questioning will be used to solve this problem.

Denote true acceptance with Y^*, which is a binary variable. This true acceptance is not observed. All we observe is the individually reported acceptance Y which may or may not be biased. The probability to accept a particular profile with attribute A at level k is thus equal to:

$$\begin{aligned} Pr(Y = 1|A = k) = Pr(Y = 1|A = k, Y^* = 1)Pr(Y^* = 1|A = k) \\ + Pr(Y = 1|A = k, Y^* = 0)Pr(Y^* = 0|A = k) \end{aligned} \quad (2.1)$$

It is very unlikely that with a true interest in the profile, acceptance would be falsely reported. We therefore assume that respondents who have true interest in the described candidate would always disclose their interest $[Pr(Y = 1|A = k, Y^* = 1) = 1]$. This is essentially a "no-means-no" rule since this assumption also implies that no person exists who declines a profile but has true interest in it $[Pr(Y^* = 1|A = k, Y = 0) = 0]$. False reporting, however, is likely if the respective profile is not acceptable $[Pr(Y = 1|A = k, Y^* = 0) \geq 0]$. The assumption of non-symmetric misreporting is consistent with the idea that the hypothetical bias is grounded in a social desirability bias, which is by definition non-symmetric. It is also consistent with cognitive dissonance (a person would like to train a person with a disability but external circumstances prevent that to happen). The reverse of this bias (a person would not like to train a person with a disability but external circumstances forces her to do so) would only be likely in the case of social enterprises that specialized on training people with disabilities with trainers who are unwilling to provide training to certain groups. Social enterprises, however, are not part of our survey. Finally, this assumption is consistent with previous empirical evidence.

In three DCEs for a private good, individuals were offered a good hypothetically and then for real (Johannesson et al., 1999; Blomquist et al., 2009). Not a single person who declined purchasing the good in the hypothetical setting switched to purchasing the good in the real setting. With non-symmetric misreporting, the last line of equation 2.1 therefore represents the hypothetical bias.

The level of each attribute is independently chosen from each other with equal probability in our experiment. This design allows us to estimate the average impact of these attributes on the individual hiring decision (which we call absolute treatment effect in the rest of the paper) without being confounded by other attributes. The hypothetical bias, however, also affects these treatment effects in the following way:

$$\begin{aligned} Pr(Y=1|A=k) - Pr(Y=1|A=j) \\ = Pr(Y^*=1|A=k) - Pr(Y^*=1|A=j) \\ + Pr(Y=1|A=k,Y^*=0)Pr(Y^*=0|A=k) \\ - Pr(Y=1|A=j,Y^*=0)Pr(Y^*=0|A=j) \end{aligned} \tag{2.2}$$

where the last two lines of equation 2.2 denote the hypothetical bias for the absolute treatment effect. To correct for the hypothetical bias, we can rewrite the true acceptance rate (assuming non-symmetric misreporting):

$$Pr(Y^*=1|A=k) = Pr(Y^*=1|A=k,Y=1)Pr(Y=1|A=k) \tag{2.3}$$

Equation 2.3 shows that we need an estimator for the true probability to accept a candidate conditional on the respondent's statement that the candidate would be accepted. It also shows that the assumption whether false reporting is associated with the attribute level itself is crucial. In case misreporting is not related to the attribute $Pr(Y^*=1|A=k,Y=1) = Pr(Y^*=1|Y=1)$, the absolute treatment effect is biased, while the relative treatment effect is unbiased:

$$\frac{Pr(Y^*=1|A=k)}{Pr(Y^*=1|A=j)} = \frac{Pr(Y=1|A=k)}{Pr(Y=1|A=j)} \tag{2.4}$$

In case misreporting is systematically related to the attribute level, the absolute as well as the relative treatment effect are biased. Furthermore, non-symmetric misreporting as well as misreporting being not related to the attribute lead to true treatment effects that are considerably smaller than the ones estimated from stated acceptance:

$$Pr(Y^* = 1|A = k) - Pr(Y^* = 1|A = j)$$
$$= Pr(Y^* = 1|Y = 1)[Pr(Y = 1|A = k) - Pr(Y = 1|A = j)] \quad (2.5)$$

As a method to correct for the hypothetical bias, we use a follow-up certainty question, where respondents are asked how sure they are to perform the questioned action in reality (Blomquist et al., 2009; Morrison and Brown, 2009). We employ this method by asking respondents how high they perceive the probability to really train a person with the shown profile in case they accepted the profile. This question is shown immediately after the respondent accepted the profile and before the next profile is shown. It is answered with a slide control, which ends were labeled with "very unlikely" (value of 0.01) and "very likely" (value of 1). The values are not shown to the respondents since we found it difficult for respondents to provide an exact number to the question. The aim of a follow-up question is to give respondents the opportunity to qualify their response given in the first question. We do not show the follow-up question to respondents who do not accept a profile because we assume they have no incentive to conceal a true positive attitude. When confronted with a certainty scale, they might qualify their response which would in fact increase social desirability bias. A further concern might be that respondents in the beginning of the DCE are not aware of all the possible attribute levels. Therefore, they accept a profile in the beginning, which they would not accept after seeing later ones due to a learning process (and vice versa). However, the respondents have the possibility to go back in the survey at any time to change any given answer.

The economic literature either uses the answer possibilities "definitely sure" and "probably sure" for the follow-up question (where only definitely sure is treated as a true acceptance), or rely on discrete scales (such as 10-point scales) that are transformed into binary indicators (for example all answers that are 9 and higher are treated as true acceptance). Experiments with real scenarios are used to validate this approach (Blomquist et al., 2009; Morrison and Brown, 2009). We assume the answer to the follow-up question to provide an unbiased estimate of the true acceptance in case the person reported to accept the respective profile. In our particular setting, we cannot rely on a realistic scenario. Nevertheless, our survey design provides us with some indicator for misreporting that can be used to validate this assumption. A later section will provide empirical evidence for the unbiased follow-up response assumption.

2.4 Data

2.4.1 Data collection

We conduct a discrete choice experiment (DCE) among all Swiss companies that provide standard dual-track VET for commercial occupations in three cantons (St. Gallen, Thurgau, and Appenzell). We focus on commercial occupations because we want to reduce the variance among our respondents. However, since commercial occupations are among the most popular occupations for dual-track VET programs (BBT, 2010), our results are valid for a large share of potential participants for S-VET.

The DCE is part of an online survey among employees who are responsible for standard VET. The addresses of the companies and the names of all responsible persons for VET were provided by the cantonal offices for professional education. In total, our database included more than 1,800 persons.

A key problem in all company surveys is the low response rate. To increase response, the survey was announced to company members of local commercial associations. To sensitize employers to S-VET, we placed an article in the journal of the Swiss Employers' Association which also announced our survey (Meisen Zannol, 2010). Most importantly, to collect all e-mail addresses from employees who are responsible for standard dual-track VET, we called all companies. This provided us with the opportunity to announce our study and to explain the goals. All in all, we collected 1,527 e-mail addresses to which we sent a personalized invitation to participate in our survey.

Field work started on November 9, 2010. We sent invitations at staggered intervals over a time period of five days. Two weeks later we sent a reminder to those persons, who had not yet responded to the survey. The survey was closed after four weeks. 759 persons (50%) participated in the survey, from which 533 persons completed the discrete choice experiment (effective response rate 35%). Our response rate is considerably higher than the response rate from comparable surveys. A company survey on behalf of the Swiss Ministry for Social Insurance yielded a response rate of only 24% (Baumgartner et al., 2004). A similar response rate of 25% was achieved in a company survey in the US (Bricout and Bentley, 2000).

Since S-VET has not yet been widely implemented in Switzerland, it is likely that respondents were not aware of the program. We therefore confronted respondents at the beginning of the survey with an explanation of the program that should help them to answer the following questions. A link to a webpage providing further information was given. The questionnaire started with a number of warm-up questions. In these questions

we asked about the opinion on possible supporting services and potential barriers to participate in S-VET. After the DCE we collected data on company and individual background variables.

2.4.2 Descriptive Statistics

A detailed description of the sample of respondents is given in Table A-2 in the appendix. Our sample included 60% women. The average age is 40.67 years (SD 10.68). Job tenure as a vocational trainer is fairly evenly distributed, and already 16% had previous experience with training an apprentice with a disability. Nevertheless, S-VET is relatively rare, with only three respondents reporting that they had any experience with it. About 95% claimed to have decision making authority (either alone or with somebody else) when hiring apprentices, which made us confident that our survey reached the right decision makers. A surprisingly high share (72%) knows a person with a mental disorder.

The majority of the respondents (47%) work in the service and administration sector, 17% work in public administration, and 13% work for banks or insurances. Company size is usually low; more than 50% work for companies with 50 or less employees. This corresponds with the economic structure of Switzerland in general and our study setting in particular, where most companies are small and medium-sized enterprises (BFS, 2009). 38% report that their companies have any experience with hiring people with disabilities, 13% report that the company has special hiring policies for people with disabilities and 10% report that the company has an equity policy or Managing Diversity guidelines. Even though, there is no representative data for Switzerland, these relatively high numbers could already indicate a selection problem: Companies who are willing to integrate people with disabilities may be more likely to participate in our survey.

On average, 1.09 (SD 1.2) of the five different profiles were accepted. A large share of respondents (41%) did not accept any profile. We test whether any of the background characteristics (individual, company, work environment) are associated with a different willingness to accept a profile (columns 3 to 5 of Table A-2). Younger and less experienced trainers are willing to accept a higher number of profiles. Non-profit organizations are willing to accept more profiles than for-profit organizations. Companies which are sensitized to the issue of employment of persons with disabilities may be more likely to accept profiles. The sensitivity to the topic is measured by three dummy variables; if the company has experience in hiring people with disabilities, if it has special hiring policies for them and if it has equity policy or Managing Diversity guidelines. However, the difference is only weakly significant for those with experience in hiring people with

disabilities. Providing training to people with mental illness may be easier in work environments where the apprentice has no customer contact, no open plan office, and flextime. However, the difference is only significant for one characteristic: For the rare case that the apprentice has no customer contact at his workplace (only 3% of the companies offer that kind of workplace), the number of accepted profiles is almost twice as high. All other differences are not significant.

2.5 Data analytic procedure

2.5.1 Outcome Variables

In our empirical analysis we are interested in the likelihood that a specific profile is accepted. Here we use the individual profiles as units of analysis (N=2,656). Two different outcome variables are used: outcomes with and without the proposed correction for the hypothetical bias. Uncorrected outcome variables simply represent the answer to the question, whether or not the person can imagine to provide training to a person with the given profile (dichotomous variable 0: No; 1: Yes).

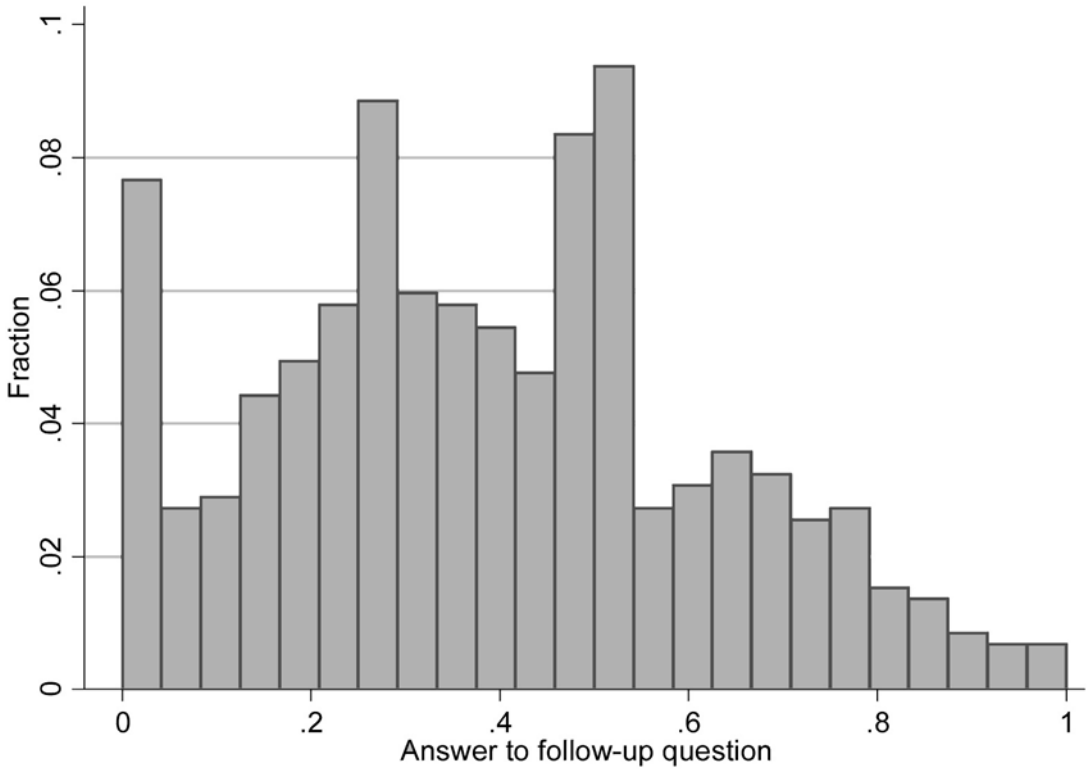
We correct for the hypothetical bias using a follow-up question. We asked respondents to estimate the probability to train this person given the person had initially accepted the profile. We assume that this answer is an unbiased estimate for the true probability of providing training conditional on people claiming that they could imagine doing so. We therefore replace the answer to the DCE for those who said they could imagine training that person with the self-reported probability in the follow-up question (see equation 3). Figure 2.1 illustrates the distribution of the self-reported probability.

2.5.2 Testing the unbiased follow-up assumption

Our prime assumption is that answers to the follow-up question are an unbiased estimate for the true probability of providing training conditional on people claiming that they could imagine doing so. We test this assumption using an indicator for misreporting that is not related to the attribute level or to the true outcome. If the proposed correction method provides an unbiased estimate for the true probability, corrected outcomes should not be associated with this exogenous indicator for misreporting.

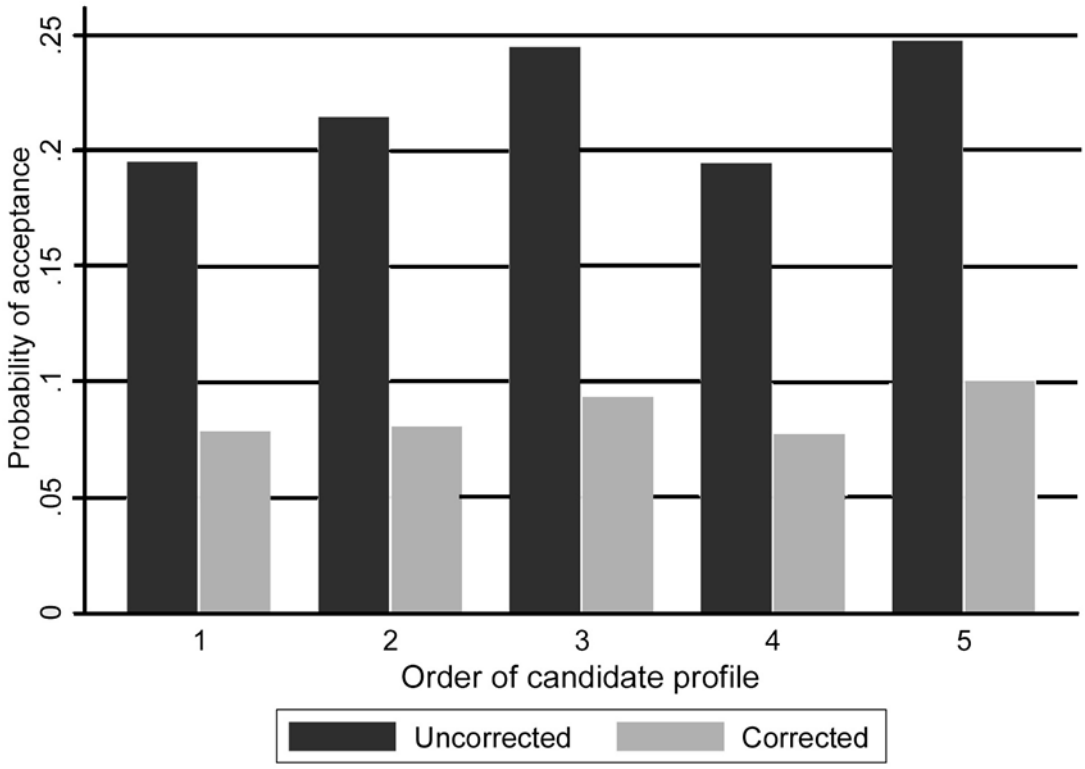
We use the ordering of profiles as such exogenous indicator. Each respondent sequentially evaluated five different profiles, where each profile was randomly generated and independent from the previous profile. Consequently, in expectation, one profile should

Figure 2.1: Distribution of answers to the follow-up question



be as "employable" as the other profile. With no misreporting, we would expect that average acceptance rates for each profile (first profile to fifth profile) would be the same. Considerable differences in acceptance rates therefore indicate a bias. This experiment is likely to be a pure "instrument" for misreporting. Ordering is random. Neither do we provide more information on S-VET programs, nor do we include further information on candidate profiles. Our results show that there are sizeable differences in acceptance rates (see Figure 2.2, dark bars). The acceptance rates of the third and fifth DCE are five percentage points higher than the acceptance rates of the first DCE. Thus, it seems that respondents felt particularly in the middle and in the end under pressure to accept a profile. We test whether there are significant differences in acceptance rates by DCE order using a standard F-test. We find significant differences when using the uncorrected results (F-test, 2671 degrees of freedom; p=0.08), but no significant differences when using the corrected answers (F-test, 2671 degrees of freedom; p=0.20). Consequently, this test points towards the validity of our main assumption, correcting for the hypothetical bias using the follow-up question results in an unbiased estimate for true acceptance.

Figure 2.2: Uncorrected and corrected probability of acceptance by order of candidate profile



2.5.3 Multivariate analysis

In our multivariate analysis, we are interested to which degree the different profile attributes affect acceptance. Note that our corrected outcome is still in the range between 0 and 1, but no longer dichotomous. Consequently, standard discrete choice models, such as Probit or Logit models cannot be used. To guarantee comparability between the empirical models, we use standard ordinary least squared (OLS) regressions on uncorrected and corrected outcomes. The model parameters correspond to the absolute treatment effects, the relation between respective model parameters denote the relative treatment effects.

We are well aware of the shortcoming of the linear probability model (LPM), particularly that some combinations of values for the independent variables can result in predictions outside the (0,1) interval. However, in our application, this is only a minor problem, since only 5.2% of the uncorrected predictions and 6.4% of the corrected predictions lie outside the interval, i.e. are below zero. A further problem is that residuals are heteroskedastic by definition. OLS estimators are still unbiased, but the standard errors, and hence the

t-values, are not. We solve this problem by estimating standard errors which are robust to heteroskedasticity. We furthermore check if marginal effects from a standard Probit and Logit model using uncorrected outcome variables (which corresponds to treatment effects since all independent variables are discrete) differ to our results from the OLS regression. The results are not reported but available from the authors upon request. They do not indicate that the usage of LPM results in a major bias.

2.6 Results

2.6.1 Overall acceptance rate

Using the uncorrected outcome variable, the estimate for the average acceptance of our profiles is 22% (95% confidence interval: 20.4%-23.5%). However, for two reasons we think that this number provides a biased estimate: *(i)* Our sample is self-selected as only 35% responded. We have very little information on those who did not respond (we only know the branch and gender), so that any *selection on observables* strategy is unfeasible. Worst-case bounds assuming that either all people who did not participate would either not accept any profile or would accept all profiles are rather large (8-73%). However, the overall high rate of companies in our sample that already have experience with training people with disabilities makes us to believe that true acceptance may be at the lower end of the worst-case bounds. Consequently, all estimates are unlikely to be externally valid, and our experiment can only provide internally valid results. *(ii)* Our main concern for internal validity is the hypothetical bias, which overestimates the overall acceptance rate, and may also bias our estimates for the average treatment effects.

Using the outcome variable corrected with the follow-up question, the acceptance rate within our sample shrinks to 8.6% (95% confidence interval: 7.9%-9.3%; worst case bounds: 3% - 68%). This number is considerably lower because almost 70% of the respondents estimated probability of acceptance to be less than 0.5, with a relatively high clustering of answers at 0.01 (see Figure 2.1). This fact documents that the hypothetical bias is severe and it seems to be the case that respondents used the follow-up question to "correct" their initial answer to the DCE. Answers to the DCE without correction would yield in a too high estimate for the true willingness to participate in S-VET.

2.6.2 Barriers to participate in S-VET

Table 2.1 presents the results from our multivariate analysis using OLS. Columns 1-3 contain the results for the model that uses the uncorrected responses to the DCE as the dependent variable, while columns 4-6 contain the results using the corrected responses. The regression coefficients can be interpreted as the absolute treatment effects, while the numbers in the brackets denote the relative treatment effects to the reference group (i.e., the respective regression coefficient divided by the constant).

Comparing absolute treatment effects for both models indicate that the hypothetical bias is indeed severe. The coefficients are up to four times larger in the uncorrected model compared to the corrected model. The relative treatment effects for corrected and uncorrected outcome variables, however, are fairly similar. Major differences are only observed for the diagnosis and age. Thus, false responses seem to be only weakly associated with different attribute levels. Consistent with our model, using uncorrected coefficients leads to significantly overestimated treatment effects.

Even though our model may not explain a large amount of variation in the response (R^2=0.09), several attributes are highly significant. The most relevant attribute for acceptance is adherence to regulations. If an applicant is rarely reliable in terms of punctuality, agreements, and regulations, she has a 20% lower chance to be accepted for S-VET compared to an applicant who fulfills this attribute almost always. Other important aspects are competency (ability to perform work in a reliable and fast way), as well as contacts with others.

Note that these relevant attributes all correspond to non-cognitive dysfunctions. Poor school performance as an indicator for cognitive dysfunctions decreases the probability to be accepted for S-VET. However, compared to non-cognitive dysfunctions, the effect is rather small.

Regarding health related variables, we find that a poor self-maintenance as well as a high number of absence days reduces the willingness to accept a profile. The overall effect, however, is considerably smaller than the effects of most non-cognitive dysfunctions. This result is in line with the literature that shows a large part of costs of mental illnesses for an employer comes from presenteeism, i.e. when the person is at work (Goetzel et al., 2004).

Our results provide no clear answer to the question whether a diagnosis should be disclosed in an application or not. We briefed all respondents that S-VET is primarily targeted at students with mental problems. To test for a possible stigmatizing effect, we provided one group of respondents with the additional information of a diagnosis

Table 2.1: OLS regression results

	Uncorrected			Corrected		
	Coef.	t-value	%	Coef.	t-value	%
Diagnosis (omitted category: not mentioned)						
Bipolar disorder	-0.02	-0.60	-2%	-0.02	-1.19	-5%
Schizophrenia	-0.07***	-2.73	-9%	-0.04***	-2.95	-12%
Borderline Personality Disorder	-0.05**	-2.07	-7%	-0.02*	-1.73	-8%
Eating Disorder	0.05*	1.87	7%	0.04**	2.47	13%
Work experience (omitted category: start of VET)						
First year	-0.03	-1.45	-4%	-0.02*	-1.70	-5%
Second year	-0.05**	-2.29	-6%	-0.01	-1.17	-4%
Third year	0.00	-0.14	0%	0.00	0.10	0%
Age (omitted category: 16-18)						
19-21	-0.05**	-2.35	-7%	-0.03***	-2.89	-11%
22-24	-0.06***	-2.89	-8%	-0.03***	-2.58	-9%
25+	-0.09***	-4.27	-12%	-0.04***	-4.29	-15%
School performance (omitted category: 5.5 good-very good)						
4 (marginal)	-0.08***	-3.82	-11%	-0.02*	-1.91	-7%
4.5 (satisfactory)	-0.05**	-2.50	-7%	-0.02	-1.62	-5%
5 (good)	-0.01	-0.49	-2%	0.01	0.53	2%
Absence (omitted category: none)						
1-2	0.00	-0.14	0%	-0.01	-0.64	-2%
3-4	-0.04**	-1.98	-6%	-0.02**	-2.33	-8%
5 and more	-0.06***	-2.87	-8%	-0.03***	-3.42	-11%
Motivation (omitted category: always)						
Sometimes	-0.07***	-3.43	-9%	-0.03***	-3.37	-10%
Rarely	-0.13***	-6.54	-17%	-0.05***	-5.59	-17%
Adherence to regulations (omitted category: always)						
Sometimes	-0.12***	-5.89	-16%	-0.04***	-4.54	-14%
Rarely	-0.17***	-9.09	-22%	-0.06***	-7.23	-20%
Competency (omitted category: always)						
Sometimes	-0.10***	-5.17	-13%	-0.04***	-4.40	-13%
Rarely	-0.14***	-7.01	-19%	-0.05***	-5.83	-18%
Flexibility (omitted category: always)						
Sometimes	-0.06***	-3.01	-8%	-0.02**	-2.06	-6%
Rarely	-0.07***	-3.65	-10%	-0.04***	-4.07	-12%
Contact with others (omitted category: always)						
Sometimes	-0.06***	-3.16	-8%	-0.02***	-2.74	-8%
Rarely	-0.14***	-7.29	-19%	-0.06***	-6.07	-19%
Self-maintenance (omitted category: always)						
Sometimes	-0.02	-1.07	-3%	-0.01	-0.96	-3%
Rarely	-0.07***	-3.44	-9%	-0.03***	-3.58	-11%
Constant	0.75***	15.75		0.30***	12.82	
N		2656			2656	
R^2		0.11			0.09	

Notes: ***, **, and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels, respectively. T-values are robust and clustered by individual respondents.

(randomly chosen from four different diagnoses), while we did not show this information to the control group. Our results show that the effect of disclosing the diagnosis varies with the diagnosis. Psychotic disorders clearly act as a deterrent. The diagnosis *Schizophrenia* for example reduces the individual willingness to accept a profile by 12% (conditional on all other chosen attributes). Patients may therefore not only be burdened with their disease itself but also with social stigma, which may be enforced by movies and the popular literature (*Dr. Jekyll and Mr. Hyde*). In contrast, acceptance of *eating disorders* is 13% higher if the diagnosis is disclosed. However, since mostly young women are affected, this could also be a gender effect.

Regarding non-disability related characteristics, we find that the age of an applicant is inversely related with the willingness to accept a profile. This demonstrates that a relatively fast integration into the labor market (providing that the individual health allows for this) is not only in the interest of the patient but also in the interest of a potential employer. Finally, our results indicate a u-shaped relationship between work experience and acceptance: Employers are more willing to take over the complete training, or to accept students in their last year of training when they already have some work experience. The overall magnitude of this effect, however, is relatively small compared to the importance of non-cognitive dysfunctions.

The previous analysis documents average treatment effects. To analyze effect heterogeneities, we stratify our sample by different firm level and respondents' characteristics (results are not presented but are available from the authors upon request). The treatment effects are fairly stable and do not vary much. One notable exception is the treatment effects for school performance. Publicly owned companies have a strong preference for good school grades, while privately owned companies do not seem to care as much about school grades.

2.7 Discussion

It is often argued that Supported Education/Employment is a (cost-)effective alternative to standard (mainly institutionalized or sheltered) rehabilitation methods, particularly for young people with mental illnesses. However, our results document that very few employers are currently willing to participate in this rehabilitation method. Even though more than 20% of our profiles were accepted and about 60% of our respondents accepted at least one profile, we believe that these numbers are grossly inflated and that the true acceptance rate is far lower. Our best estimate for the acceptance rate within our

sample is 9%. However, we have to keep in mind that the response rate was only 35%, and that it is very likely that companies, who are in principle willing to participate in Supported Vocational Education and Training, are also more likely to respond. The overall acceptance rate could thus be as low as 3%.

Our results indicate that non-cognitive dysfunctions related to psychological disorders are the main deterrents. These results are in line with the literature arguing that a substantial part of the costs of mental illnesses for an employer is driven by presenteeism (i.e., when the person is at work).

Supported Education/Employment programs are therefore unlikely to be a magic bullet. A wide implementation of these programs is likely to fail on the unwillingness of companies to participate in this measure. This is in spite of the fact that no direct costs occur to employers. Providing additional incentives, for example in form of subsidies or legal requirements, may therefore be needed.

Our study has been tailored to the Swiss Vocational Education and Training system. Yet, many other countries have similar dual-track education systems (such as Germany, Denmark, Czech and Slovak Republic, or Austria). Our results may therefore be particularly relevant for these countries. In addition, because we assume non-cognitive dysfunctions to be similarly important for the acceptance in other types of supported education or supported employment programs, our findings might generalize to these programs as well. Further research may be needed to substantiate this belief.

2.8 Appendix

Table 2.A1: DCE attributes and levels

Attribute	**Description**	**Levels**
Diagnosis	Diagnosis	Not mentioned Bipolar disorder Schizophrenia Eating disorder Borderline personality disorder
Work experience	Current year of Supported Vocational Education and Training	Start of VET First year Second year Third year
Age	Age category	16-18 19-21 22-24 25+
School performance	Average grade in school (previous year of school)	4 (marginal) 4.5 (satisfactory) 5 (good) 5.5 (good - very good)
Absence	Average days of absence due to the illness in the past month	0 1-2 3-4 5 and more
Motivation	Student is self-motivated and willing to learn	Rarely Sometimes Almost always
Adherence to regulations	Student is reliable in terms of punctuality, agreements, and regulations	Rarely[a] Sometimes Almost always
Competency	Student carries out work quickly and reliably	Rarely Sometimes Almost always
Flexibility	Student is flexible to new tasks or to changes in the environment	Rarely Sometimes Almost always
Contact with others	Student becomes part of the work team and acts appropriately in social contacts	Rarely Sometimes Almost always
Self-maintenance	Student pays attention to his/her own health, knows his/her limits and gets help when needed	Rarely Sometimes Almost always

[a] Originally, the ICF-Mini Rating has a 5-digit scale, ranging from *no* to *complete*. We found this difficult to interpret for our purposes and therefore changed it to a time-perspective. The two most extreme levels (in the sense of *Very rarely* and *Never*) were not included because we thought a person with those levels is not employable.

Table 2.A2: Descriptive statistics (background variables)

Variable	Proportion	Number of accepted profiles	Std. Dev.	p-values of difference
Individual information				
Age				
40 and older	0.57	1.02	1.16	0.035**[a]
Younger than 40	0.43	1.25	1.26	
Gender				
Female	0.6	1.07	1.16	0.691[b]
Male	0.4	1.16	1.28	
Years of job experience as VET trainer				
0-2	0.16	1.4	1.47	0.049**[b] (4)
3-5	0.26	1.07	1.11	
6-10	0.25	1.22	1.21	
11-20	0.21	0.86	1.1	
>20	0.12	0.98	1.13	
Experience to train persons with disabilities				
Yes	0.16	1.15	1.19	0.593[a]
No	0.84	1.1	1.21	
Previous experience with S-VET				
Yes	0.01	0.67	0.58	0.644[a]
No	0.99	1.11	1.21	
Decision-making authority for selecting apprentices				
None	0.05	1.19	1.11	0.799[b] (2)
Joint with others	0.84	1.11	1.23	
Alone	0.11	1.03	1.08	
Knows person with mental disorder				
Yes	0.72	1.12	1.22	0.799[a]
No	0.28	1.08	1.19	
Company background				
Branch				
Service & Administration	0.47	1.25	1.32	0.217[b] (5)

Continued on next page

Table 2.A2 – *Continued from previous page*

Variable	Proportion	Number of accepted profiles	Std. Dev.	p-values of difference
Public Administration	0.17	1.01	1.04	
Trust / Real estate trust	0.06	1	1.29	
Bank / Insurance	0.13	0.81	0.92	
Mechanical, Electrical & Metal	0.08	0.98	1.1	
Other	0.09	1	1.2	
Non-profit organization				
Yes	0.11	1.34	1.31	0.100*[a]
No	0.89	1.06	1.19	
Public or semi-public ownership				
Yes	0.24	1.13	1.12	0.383[a]
No	0.76	1.09	1.23	
Number of employees				
<10	0.18	1.04	1.19	0.212[b] (4)
10-50	0.41	1.13	1.18	
51-100	0.13	1.39	1.42	
101-500	0.19	1.05	1.16	
>500	0.09	0.84	1.06	
Experience with hiring people with disabilities				
Yes	0.38	1.22	1.24	0.078*[a]
No, Don't know	0.62	1.04	1.18	
Special hiring policies for people with disabilities				
Yes	0.13	1.29	1.39	0.349[a]
No, Don't know	0.87	1.07	1.17	
Managing Diversity or equity policies				
Yes	0.10	1.16	1.38	0.895[a]
No, Don't know	0.90	1.09	1.18	
Number of apprentices in commercial occupations				
0	0.08	1.26	1.19	0.256[b] (5)
1	0.33	1.25	1.28	
2	0.16	1.14	1.33	

Continued on next page

Table 2.A2 – *Continued from previous page*

Variable	Proportion	Number of accepted profiles	Std. Dev.	p-values of difference
3-5	0.26	0.95	1.09	
6-10	0.11	0.96	1.08	
>10	0.06	0.91	1.06	
Work environment				
Apprentice has customer contact				
Yes	0.97	1.07	1.18	0.007***[a]
No	0.03	2.06	1.57	
Open plan office				
Yes and partly during apprenticeship	0.63	1.09	1.21	0.530[a]
No	0.37	1.14	1.19	
Flextime				
Yes	0.52	1.13	1.22	0.755[a]
No	0.48	1.08	1.19	

Notes: Numbers of observations vary between 516 and 531 because not every participant provided information on every background variable. In column 5 p-values are reported from a test for significance of the difference in the number of accepted profiles reported in column 3. The test is non-parametric because the dependent variable (number of accepted profiles) is not normally distributed. ***, **, and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels, respectively.

[a] P-value from a Wilcoxon-Mann-Whitney test.

[b] P-value from a Kruskal-Wallis one-way analysis of variance; degrees of freedom in parentheses.

3 Hiring Subsidies for People with a Disability: Helping or Hindering? - Evidence from a Small Scale Field Experiment

joint with Eva Deuchert

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 is implemented in a sample of clients from employment consulting services. Our results reveal that the subsidy is ineffective or even counterproductive in a group of adolescents who are at the end of their vocational training program, but may increase call-back rates in a group of clients of job coaching services.

3.1 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

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

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.

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

[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.

and Spiegelman, 1987). Second, in situations where field experiments are not feasible, natural experimental designs are used for ex-post evaluations (e.g. Blundell et al., 2004; Boockmann et al., 2012; Hamersma, 2008; Huttunen et al., 2013; 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; Dorsett, 2006; Forslund et al., 2004; Jaenichen and Stephan, 2009; 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).

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 that the subsidy is ineffective or even counterproductive in a group of adolescents who are at the end of their vocational training program. A positive (but not significant) effect is identified for 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 3.3 describes the study design, and section 3.4 the two different samples of study participants and the number of applications sent. Results are presented in section 3.5. The final section concludes.

3.2 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 3.1. Hiring subsidies are implemented because policy makers assume that many firms employ a person with a disability only if a subsidy is paid but not if it is unavailable (*compliers*). The benefits (e.g. reduced DI payments or higher

[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.

Table 3.1: Possible effects of hiring subsidies

		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)

tax revenue) are typically set such that they 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] violations are likely to bias the estimates for

[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.

effectiveness. In our case this is unlikely to be the case since each person is used as their own control group (see section 3.3). Thus, as long as not all applications are sent to the same potential employers, *SUTVA* violations will not affect effectiveness estimates.

3.3 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.

3.3.1 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 3.2). 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

Table 3.2: Study design

	Group 1	**Group 2**
Sample population	Graduates from 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 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	January 2012-May 2013
Geographical region	Appenzell, Lucerne St. Gallen, Thurgau	St. Gallen

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 a training 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 discussing the training grant).[7] In wave 2, we thus have three different

[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.

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

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]

When analyzing the results of the first wave (which was conducted half a year before the start of wave 2), we realized effect heterogeneities according to different application types: applications to advertised job offers and unsolicited applications (see section 3.5). 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 all these differences between the groups by providing pooled as well as stratified results (see section 3.5).

3.3.2 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

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.

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; Carlsson and Rooth, 2007; Drydakis, 2009; Oreopoulos, 2011; Riach and Rich, 2002). 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, 2004a). 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 be 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.

3.4 Data and descriptive statistics

In total, 52 individuals participated in our experiment, 39 in wave 1 and 13 in wave 2.[9] 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 3.3). 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 3.3 for details).

All participants received an answer sheet to document the calendar date of a company's reaction (invitation for 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 believe it is highly unlikely that a company would react in any way (especially positively) to an application after more than seven weeks. In fact, our results show participants received hardly any response after five weeks (see the survival functions in Figure 3.A1 in the appendix). There is no significant difference in the timing of response between the treatments.

From a total of 384 applications, 24 (6.25%) resulted in a call-back for a job interview (see Table 3.3). 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. 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

[9] When we started the research project, we expected a higher number of participants in wave 2 (60-80). Since we did not contact participants directly (the initial contact was via DI case workers), we do not know exactly the reason for the disappointing low number. It may be the case that potential participants were discouraged because the workshop did not take place at the local DI office but at the university. Alternatively it may be also the case that many DI case workers did not call attention to our study (which is very likely since the majority of participants were sent by only a very small number of case workers).

Table 3.3: 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.

(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 3.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 acquired their disability after they completed their formal education. Vocational education is lower for participants of wave 2, 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 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. 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 3.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 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

[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).

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

[12] TREE (TRansition from Education to Employment) 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.

problem (column 4 of Table 3.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.

Descriptive statistics of the 384 employers receiving the applications are limited and shown in Table 3.A2 in the appendix. More than four out of five companies can be assigned to the service sector, while the rest is in the industrial sector. A similar ratio is found for profit organizations compared to non-profit organizations. There are no significant differences in companies' chararacteristics between the waves.

3.5 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 3.4). Because we compare the same subjects with themselves under different treatments, we estimate average call-back rates for applications with the same treatment for each participant and use a non-parametric Wilcoxon signed ranks test to test for statistical significance.[14] 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 statistical difference between the two waves: 0.39). The negative effect in wave 1 is predominantly driven by unsolicited applications (see Table 3.A3 in the appendix). As been discussed in section 3.3, some participants of wave 1 could not find suitable offers and therefore wrote only unsolicited applications. Considering only participants who wrote both types of applications yields an even higher effect (-0.175, p=0.158).

For wave 2 we do not find similar results. The treatment effects for regular and unsolicited applications are both positive, yet not significant. The design of wave 2 furthermore allows isolating the signaling effect from the incentive effect of the subsidy. The incentive effect can be estimated by comparing call-back rates for applications disclosing the disability but not the subsidy (i.e. including a letter from the local DI office

[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 wave 2009.

[14] Note that in this case using every application as an independent observation (and therefore using e.g. a two-sample t-test) could result in overestimation of p-values (see e.g. Sainani, 2010).

Table 3.4: Treatment effects

Difference in call-back rate within applicant	Notification of subsidy vs. No notification	Notification of subsidy vs. General support	Notification of general support vs. No notification
Effect type	**Total**	**Incentive**	**Signaling**
Full sample	-0.005		
	(0.953)		
Wave 1	-0.019		
	(0.534)		
Wave 2	0.038	0.051*	-0.013
	(0.261)	(0.084)	(0.317)

Notes: P-values of Wilcoxon signed rank test in parentheses. * indicates statistical significance at the 10 percent level.

that offers general support only) and for applications disclosing the disability and the subsidy (see column 2 in Table 3.4). The effect is 5.1 percentage points and significant at the 10% level. Assuming that offering general support has only a minor impact on a 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. For wave 2, this signaling effect is relatively small and insignificant (see column 3 of Table 3.4).

Assuming the incentive effect would be of a similar magnitude for participants of wave 1, the signaling effect would be sizeable (-7 percentage points). The reason for this effect heterogeneity is likely to be grounded in the underlying differences in characteristics. Participants of wave 1 have just completed vocational education and training, typically search for employment in their profession, and have often no major gaps in their CV. For these individuals, a subsidy aimed at compensating higher initial costs for on-the-job training may not be needed. The signaling effect, however, is likely to be higher for this group if potential employers would not guess a health problem from the CV. Participants of group 2, in contrast, have major gaps in their CV and often search for employment outside their prime profession. For these individuals it may be necessary to compensate for any higher training costs. The signaling effect (in addition to the signal from gaps in the CV) is likely to be smaller.

Given the small sample size of participants, looking for further effect heterogeneities with respect to individual 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. To test for significant differences we regress the call-back rate on a treatment dummy (=1 if notification of subsidy is included; =0 if any notification is missing) with fixed effects and weighted by the inverse number of applications from a participant; separately for two groups with and without the particular characteristic. Table 3.A4 in the appendix reports the coefficients of these separate regressions and its difference. To test for significance of the difference a two-sample mean-comparison test is used. The effect of the hiring subsidy seems to be particularly negative for firms in the industrial sector compared to no effect in companies in the service sector. This difference is significant at the 10%-level. Thus, if a job requires physical strain, which is usually common in the industrial sector, an employer might be particularly deterred by a notification of a disability.

3.6 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 may depend on the respective target group. While the subsidy seems to be 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. Unfortunately, these heterogeneities are not statistically significant, which could be due to the low sample size. Larger-scale studies are needed to substantiate these current findings.

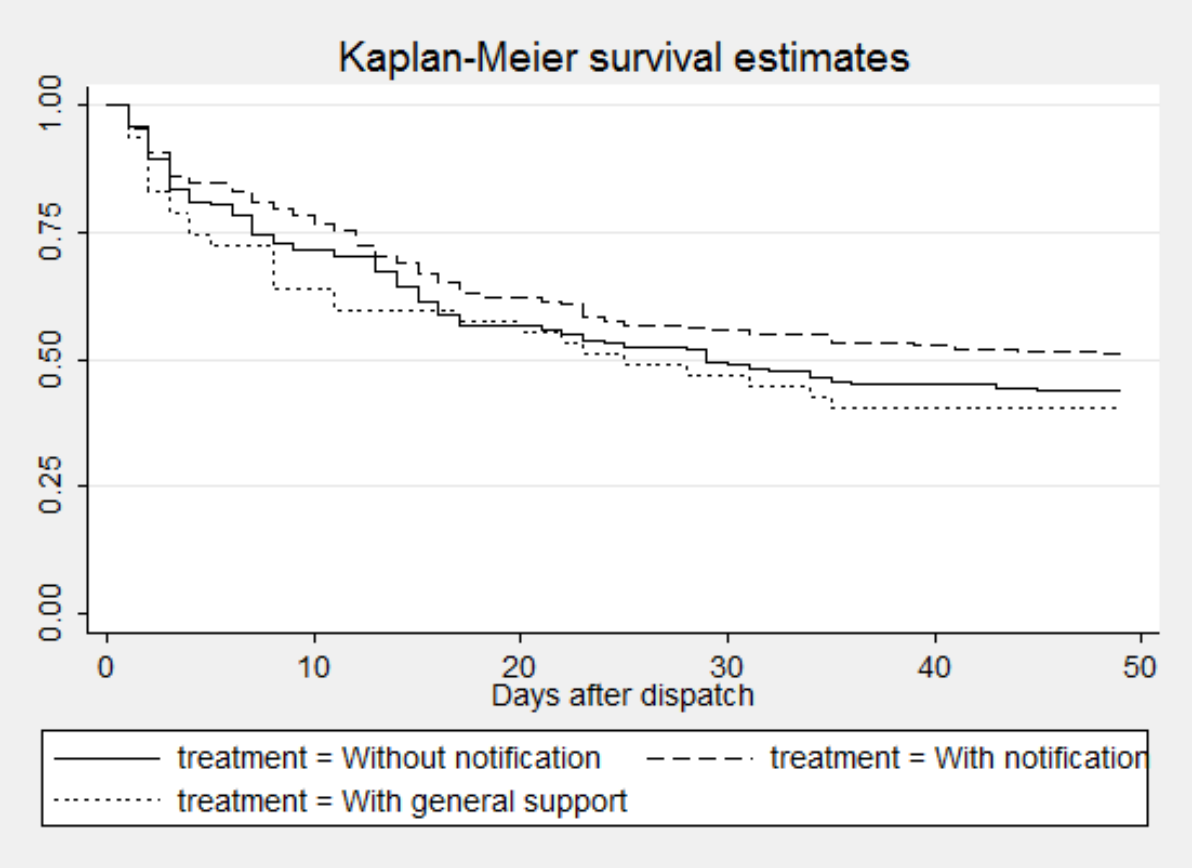
Different treatment effects may be grounded in the difference between the two groups: The first group consists predominantly of young people who acquired their disability before they completed their education and who are just at the end of their vocational training program. For this group, the training 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. Policy makers should then be careful to whom subsidies are targeted: Subsidies

may be ineffective or even counterproductive if expected productivity loss is relatively low (or in other words, when the subsidy is not needed) and the disability is not directly observable.

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 firms were willing to accept a candidate with a handicap anyway. Second, and far more relevant, our sample may not be representative for the general population eligible for training grants. Our results are thus causal for the chosen groups but may not correspond to the average treatment effect.

3.7 Appendix

Figure 3.A1: Kaplan-Meier survival function - Days until rejection



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 3.A1: Descriptive statistics: Participants

	Wave 1 (N=39)	Wave 2 (N=13)	TREE (N=1750)	SHP (N=1810)
Demographics observable from CV				
Male	0.49	0.53	0.48	0.43
Age	21.53	44.90***	19.55**	46.78
Married	0.06	0.38***	0.01***	0.58
Swiss	0.88	0.58**	0.81	0.93†††
German language	0.88	0.69	0.50***	0.74
Non-german family name	0.33	0.46		
Work & education history				
Finished apprenticeship	0.97	0.69***		
Duration of apprenticeship in years	2.13	3.11***	3.14***	
Finished school in regular time	0.52	1***		
Has ever visited special school	0.21	0*		
Years of work experience				
in sustainable employment	1.30	20.96***		
Gaps in CV	0.36	0.85***		
Unexplained gaps in CV (in months)	5.59	9.85		
Explained gaps in CV (in months)	1.51	32***		
Learned profession	%	%		
Industrial sector	67	69		
Service sector	33	31		
Importance of job characteristics				
Short commute to work	2.64	2.69	2.62	
Working part-time	2.16	2.85*	2.08	
Job security	3.69	3.61	3.62	
Able to learn something new	3.34	3.15	3.36	
Personality traits				
Self-efficacy	2.88	2.96	3	
Social Support	3.41	3.31	3.6	
Self esteem	4.02	3.9	4.1	
Positive attitude towards life	4.41	3.92**	4.66**	

Continued on next page

Table 3.A1 – *Continued from previous page*

	Wave 1 (N=39)	Wave 2 (N=13)	TREE (N=1750)	SHP (N=1810)
Self-responsibility/locus of control	4.89	4.31**		
Big Five Personality traits				
Extraversion	0.33	0.23		1.32††
Agreeableness	1.22	1.23		1.37
Conscientiousness	1.53	2.31*		1.94
Neuroticism	0.08	0.23		-1.04††
Openness	1.28	0.77		1.12
Perception about disability				
Self-perceived disability				
or health impairment	0.35	0.92***		
Type of disability		%		
Physical		92		
Mental		33		
Intellectual		0		

Source: own calculations, TRansition from Education to Employment (TREE) and Swiss Household Panel (SHP, wave 2009).
Notes: Numbers reported are means unless otherwise stated. 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 Five 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). Number of observations in wave 1 and 2 may be lower as some characteristics were not observed or some questions were not answered. 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). 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 3.A2: Descriptive statistics: Employers

Characteristic	**Wave 1**	**Wave 2**
For profit organization	0.880	0.841
Service sector	0.824	0.828

Notes: Classifications are based on authors' assessment of information available from the internet and the job ad. There are no agricultural companies so that a company belongs to the industrial sector if not to the service sector.

Table 3.A3: Treatment effects for unsolicited and regular applications

Difference in call-back rate within applicant	Notification of subsidy vs. No notification	Notification of subsidy vs. General support	Notification of general support vs. No notification
Effect type	**Total**	**Incentive**	**Signaling**
Unsolicited applications			
Full sample	-0.043		
	(0.327)		
Wave 1	-0.108*		
	(0.084)		
Wave 2	0.038	0.038	0
	(0.317)	(0.317)	-
Regular applications			
Full sample	0.036		
	(0.323)		
Wave 1	0.044		
	(0.415)		
Wave 2	0.018	0.040	-0.021
	(0.597)	(0.158)	(0.317)

Notes: P-values of Wilcoxon signed rank test in parentheses. * indicates statistical significance at the 10 percent level.

Table 3.A4: Effect heterogeneities by employer characteristics

	Notification of subsidy vs. No notification		
Characteristic	Yes	No	Difference
Service sector	0.017	-0.115*	0.132*
	(0.032)	(0.056)	(0.074)
For profit organization	-0.009	-0.055	-0.046
	(0.037)	(0.091)	(0.097)

Notes: Each cell in the first two columns reports the coefficient from a fixed-effect regression of the call-back rate on the treatment weighted by the inverse number of applications from a participant. Standard errors in parentheses. Standard errors for the difference are estimated with two-sample t-tests. * indicates statistical significance at the 10 percent level.

4 The Effect of Cutting Disability Insurance Benefits on Labor Supply in Households

Abstract

Previous empirical literature has shown a substantial extent of work disincentives in the Disability Insurance (DI). While its focus has been on the inflow into DI and on increases in benefits, this study tests whether the results are symmetric for existing beneficiaries and for a benefit cut. I use exogenous variation created by a partial benefit reduction for married DI beneficiaries. The richness of the dataset allows me to look at the behavioral response on labor market participation from the spouse and not only from the beneficiary. Using a difference-in-differences methodology, I find no effect on labor supply or earnings for both members of the couple. If anything, there might be spillover effects into means-tested social insurance. The results indicate policy makers should focus on the prevention of inflow into DI to cope with non-decreasing stocks of DI beneficiaries.

4.1 Introduction

The purpose of disability insurance (DI) is to guarantee individuals a certain standard of living if their working capacity is limited by a long lasting health-related problem. Yet, asymmetric information between the insurer and the claimant leads to moral hazard. For the insurer it is not possible to perfectly assess whether a claimant is truly disabled. The DI thus distorts work incentives and people with a large disutility of work may select into the DI. In this sense, DI differs from the unemployment insurance (UI) where people with a large disutility of work can be much better monitored through job search requirements. If moral hazard is substantial in the DI, variation in DI eligibility or in

benefit generosity should lead to variation in labor supply (for a theoretical model, see e.g. Halpern and Hausman, 1986).

As a second difference to UI, the outflow from DI (for other reasons than death or retirement) is much lower at about 1% (OECD, 2010). While the unemployed usually run out of benefits after a given period, DI benefits are granted permanently. In consequence, the DI has a much higher stock relative to the flows both in and out of the stock. With permanent benefits and low outflow, it is then no surprise that this system can be very expensive. For policy makers it is therefore crucial to know of effective measures to control the stock of DI beneficiaries.

By exploiting policies or policy reforms which applied to only a subset of the population, the previous literature has mainly addressed the inflow into DI and variation in DI eligibility. Most of these studies show substantial work disincentive effects (see e.g. Gruber, 2000; Staubli, 2011). This study, however, analyzes the behavior of the stock of beneficiaries to a partial reduction in benefits. I exploit exogenous variation by a policy reform of the Swiss DI that partially cuts benefits for married individuals only. Prior to the reform, married DI beneficiaries fulfilling very low eligibility criteria could draw an additional pension (AP) for their spouse, unconditional on the spouse's labor supply. Starting in 2004, new accompanying benefits for spouses were no longer granted. In 2008, all existing benefits for spouses were abolished. Beneficiaries lost up to 23% of their DI benefits, on average about 300 CHF per month, which is equal to about 5% of total household income. I apply a difference-in-differences methodology comparing beneficiaries who entered DI just prior to the first revision losing the AP in 2008 to beneficiaries who started to draw DI benefits just after the first revision never receiving an AP.

Another contribution of this study is to analyze the behavioral response not only of the beneficiary but also of the spouse. Today, many decisions in a household - especially on labor force participation - are taken in consideration of all household members and are dependent on total income of the household. For the main beneficiary, changes in labor supply could affect the amount of DI benefits even further. In contrast, for the spouse, the partial reduction through the abolition of the AP is a pure income effect being welfare neutral.

Unlike the previous literature on inflow, the results show no effects neither on employment nor earnings for both the beneficiary and the spouse. Due to the loss in income, part of the population seems to newly qualify for means-tested supplementary benefits. The reason why the effects for in- and outflow are not symmetric may be that

the beneficiaries' stock of human capital depreciates so rapidly that they are no longer able to find employment after a certain time in the DI. Another reason may be that the leisure-consumption trade-off is different for the beneficiary couple. The results indicate that policy makers should address the prevention of inflow into DI rather than spending too many resources on interventions trying to increase the outflow.

The paper proceeds as follows. The next section gives information on the background, including a literature review and the institutional setting of social insurances as well as of the policy reform in Switzerland. Section 4.3 describes the data. In section 4.4 the identification strategy is outlined. Section 4.5 presents the results, while section 4.6 concludes.

4.2 Background

4.2.1 Literature review

The literature review should give a broad overview of studies investigating the work disincentive effect of the DI and identify the gap which this study aims to close.The work disincentive effect of the DI has been most extensively studied in the United States. Bound (1989) was the first to argue that when using cross-sectional variation in potential DI benefits relative to previous earnings, the estimated elasticity of labor force participation with respect to benefit generosity would be inflated. The reason is that potential DI benefits are likely to be endogenous due to their relation to past earnings. Bound uses rejected disability applicants instead of non-recipients as a group to construct the counterfactual of DI recipients. As this group may still be different in many characteristics from DI recipients, Bound interprets the results as an upper bound for the behavior of DI recipients if those had not received DI benefits. This approach is still very popular as numerous recently published articles demonstrate (see e.g. von Wachter et al., 2011; Giertz and Kubik, 2011; Singleton, 2012).

Another way to quantify work disincentives in the DI is to rely on exogenous variation created by quasi-experiments in form of policies or policy reforms which apply to only a subset of the population. Some researchers estimated the effect of tighter or less tight eligibility criteria for DI benefits or of an increase in benefit generosity on labor supply in different countries (Staubli (2011) for Austria, Campolieti (2004) and Gruber (2000) for Canada, Karlström et al. (2008) for Sweden, and Chen and van der Klaauw (2008) and Duggan et al. (2010) for the United States). Other researchers used a change or

regional variation in screening stringency to evaluate the effect on labor supply (Autor and Duggan, 2003; Gruber and Kubik, 1997; Mitra, 2009). Most of them are able to quantify considerable work disincentive effects in the DI.[1]

While all these studies look at the effect on the inflow into the DI or on employment at the time of application, studies about the effect on existing beneficiaries are rare. Outflow from DI is generally low across all OECD countries; only around 1-2% of all beneficiaries leave the DI annually for reasons other than death or retirement (OECD, 2010). Two reasons have been identified for the low outflow: (i) There may be limited access to vocational rehabilitation and employment integration measures. A couple of countries have implemented special rehabilitation and integration measures targeting DI beneficiaries. The evaluation of these projects in the US and the UK, however, delivered disappointing results indicating low take-ups and no or only small effects on outflow (Adam et al., 2010; Clayton et al., 2011; Kornfeld and Rupp, 2000; Thornton et al., 2007). (ii) Due to means testing, many DI systems may generate considerable lock-in effects. Expanding work efforts reduces benefit levels and the implicit tax rate on labor supply can be quite a substantial disincentive for return to work. Various policies have therefore been introduced to encourage beneficiaries to return to work by reducing this implicit tax rate. In the US, for example, DI beneficiaries are also covered by health insurance. Thus, they face a tradeoff between work and combined cash as well as health coverage benefits. States were given the authority to expand health insurance coverage to include persons with disabilities at higher income levels. Yet, as Gettens (2009) shows, the effectiveness of this expansion on employment and DI benefit participation is small.[2] Campolieti and Riddell (2012) demonstrate positive effects of the introduction of an earnings disregard in Canada; increasing the propensity to work for men by 5.1 to 5.7 percentage points. The effect for women is even larger (7.9-9.5 percentage points) but not as precisely estimated. However, there is no effect on DI in- or outflow. They also evaluate the introduction of automatic reinstatement provisions whereby former recipients could remain eligible for DI when taking up work. This new measure had no effect on any of these outcomes.

Autor and Duggan (2007) provide an explanation for the low effects on existing beneficiaries: Beneficiaries may prefer leisure over labor even if work is not implicitly

[1] Yet, Staubli (2011) and Karlström et al. (2008) are also able to identify substantial spillover effects into other social insurances (unemployment or sickness insurance).

[2] The UK "Pathways to Work" program also included a financial incentive to return to work in addition to job assistance services. Exploiting regional variation the results show that the program has accelerated the outflow from DI benefits, but only for those individuals who would have left benefit roles in less than a year in any case (Adam et al., 2010). In addition it remains unclear which aspect of the program has contributed to this decline.

taxed. The authors exploit a change in the DI program for veterans, where veterans who served in the Vietnam War could increase DI benefits due to the inclusion of diabetes on the list of conditions. Because these benefits are not work-contingent or means tested, the estimated decrease in labor force participation is due to a pure income effect. Marie and Vall Castello (2012) are able to replicate this finding in a Spanish setting.

Empirical studies on the effect of exogenous reductions in DI benefits are scarce. The only exception is the removal of drug addiction as a disabling condition in the US in 1996, which fully terminated benefits of approximately 100,000 individuals. Most recently, Moore (2011) estimates considerable employment increases by 20-30 percentage points for this population. However, to the best of my knowledge there is no study which analyzes the effect of only a partial benefit cut applying to existing beneficiaries unconditional on any condition or type of disability.

Additional (self-)insurance against work-limiting disability can be provided by the spouse through the added worker effect. As a result of a negative income shock the spouse might increase her labor supply. While the added worker effect has mainly been studied in the unemployment literature (see e.g. Cullen and Gruber, 2000), the relationship between DI and spousal labor supply has only recently been analyzed. Spousal labor force participation might be higher in absence of a DI. Using longitudinal data, Chen (2012) shows that spousal labor force participation decreases in the long term as soon as their husbands are granted DI benefits. Using quasi-experimental variation, Duggan et al. (2010) also find a reduction in spousal labor supply if relaxed eligibility criteria induce an increase in the propensity of their husbands to enroll into DI.

The contribution of this study is the combination of three topics which have not yet gained much empirical attention. First, I look at the effect of a partial reduction of DI benefits unconditional of the type of disability. Second, I focus on existing beneficiaries. Third, I observe the labor supply response to this reduction not only from the beneficiary but also from the spouse. The previous literature has focused on the inflow and on increases in benefits. This study contributes in evaluating whether a reduction in benefits leads to symmetric results.

4.2.2 Institutional setting

The Swiss Disability Insurance as part of the Social Insurance system

This section gives an overview of the Swiss Disability Insurance (DI) which is a major part of the Swiss Social Insurance system. It is important to be acquainted with the

particular features of the system as some may also be affected by the revision, which is described in the next section, through spillover effects.

The Swiss DI program is similar to the social security disability insurance (SSDI) program in the United States.[3] Both are mainly financed through payroll taxes and pay benefits not only to the disabled worker, but also to dependent children and in some cases also to the spouse. In Switzerland, there are three conditions for the eligibility for benefits from the DI program: health impairment, working incapacity and a causal relationship between the two (BSV, 2003). The working incapacity has to last for at least one year. Unlike in the US, the Swiss program differentiates between ordinary and extraordinary benefits. In order to be awarded with an ordinary pension, the applicant must have worked at least one year (three years since 2008) in sustainable employment. Extraordinary pensions are granted mainly for individuals with a congenital condition who were never able to work and to contribute to the pension system. Another important difference to most other DI programs is the method the amounts of benefits are calculated, which is dependent on the degree of disability and leads to a partial benefit system. The degree of disability is calculated by comparing hypothetical and reasonable earnings without a disability with those with a disability in any job:

$$\text{Disability degree} = \frac{\text{Earnings without disability} - \text{Reasonable earnings with disability in any job}}{\text{Earnings without disability}}$$

If this ratio is smaller than 40%, the claimant does not receive any benefits. If it is bigger than 70%, she receives a full pension. Between a degree of disability of 40% and 70% the type of pension is leveled in steps of ten percentage points.[4] Table 4.1 illustrates the type of disability pension and minimum and maximum amounts of benefits with respect to the degree of disability. The benefit amounts depend on how many years the person has been insured and on her past earnings. Unlike in the US, DI beneficiaries are allowed to work without losing their benefits. The benefits are reduced only if the degree of disability changes to an extent that the new degree of disability falls into a lower category.

Similar to the supplemental security income (SSI) in the US, there is also a means-tested supplemental income program in Switzerland. If benefits from DI and other

[3] The SSDI program has received most of the attention in the economic literature. For a short description of it, see e.g. Bound and Burkhauser (1999).

[4] These steps imply considerable disincentive effects where a marginal increase in labor supply which would lead to a reduction in the degree of disability by 1 percentage point could yield in a loss of a quarter pension.

Table 4.1: Degree of disability and amounts of benefits

Degree of disability	Type of pension	Amount of monthly pension in CHF	
		minimum	maximum
<40%	no pension	-	-
40-49%	quarter pension	277	553
50-59%	half pension	553	1105
60-69%	three quarter pension	829	1658
≥70%	full pension	1105	2210

Notes: Amounts reported are effective from January 2007. 1 CHF = 1.61 € in January 2007.
Source: BSV (2008b)

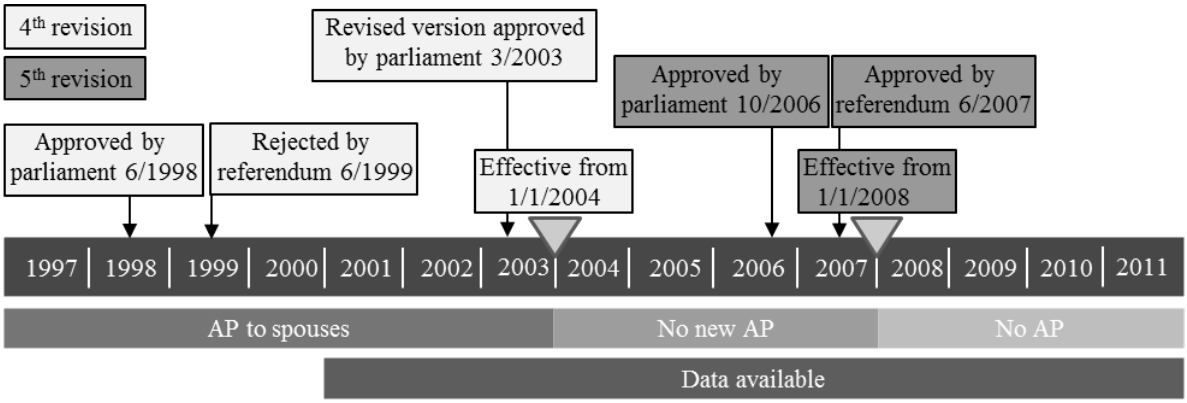
income fail to cover basic living costs, a Swiss resident can apply for *Ergänzungsleistungen* (Supplementary Benefits). Unlike in the SSI, eligibility is conditional on DI benefit.

Together with the Old Age and Survivors Insurance and the Unemployment Insurance, the DI and Supplementary Benefits form the first pillar of the Swiss social insurance. The second pillar is constituted by occupational pension plans and accident insurance. Every employed worker is required to individually contribute to an occupational pension plan through payroll taxes. This occupational insurance account is filled over the workers lifetime and managed by a private insurance company. In case of disability onset or retirement, a monthly benefit is granted, whose amount depends on the stock of funds as well as on the degree of disability assessed by the DI. In addition, a similar system exists for accident insurance. Those who are not employed pay their premiums through health insurance. The goal of policy makers is to achieve a replacement rate of 60% with benefits from the first and second pillar. The third pillar in Switzerland includes voluntary, tax-deductible private insurance plans, whose benefits are granted in the same way as in the second pillar.

The additional pension for spouses and its abolition

Up to the 4^{th} revision of the DI act, the DI awarded an additional pension (AP) to a spouse if the person eligible for an ordinary DI pension was employed prior to the onset of work incapacity and if the spouse is not his-/herself eligible for a DI benefit. The AP was therefore granted unconditionally of the spouse's earnings. The reason for introducing this pension in the first place was the compensation of the part of the earnings which was lost due to the work incapacity by one member of the couple and was destined for

Figure 4.1: Timeline of the abolition of the additional pension (AP)



the subsistence of the conjugal community. The amount of the AP was set to be an additional 30% of the DI benefit, which was on average about 300 CHF per month or about 5% of total household income.

Due to the disproportionally rising number of DI beneficiaries in the nineties, which led to an imbalance in the system, the federal government decided in 1997 to cut mainly on the expenditure side. The policy reform, which was approved by the parliament in June 1998, included the abolition of the quarter pension and no new grants of APs to spouses.[5] Yet, lobbyists were able to request a referendum and the Swiss people disapproved the reform in June 1999 in a vote with an unusual high no-share of 70%. There was wide consensus that the failure of the reform was mainly due to the planned abolition of the quarter pension level. Following the referendum, the federal government adapted the reform accordingly and the parliament passed the law in March 2003. Moreover, the parliament added the introduction of a three-quarter pension. Since no referendum was requested, the reform became effective starting January 1^{st}, 2004. Soon after, the 5^{th} revision was drafted as the imbalance in the system had continued to increase. The parliament discussed the law in 2005-2006, finally passing it in October 2006. Again, a referendum was requested but the Swiss people approved the reform in June 2007 with a yes-share of 59%. The government announced on September 28^{th}, 2007 that the reform was going to be implemented on January 1^{st}, 2008. The reform contained the abolition of all still existing APs for spouses. As a summary, Figure 4.1 illustrates the timeline of the abolition of the AP.

[5]Another incentive to abolish the additional pension for spouses was the fact that it was the only benefit left in social insurance that was conditional on marital status.

Couples affected by the abolition of the AP can close the income gap caused by the abolition by individually or jointly increasing their labor force participation (LFP) on the intensive or extensive margin. If the effect on LFP is strong enough, this may lead to a further reduction of DI benefits (due to the partial DI system) or even an increase in DI outflow. However, the abolition might also cause spillover effects into other parts of the social insurance system.[6]

4.3 Data

The data are administrative covering the full Swiss population of DI beneficiaries between the years 2001 and 2011. Yearly information in December is available for the main beneficiary and the spouse (if any) on DI benefits (including benefits for dependent children), labor market status (employed or self-employed), yearly earnings, and additional social insurance payments (i.e. Supplementary Benefits and Helplessness Allowance). Background characteristics of the main beneficiary include age, age at entry into DI, gender, canton of residence, marital status, number of minor children, nationality, degree and type of disability (four categories: congenital, accident, mental, and other illness). Since neither information on the application process nor on the outsiders (i.e. non DI beneficiaries) is available, any analysis on the inflow into the DI (e.g. benefit denial rate or effect of self-screening) is not possible.

In the analysis I exclude individuals older than 55 at the time of their entry into DI as their labor supply behavior might be affected by the additional option to enter (early) retirement. The statutory retirement age in Switzerland for men is 65, while it changed for women from 63 to 64 in 2005.[7] In addition, I also disregard people living not in Switzerland since they face different labor markets and institutions. The third group excluded are beneficiaries of an extraordinary pension because they may have a significantly different employability. These people receive a pension although they do not fulfill the requirement of having contributed to the system for at least a year because their working incapacity had been limited before they reached the working age. Table 4.A1 in the appendix gives an overview of the sampling.

[6] These parts include supplementary benefits, old-age and unemployment insurance, and benefits from the second pillar.

[7] The reason why many years before the statutory retirement age are excluded lies in the way treatment groups are defined (see next section).

4.4 Identification strategy

The empirical strategy exploits the fact that not every DI beneficiary was affected by the abolition of the additional pension (AP). I can therefore use the unaffected DI beneficiaries to control for the counterfactual situation in a difference-in-differences (DiD) approach. This approach relies heavily on the assumption that both the treatment and control group follow a similar trend so that in the absence of the policy reform, the two groups would evolve in the same way (common trend assumption). This assumption is more plausible the more similar the control group is to the treatment group. How crucial the definition of the control group can be in the field of disability research is shown by Gupta and Larsen (2008, 2010), whose results differ depending on the control group.[8]

As illustrated by Figure 4.1 the AP to spouses was abolished in two steps. I use the first step to define the treatment and control group, while the second step provides before and after observations. A rather naive approach would then be to look at beneficiaries who were married at the beginning of their spell in the DI and compare those who entered DI prior to January 1^{st} 2004 and lose their AP by January 1^{st} 2008, with those who entered DI after January 1^{st} 2004 and have never received any AP. However, beneficiaries on the DI rolls for a decade or more might not only have different background characteristics but also different human capital stocks than beneficiaries who entered DI only recently. This should make clear it is crucial to compare beneficiaries with equally long time on DI rolls. The duration on the DI rolls measured in years is therefore the parameter used to define the treatment and control group.[9] Depending on the duration in DI, beneficiaries with more than four years in DI in 2008 are treated (i.e. they lose the AP), while beneficiaries with less than four years in DI in 2008 are controls (i.e. they have never received an AP). Using this definition I do not observe the same individuals over the years. Instead, I compare in each year different cohorts as they move along the time axis. However, depending on the year of observation group substitution need to be taken into account. This issue can be most conveniently explained by an example, illustrated in Table 4.2.

[8] The authors evaluate the introduction of the Danish Flexjob scheme using a DiD approach. To be eligible for Flexjob, individuals must have a long-term disability (at least three years) and a reduction in working capacity. In the earlier draft (2008) the general population was used as the control group, resulting in only modest employment effects. In the later version (2010), when long-term disabled without reduction in working capacity and short-term disabled individuals (based on self-classification) were used as control groups, the authors find a substantial positive employment effect.

[9] It might seem more straightforward to use the year in which beneficiaries started to draw DI benefits to define treatment and control groups. However, identification then fails due to a non-linear trend (see a note in the appendix for more details).

Table 4.2: Group substitution between treatment and control group

Year of observation	2005	2006	2007	2008	2009	2010
Five years in DI						
Receive AP	2000[1]	2001	2002			
Ever received AP				2003		
Never receive AP					2004	2005
Three years in DI						
Receive AP	2002	2003				
Ever received AP						
Never receive AP			2004	2005	2006	2007

Notes: [1]Year of entry into DI

Consider a beneficiary with five years in the DI observed in the year 2007 which means she entered DI in 2002 and therefore receives an AP in 2007 (upper panel of Table 4.2). One year later in 2008, a beneficiary with five years in the DI (thus entering DI in 2003) loses the AP, which is about a 23% reduction in DI benefits of a household without minor children. In 2009, a beneficiary with five years in DI entered the DI in 2004, so she has never received any AP. Thus, beneficiaries with five years in DI can no longer be used as treated after the year 2008. Similarly for the control group, consider a beneficiary with three years in DI in the lower panel of Table 4.2. Persons fulfilling this requirement can only be used from 2007 and after. In the year 2007 they entered DI in 2004 being the first cohort who never receives an AP. Taking these two groups together, I can only use the years 2007 and 2008 (to observe the outcome variable), rendering for each group one observation prior and one after the policy change. For all other years there is no counterfactual group available (i.e. either both groups receive the AP or both groups have never received the AP).

By altering the time spent in DI I can access more pre- and post-reform years of observation. On the one hand, using beneficiaries with, for example, six years in DI allows me to additionally use the year 2009 as a post-reform year. On the other hand, using beneficiaries with two years in DI opens the year 2006 as an additional pre-reform year. Table 4.A2 in the appendix gives an overview of all possible combinations. Beneficiaries with nine years in DI or more do not increase the number of post-reform years of observation since data from 2012 are not available. As mentioned above these persons might also have very different stocks of human capital and other unobservable characteristics than beneficiaries with a shorter spell in DI. For this reason, all beneficiaries

with more than eight years in DI are excluded from the analysis.

Observations from beneficiaries with two years in the DI and even more so those with one year in DI need special attention. The year in which a beneficiary shows up for the first time in the data may not correspond to the year in which she actually enters DI in the sense of receiving payments. The reason is that payments are made retrospectively up to the time of application which is not observed in the data. The age at entry into DI allows to identify in which year the applicant started to receive benefits. In that sense, observations after one year or even two years are incomplete.[10] As Table 4.A3 in the appendix shows, beneficiaries who appear in the data one or two years after the start of their spell are significantly different in background characteristics from beneficiaries who are observed after three years or later. Therefore, I use beneficiaries with one or two years in DI as a control group in robustness checks only.

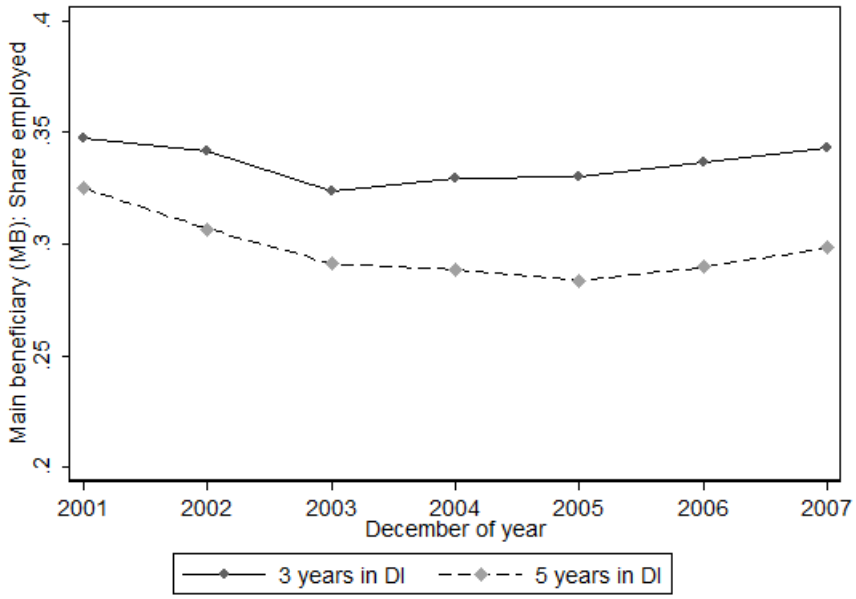
The underlying assumption is that cohorts with equally long time on the DI rolls have common labor market outcomes over time. Although the common trend assumption is not testable, it may gain credibility if a parallel trend prior to the reform can be observed. Figure 4.2 plots the employment rate of the main beneficiaries with five and three years in DI prior to the reform. The two lines decrease similarly in 2002 and 2003 and slowly recover thereafter. The two groups seem to be affected in the same way by the short recession in Switzerland after the burst of the dot-com bubble. Figures 4.A3-4.A6 in the appendix plots the two lines for the other outcomes analyzed. The pattern is very similar and parallel for all outcomes.

Another way to check the plausibility of the common trend assumption is to run placebo tests at a random point in time other than at the time of the implementation of the reform. A significant difference between the treatment and control group after this artificial policy reform would cast serious doubt on the causality of the real reform. The results of this test are discussed in section 4.5.2.

If the placebo date is chosen close to the date of the actual policy change, a significant difference could also be viewed as an anticipation effect. This anticipation effect would violate the assumption of an exogenous treatment. Since I use two revisions, I need to provide evidence that there are no anticipation effects at both thresholds. At the first threshold, I need to rule out that individuals selected into treatment by bringing their DI application forward prior to January 1^{st} 2004 (in order to still benefit from the AP). Although applying for DI benefits is a timely issue as there is a waiting period of at least

[10] There are only few beneficiaries who appear in the data for the first time more than three years after their estimated start of the spell.

Figure 4.2: Employment rate of main beneficiary prior to the policy change

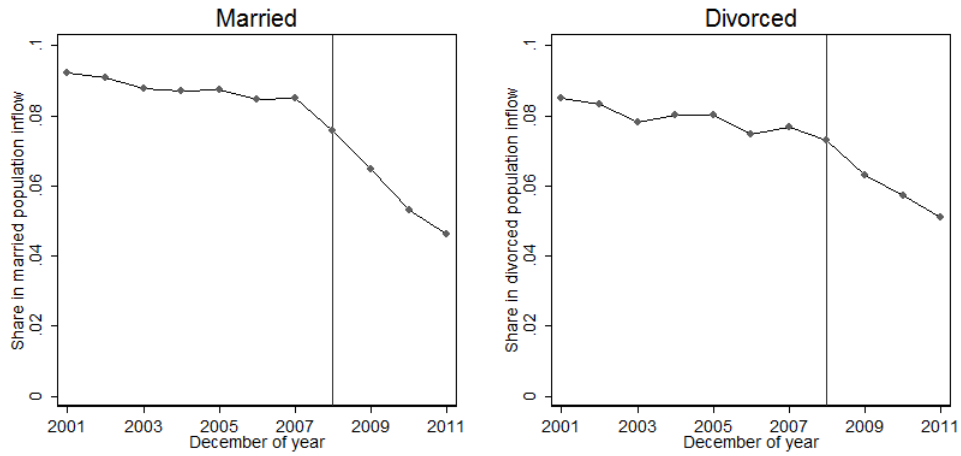


one year, this problem might be relevant since there was a long timespan between the announcement of the 4^{th} revision and its implementation (see Figure 4.1). However, the final revision was adopted only nine months prior to its implementation. If selection into treatment was substantial, a significant spike in the inflow into DI should be observed in the years before 2004 and a drop-off thereafter. Figure 4.3 plots the inflow for two groups over the observable years. The picture on the left shows the inflow of married beneficiaries measured by the share of beneficiaries with five years in DI over the total married population in the DI. If selection into treatment was present, the inflow should increase prior to the year 2008 as people with five years in DI before 2009 entered DI in the regime where they still receive the AP (before 2004). Yet, the inflow decreased from the start of the observation period in 2001 (i.e. inflow in 1996). Although the decline is steeper after the revision, it is rather reflecting a general trend since the decrease is similar in the group of divorced beneficiaries which are not affected by the policy change (see picture on the right of Figure 4.3).

As a second check for selection into treatment, I compare the observable characteristics of beneficiaries who entered DI just prior to the policy change with those who entered just after. If the characteristics significantly differ, this would be a strong sign for an anticipating reaction of some part of the population.[11] Under the assumption that in every

[11]There should not be any anticipating effects with respect to a change in civil status. The AP could

Figure 4.3: Share of inflow into DI of married and divorced beneficiaries with five years in DI



Notes: The y-axis reports the following ratio: $\frac{\text{Married (Divorced) population with 5 years in DI}}{\text{Married (Divorced) population in DI}}$

year people with similar characteristics appear later in the data (which is substantiated by Table 4.A3 in the appendix), Table 4.A4 in the appendix compares beneficiaries who entered DI in 2003 with those who entered one year later, for whom no AP was granted. Except for mental illness, no significant difference between the two groups can be detected. In fact, the share of persons with a mental illness is lower before the policy reform. Under the assumption that suffering from mental illness is easier to pretend than other types of disability, this share should be higher if selection into treatment was present. Similar to Figure 4.3, columns 5-7 report the same analysis for the non-affected group of divorced beneficiaries. Other placebo groups are two groups who both entered DI before the policy change (e.g. 2001 vs. 2002, see column 8) or both after (2005 vs. 2006, see column 9). In all three comparisons there are hardly any significant differences between the groups. Although selection into treatment cannot be fully ruled out, all these analyses reveal that it is not likely to be a relevant confounder of the results.

At the second threshold, there could be anticipatory effects when beneficiaries started to increase their labor supply prior to the 5^{th} revision in anticipation of the reduction in benefits. However, this is not very likely as the transposition date was announced only three months earlier (see section 4.2.2). If there was any anticipation, it would result in a downward bias of the effect anyway.

still be awarded if the marriage was registered after the ordinary DI benefit had been awarded. Therefore, there was no incentive to bring a marriage forward.

Analytically I estimate regressions of the following type:

$$y_{it} = \alpha + \beta Treat_{it} + \gamma Post_t + \delta(Treat_{it} \times Post_t) + X'_{it}\theta + \epsilon_{it} \tag{4.1}$$

where y_{it} is the outcome variable of interest for person i at time t. $Treat$ is an indicator for the treatment group. In the baseline regression, $Treat$ is 1 if the main beneficiary has drawn DI benefits for five to eight years. $Treat$ is 0 if this period is equal to three years. t is then $\in \{2007, 2008\}$. $Post$ is an indicator for the year of observation (1 if 2008 and after; 0 before 2008), while X is a set of covariates, which include age, gender, foreigner status, number of minor children, cantonal dummies, and a dummy for mental illness.[12] The coefficient of interest is always δ which measures the effect of the abolition of APs on beneficiaries relative to beneficiaries who had never drawn an AP using variation over time persons spent on the DI rolls. The main identifying assumption is that there are no unobserved changes between beneficiaries with five to eight years in DI compared to three years in DI which are correlated with the policy change or the outcome variables.

The remaining identifying assumptions for a DiD strategy are the following: (i) If covariates are included in the specification to increase precision, these covariates must be exogenous, i.e. unrelated to the policy reform. The covariates I use are exogenous since most are determined prior to the reform. (ii) Almost all outcome variables have limited support so that the predicted expected potential outcomes could lie outside of it. But in any case, the group difference prior to treatment is never as big as to an extent that the adjustment for common trends would lead to a prediction outside of the support.[13] (iii) This leads to the next concern of the functional form dependency of the DiD approach. The common trend may hold for the level of an outcome variable, but not for a logarithmic transformation of it (Lechner, 2011). Earlier given plausibility checks for the common trend assumption (Figures 4.2 and 4.A3-4.A6) as well as placebo tests in section 4.5.2 need to be viewed with this in mind. (iv) A more fundamental assumption applying to all typical microeconometric evaluation strategies is the stable unit treatment value assumption (SUTVA). The policy reform must not have any general equilibrium effect in the sense that the employment probability of an individual should be independent of the treatment status of other individuals. SUTVA is likely to hold as

[12] Including covariates in a linear fashion as in equation (4.1) implies a more restrictive common trend assumption. However, since the results are very similar with a specification without covariates, a substantial bias is not likely.

[13] Using a nonlinear model may be attractive but it would yield a modified common trend assumption. It may be difficult to argue why the common trend assumption is credible for a nonlinear transformation but not for the linear model.

the policy reform affected only a small group of persons relative to the general population. Although jobs for people with a disability might not be abundant, the persons in the two groups are not geographically clustered, so they are likely to not compete on the same local labor market.

4.5 Results

4.5.1 Main results

Table 4.3 reports summary statistics of the sample used for the baseline evaluation. Each variable's mean is grouped by treatment status and compared before and after January 1^{st} 2008, the date of the policy change. The treatment group includes beneficiaries who have drawn DI benefits for five to eight years, while members of the control group have spent three years on the DI rolls. Panel A shows the effect of the revision on various outcome variables. The average monthly amount of the AP, which was canceled, is 292 CHF.[14] While the amount of the ordinary DI benefit and in combination the degree of disability also decreased significantly in the treatment group, these two outcomes decreased in a similar way in the control group. Both the employment rate and earnings of the main beneficiary and the spouse increased in both groups. The abolition of the AP might induce spillover effects into other social insurances which might have an effect on the decision whether to change the labor supply. If the couple passes a means test, it can request supplementary benefits (SB). Some DI beneficiaries might additionally qualify for helplessness allowances (HA) if they need assistance to perform activities for daily living. However, comparing the two groups before and after the policy change the same pattern as for the employment effect can be observed for the additional social insurance payments. As a side effect it is interesting to see whether the abolition of the AP led to an increase in the dissolution of marriages. However, the separation rate is rather constant over the period for both groups.

Panel B of Table 4.3 reveals whether the groups differ in observable background characteristics. Remember that every of the four columns represents different individuals. Except for the fact that the share of people with a mental illness has significantly increased in 2008 in the treatment group compared to the control group, changes between the two periods in all other characteristics are not significant.

[14]Note that 7% of beneficiaries in the control group also received an AP in 2007. The reason lies in the fact that in the administrative data the age at entry into DI is not precise but estimated only.

Table 4.3: Descriptive statistics by treatment status

	Treated (5-8 years in DI)		Control (3 years in DI)		
	Before 2007	After 2008	Before 2007	After 2008	Difference in Differences
A. Outcomes					
Monthly amount of AP in CHF	292.3	0***	28.26	0***	-264.1†††
Monthly amount of DI in CHF	1353	1300***	1269	1220***	-3.658
Degree of disability	76.56	76.01***	74.16	73.73	-0.130
Outflow from DI	0.036	0.038	0.019	0.024*	0.003
Main beneficiary (MB): Employed	0.277	0.282	0.343	0.359	-0.010
MB: Yearly earnings in CHF	6518	6709	8441	9403**	-771.8†
Spouse: Employed	0.646	0.659***	0.754	0.758	0.010
Spouse: Yearly Earnings in CHF	35473	36442**	40684	40921	733.1
Household (HH): Share receiving Supplementary Benefits	0.166	0.181***	0.137	0.146	0.006
HH: Yearly amount of Supplementary Benefits in CHF	3191	3439***	2718	2867	98.63
HH: Share receiving Helplessness Allowance	0.045	0.048	0.051	0.051	0.003
HH: Yearly amount of Helplessness Allowance in CHF	34.40	35.91	37.01	38.32	0.206
No longer married	0.067	0.067	0.029	0.026	0.003
B. Covariates					
Age	51.78	51.90*	48.90	48.93	0.086
Female	0.441	0.442	0.427	0.419	0.010
Foreigner	0.432	0.434	0.428	0.426	0.004
Number of minor children	0.729	0.720	0.883	0.891	-0.016
Mental illness	0.333	0.346***	0.378	0.369	0.022††
N	30,024	29,248	5341	4290	68,903

Notes: All amounts in CHF are adjusted for inflation. ***,**,*: statistically different from *Before* at 1, 5 and 10 percent, respectively. †††,††,†: statistically different at 1, 5 and 10 percent, respectively.

I now turn to the regression framework outlined in section 4.4. Results from the OLS estimation of equation (4.1) are summarized in Table 4.4. The dependent variable is the employment rate of the main beneficiary. The top panel of the table reports the results from the baseline regression where treatment and control groups are defined as in Table 4.3. Remember from Table 4.2 that in this setup only one post reform year can be observed. If beneficiaries with five years in the DI are excluded from the treatment group, an additional post reform year becomes available. The first column of the second panel of Table 4.4 therefore repeats the same regression for the diminished sample as a consistency check, while the second column now reports the result for the post period of 2008 and 2009. The third and fourth panel decrease the treatment group each by another year, so even more long-term effects can be analyzed. All regressions include the covariates mentioned, which are not reported due to space constraints. Overall, the effect on the employment rate of the main beneficiary is close to zero, which is rather consistent for the different treatment groups and over time.

Although there is no effect on the employment rate on the extensive margin, effects on earnings may partly reflect changes on the intensive margin. Table 4.A5 in the appendix therefore reports the same analysis for yearly earnings as the dependent variable.[15] Again the results are consistent and not significantly different from zero.

As a result of the partial income loss of the household due to the abolition of the AP, the healthier member of the household may increase labor supply rather than the disabled member. Tables 4.A6 and 4.A7 in the appendix report the results for the employment rate and earnings of the spouse, respectively. The baseline employment rate is more than twice as high for spouses than for the main beneficiary (about 64% vs. 27%). However, there is again no significant difference between households who lost their AP and those who never had one. The same pattern can be observed for earnings.

If both members of the household did not react to the partial benefit cut by altering labor supply, households may react by applying for supplementary benefits (SB). SB are granted if the household passes a means test. Table 4.A8 in the appendix shows the effect on the share of DI beneficiaries additionally drawing SB. While the change in this share is close to zero in all different treatment groups in the year after the abolition, the effect is stronger and becomes significant in the years after 2008 as the screening of the application may need time.[16] Comparing these effects to the baseline share before the

[15] Earnings are set to 0 if the person is not working. Results are similar when only persons who work are considered.

[16] Statistical significance is not upheld when only beneficiaries with eight years in DI are used as the treatment group but the size of the effect is hardly reduced.

Table 4.4: Impact on employment rate of main beneficiary

Dependent variable:	Number of years post reform			
MB: Employed	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Treated: 5-8 years in DI - Control: 3 years in DI				
Treat × Post	-0.007			
	(0.010)			
Treat	-0.072***			
	(0.007)			
Post	0.014			
	(0.009)			
Mean (Treated)	0.277			
N	68,903			
R^2	0.090			
Treated: 6-8 years in DI - Control: 3 years in DI				
Treat × Post	-0.006	-0.009		
	(0.010)	(0.009)		
Treat	-0.082***	-0.082***		
	(0.007)	(0.007)		
Post	0.014	0.014*		
	(0.009)	(0.008)		
Mean (Treated)	0.269			
N	54,475	79,491		
R^2	0.093	0.089		
Treated: 7-8 years in DI - Control: 3 years in DI				
Treat × Post	-0.006	-0.009	-0.014	
	(0.011)	(0.009)	(0.009)	
Treat	-0.090***	-0.090***	-0.090***	
	(0.007)	(0.007)	(0.007)	
Post	0.014	0.014*	0.018**	
	(0.009)	(0.008)	(0.008)	
Mean (Treated)	0.264			
N	39,349	57,628	74,494	
R^2	0.095	0.090	0.086	
Treated: 8 years in DI - Control: 3 years in DI				
Treat × Post	-0.007	-0.011	-0.015	-0.016*
	(0.012)	(0.010)	(0.010)	(0.009)
Treat	-0.097***	-0.096***	-0.096***	-0.096***
	(0.008)	(0.008)	(0.008)	(0.008)
Post	0.014	0.014*	0.019**	0.015**
	(0.009)	(0.008)	(0.008)	(0.007)
Mean (Treated)	0.262			
N	24,180	35,008	45,247	54,302
R^2	0.097	0.094	0.089	0.085

Notes: All regressions include covariates (age, gender, foreigner status, number of minor children, cantonal dummies, and a dummy for mental illness). Number of pre-reform years is always one (2007). Robust standard errors in parentheses. ***,**,*: statistically different at 1, 5 and 10 percent, respectively.

policy change of about 17%, they are sizeable. The increase of 1.5 percentage points corresponds to an increase in the share of about 9%. There is no effect (results not reported) on helplessness allowance, which is expected as applicants in order to be eligible for it, would need to prove a sudden change in health status newly requiring assistance in performing daily activities. There could be more spillover effects into other social insurances, for which information in the data is lacking. Potential concerns from this missing data issue are discussed in the next section.

The revision might have different effects depending on the background characteristics of the beneficiaries. Moral hazard might be more present in hard-to-diagnose cases like mental illness or for beneficiaries with only a partial disability. To check for effect heterogeneities I split the sample into two groups estimating the treatment effect for them separately. However, there is no effect heterogeneity neither between people with a full benefit and persons with a degree of disability below 70% nor between beneficiaries with or without a mental illness on any of the outcome variables considered (results are not reported). There are also no significant differences between beneficiaries below and above the median income. The only characteristic for which consistently different effects are observed is the share of beneficiaries receiving SB of foreigners compared to non-foreigners. Table 4.A9 in the appendix replicates Table 4.A8 for foreigners and non-foreigners. Although the effects are not significantly different between the two groups, the increase in the share seems to be almost entirely driven by foreigners. A possible explanation may be that foreigners do not have the same social network as natives on which they can rely on if they risk of falling into poverty.

4.5.2 Robustness checks

The results so far may be driven by the particularity of the control group which is restricted to include beneficiaries with three years in the DI. As mentioned in section 4.4, beneficiaries with one or two years in the DI have been excluded because they seem to differ significantly in background characteristics from later inflows. However, as a robustness check Table 4.A10 in the appendix reports employment effects from regressions with control groups which include beneficiaries with two to three or one to three years in DI. For the sake of comparison, the first column reports the baseline estimate from Table 4.4 when only beneficiaries with three years in DI are used as a control group. Remember that increasing the number of years in DI in the control group allows for looking at more than one pre-reform year. The effects are very similar for both control groups and with more pre-reform years. Results with other treatment groups (six to eight, seven to eight,

and eight years in DI) are not reported as the results are very similar as well. Table 4.A11 repeats the same exercise for the share of beneficiaries with supplementary benefits. The effects are slightly bigger and significant if the control group is increased and more pre-reform years are included.

As mentioned in section 4.4, placebo tests, either with an unaffected group or at a random point in time other than at the time of the implementation of the reform, may help to substantiate the plausibility of the identifying assumptions of the DiD approach. If beneficiaries who are for a shorter time on the DI rolls compared to those who are on the DI rolls for a slightly longer period have different outcomes at another point in time or from an unaffected group, the difference between the years 2007 and 2008 might not be caused by the revision but is rather reflecting a general trend. Table 4.A12 in the appendix reports placebo treatment effects for a policy change that would have abolished all remaining AP by January 1^{st}, 2006 using various treatment groups. The effects are not significantly different from zero for any of the three outcomes (employment rate of main beneficiary and spouse, share with SB). Note that only one post-reform year (i.e. 2006) can be observed because more years would interfere with the true reform. Members of the "control" group with three years in DI in 2007 would differ from previous years since they would have never had access to an AP.[17] This issue does not evolve when using another group which was not affected by the policy change, namely divorced beneficiaries.[18] The upper (lower) panel of Table 4.A13 in the appendix therefore reports treatment effects as in Table 4.4 (4.A8) but for divorced beneficiaries so that neither the treatment nor the control group has ever drawn an AP. The effects for both outcomes (employment rate and share with SB) switch sign but are not significant. Since divorced beneficiaries have no spouse by definition, I cannot run placebo regressions for the outcomes of the spouse.

Another concern stemming from the way the treatment and control group are defined may be selective outflow. Beneficiaries who survive in the DI for more than five years may be different from those who leave the DI rolls earlier. The data do not allow to control for the reason why a person leaves the DI. The sampling of disregarding persons who entered DI after the age of 55, however, excludes the main reason for leaving the DI, which is reaching the statutory retirement age. In addition, the death rate is likely to be similar between the two groups. Statistics from the Swiss Federal Social Insurance Office reveal outflow for other reasons than death or retirement is very low and stable in the observation period at about 1% (BSV, 2008a). Figures 4.A7 and 4.A8 in the

[17]Note that in this placebo test both groups should receive an AP at all times.
[18]Divorced beneficiaries are more similar in background characteristics than singles.

appendix show the yearly outflow rate (i.e. probability to exit in the next year) and the cumulated outflow rate in the pre-reform period for the two groups, respectively. While the yearly outflow rates are very similar between the groups, the cumulative outflow rate of beneficiaries with five years in DI is higher by definition but the two evolve very parallel over time. In addition, Table 4.A14 in the appendix checks whether the background characteristics of beneficiaries leaving DI after three years are different from those of beneficiaries leaving DI after five to eight years. The probability to enter DI increases by age and Table 4.3 has already revealed that members of the treatment group are older. Since the administrative data only record benefits to minor children, the children of older beneficiaries have more often reached majority age and are no longer recorded. The other characteristics do not significantly differ between the two groups.

The two revisions of the DI act included other changes which could potentially interfere with the analysis of the abolition of the AP. The 4^{th} revision included a duplication of Helplessness Allowance as well as the introduction of medical screening services and of the three-quarter pension. Helplessness Allowance is granted when a person needs help in performing daily activities. Since the amount was also duplicated for existing beneficiaries, both treatment and control group are affected by this change in the same way. The medical screening services may have an effect on the inflow after its introduction. The introduction of the three-quarter pension, where some beneficiaries received more benefits while some received less, could have an effect on the inflow and on the labor force participation. However, while differences in inflow have already been analyzed (see Table 4.A4), it is highly unlikely that differences in labor supply would evolve exactly four years after the revision and therefore coincide with the 5^{th} revision. Concerning other changes within the 5^{th} revision, additional incentives for reintegration into the labor market were introduced. Yet, these incentives are available unconditional on the time in DI so that treatment and control group were affected by this change in the same way. The remaining changes applied only to the inflow into DI, which is no longer an issue when focussing on existing beneficiaries.

Spillover effects of the revision into other social insurances have already been tested for which the data provides information, namely Supplementary Benefits and Helplessness Allowances. If the DI beneficiary receives an additional benefit from an occupational or accidental insurance plan from the second pillar, the amount of this pension increases after the revision if it has been previously capped to prevent overcompensation. Such an increase would potentially inhibit an effect on labor supply. The data do not include any information on benefits from the second pillar of social insurance. Information from the

Swiss Household Panel, a representative data set of the Swiss population, reveals that about a third of all DI beneficiaries draw an additional pension from an occupational insurance (OI) plan. To get a sense of how many OI benefits increased due to the abolition of the AP, I contacted two insurance companies which had approximately 350 beneficiaries with an AP from DI listed on their rolls by January 1^{st}, 2008. Since no single benefit has been revised, significant confounding effects through the second pillar are very implausible. Spillover effects into unemployment insurance (UI) can not be assessed with the data, either. However, beneficiaries need to have contributed to the system prior to claiming benefits. Therefore it is also rather unlikely that the inflow into UI considerably increased for the beneficiaries losing the AP.

4.6 Conclusion

A partial benefit reduction for existing DI beneficiaries in Switzerland has not been successful in raising neither the employment rate nor the earnings of married beneficiaries. Similarly, there is no reaction in employment or earnings from the other healthy member of the married couple. If anything, this policy change may have led to spillover effects into other social insurances as some beneficiaries newly qualify for additional means-tested supplementary benefits.

These results are in contrast to the results from studies which look at the effect of tighter eligibility criteria on applicants and find significant increases in employment rates. The reason why the effects are not symmetric for existing beneficiaries may be twofold: (i) The beneficiaries' stock of human capital cannot be maintained on a level to still find employment after some years on the DI rolls. (ii) But the fact that not even the healthy person of the couple reacts may call for a different explanation. For the spouse the partial benefit cut is a pure income effect. DI benefits of the couple are unaffected by an increase in the labor supply of the spouse. These couples may thus value leisure much more than their working peers. Further research is needed to understand the mechanisms which drive the behavior of existing beneficiaries and of their spouses.

At this stage, the policy implications from this study are in line with the recommendations from Autor and Duggan (2010) for the United States. Autor and Duggan argue that the DI system should provide services with the goal of enabling people with disabilites to remain in employment, keeping them away from entering DI. While these services may not come at a low cost, they might pay off in the future in the light of lacking effective measures to decrease the expensive stock of DI beneficiaries.

4.7 Appendix

Alternative definition of treatment and control group

In section 4.4 the duration on the DI rolls was introduced as the parameter to separate the treatment from the control group. It might seem more straightforward to use instead the year in which the person started to draw a DI benefit. In that sense, members of the treatment group would be beneficiaries who entered DI just before 2004, while members of the control group would be beneficiaries who entered DI just after 2004. These persons would then be followed over the years (not a different person in every year). Figure 4.A1 plots the employment rates for the treatment (entered DI in 2003) and control group (entered DI in 2005) over time. This definition might look attractive as the two lines converge after 2007 when the treatment group loses the AP. However, as Figure 4.A2 demonstrates, the effect which would be attributed to the reform is only a non-linear time trend. In Figure 4.A2 the variable on the x-axis (December of year in Figure 4.A1) is replaced by the years in DI. The employment rates of DI beneficiaries decrease in a similar non-linear way over time for both groups. This pattern is similar when other years are used for the definition of the groups.

Figure 4.A1: Employment rates of alternative treatment and control group

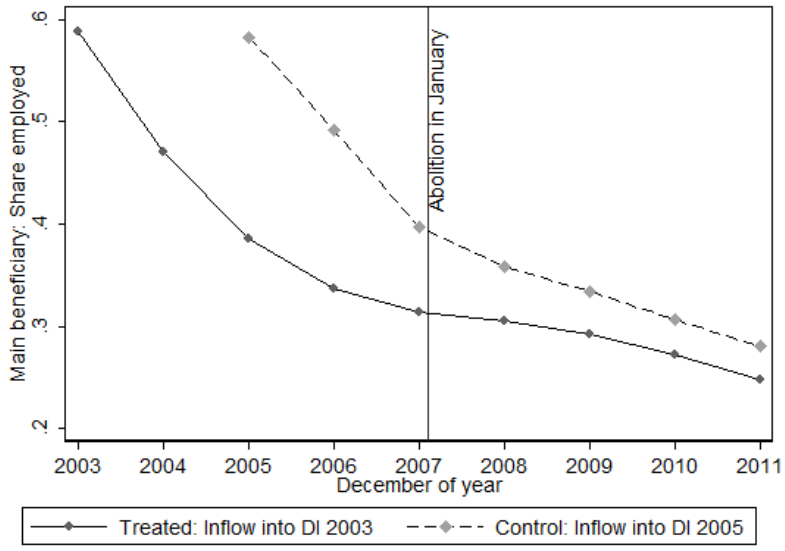


Figure 4.A2: Employment rates of alternative treatment and control group with alternative x-axis

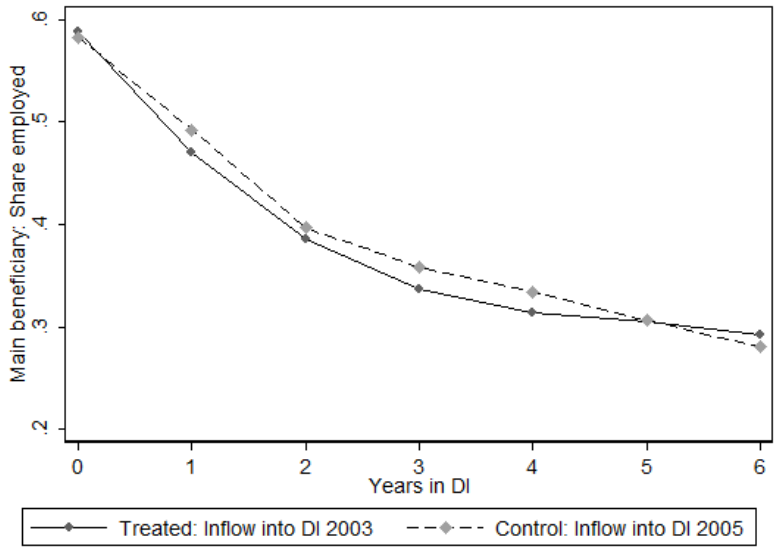


Figure 4.A3: Earnings of main beneficiary prior to the policy change

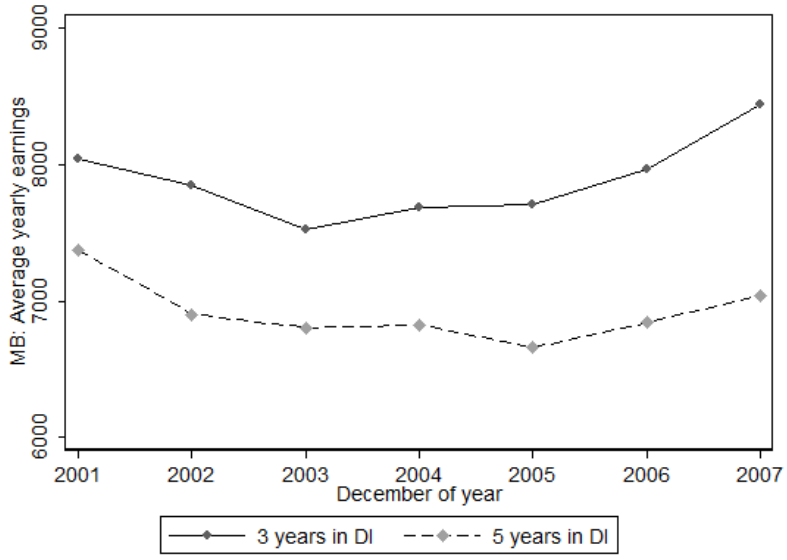


Figure 4.A4: Employment rate of spouse prior to the policy change

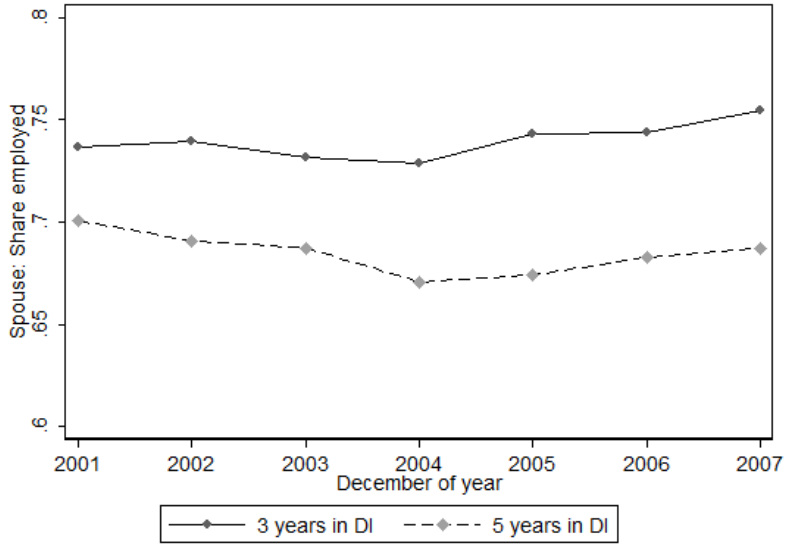


Figure 4.A5: Earnings of spouse prior to the policy change

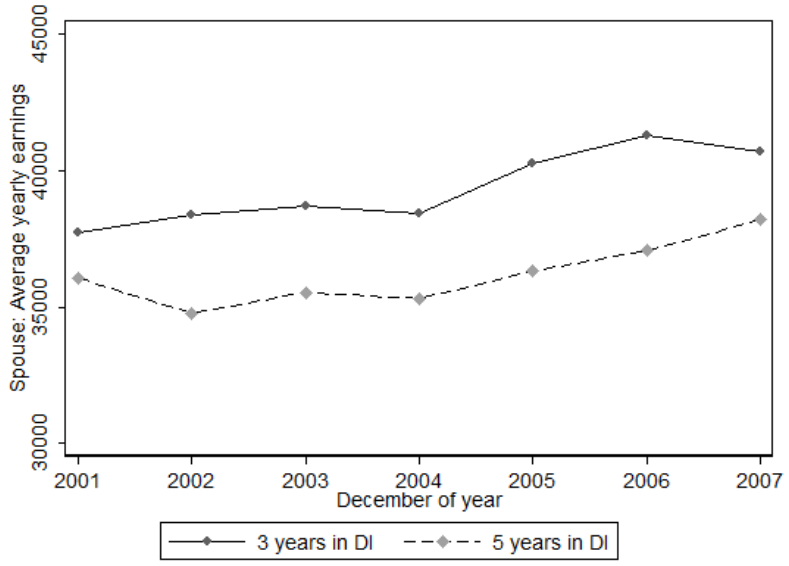


Figure 4.A6: Household's share with supplementary benefits prior to the policy change

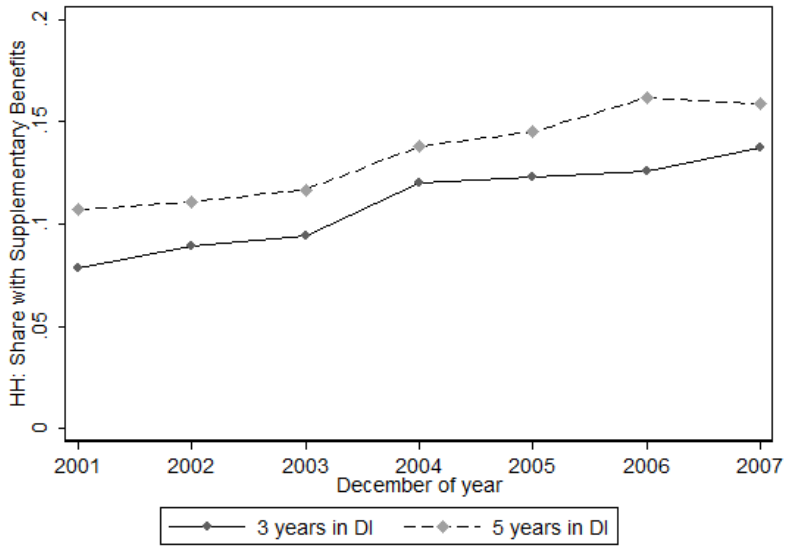
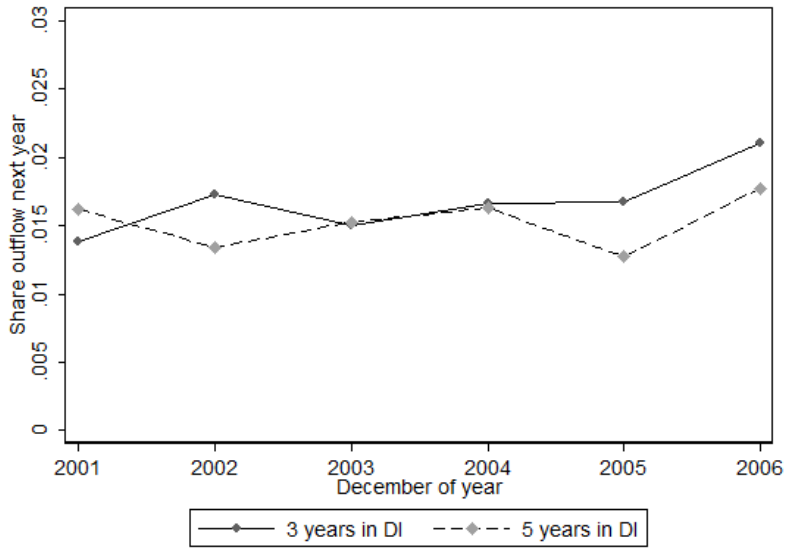
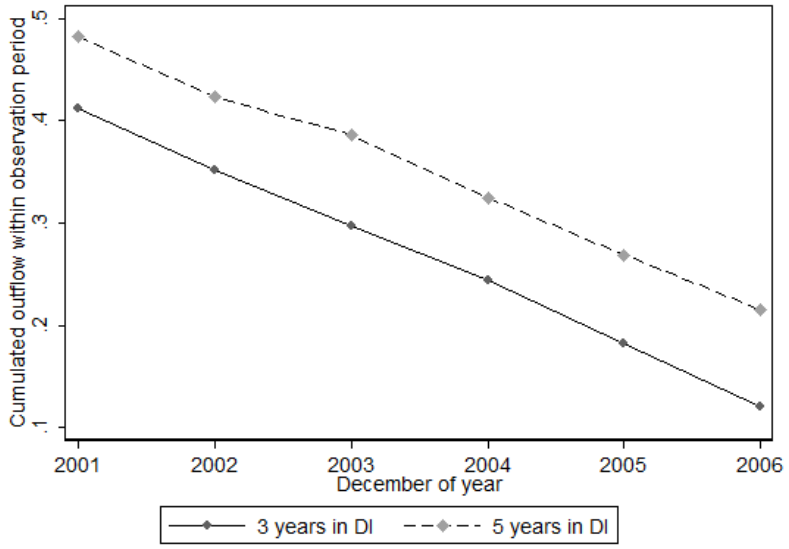


Figure 4.A7: Probability to exit DI in the next year



Notes: Ordinary retirement is excluded as a reason for exit from DI.

Figure 4.A8: Cumulated outflow over observation period



Notes: Ordinary retirement is not excluded as a reason for exit from DI.

Table 4.A1: Sampling of DI beneficiaries

Beneficiaries entering DI	before 2004	2004-2007
Total	346,604	74,714
Non-married	177,358	37,509
Married	169,246	37,205
Older than 55	31,239	13,695
55 and younger	138,007	23,510
Living abroad	28,966	3,476
Living in Switzerland	109,041	20,034
Extraordinary Benefits	1,765	59
Ordinary Benefits	107,276	19,975

Notes: All characteristics are measured in the year of first appearance in the data.

Table 4.A2: All possible treatment and control groups

Years in DI	**Year of observation**	**2004**	**2005**	**Before 2006**	**2007**	**After 2008**	**2009**	**2010**	**2011**
8	Start year of DI	1996	1997	1998	1999	2000	2001	2002	2003
7		1997	1998	1999	2000	2001	2002	2003	
6		1998	1999	2000	2001	2002	2003		
5		1999	2000	2001	2002	2003			
4		2000	2001	2002			2005	2006	2007
3					2004	2005	2006	2007	2008
2				2004	2005	2006	2007	2008	2009
1			2004	2005	2006	2007	2008	2009	2010

Table 4.A3: Background characteristics by time of appearance in the data

Year of inflow	2005			2006			2004	2003	2002
Year of first appearance	3 years or later	within 2 years	Difference	3 years or later	within 2 years	Difference	Difference	Difference	Difference
Age at entry	43.51	45.33	1.83***	43.72	45.12	1.39***	2.04***	1.65***	3.11***
Age	48.24	47.45	-0.800***	48.33	47.35	-0.979***	-0.841***	-1.33***	-0.793***
Female	0.434	0.411	-0.024	0.421	0.417	-0.003	0.014	0.023	-0.060***
Foreigner	0.470	0.424	-0.046***	0.443	0.410	-0.032*	-0.064***	-0.057***	-0.044**
Number of minor children	0.950	0.949	-0.001	0.964	0.987	0.024	-0.049	-0.024	-0.051
Mental illness	0.391	0.347	-0.043***	0.392	0.367	-0.025	-0.032**	-0.020	-0.060***
N	1372	3836		1046	3302				

Notes: ***,**,*: statistically different at 1, 5 and 10 percent level, respectively.

Table 4.A4: Descriptive statistics by year of entry into DI

Year of entry and year of observation								
	Married			Divorced			Married	
	2003	2004	Difference	2003	2004	Difference	2001 vs. 2002 Difference	2005 vs. 2006 Difference
Age	45.93	45.77	-0.159	47.54	47.24	-0.306	0.311	-0.257
Female	0.402	0.415	0.013	0.604	0.585	-0.019	-0.003	-0.008
Foreigner	0.402	0.415	0.013	0.182	0.194	0.012	0.031	-0.035
Number of minor children	0.929	0.920	-0.009	0.480	0.534	0.055	0.032	0.031
Mental illness	0.286	0.328	0.042**	0.466	0.475	0.009	0.015	0.015
Main Beneficiary (MB): Employed	0.599	0.567	-0.033	0.610	0.594	-0.159	0.039**	-0.029
Spouse: Employed	0.814	0.817	0.002	-	-	-	0.014	0.000
MB: Yearly earnings	14665	16511	1845	14711	15043	331.4	980.8	-3527
Spouse: Yearly earnings	41802	41639	-163.8	-	-	-	372.9	2009
MB: Monthly DI benefit	1292	1284	-8.08	1380	1368	-11.6	-13.2	56.7*
Degree of disability	78.24	77.48	-0.764	78.22	77.50	-0.724	0.614	1.41
Share with supplementary benefits	0.025	0.026	0.002	0.098	0.078	-0.020	0.002	0.003
Share with helplessness allowance	0.041	0.040	-0.001	0.019	0.027	0.008	-0.007	-0.000
N	1302	1112		369	335		1364/1359	904/515

Notes: **,*: statistically different at 5 and 10 percent level, respectively.

Table 4.A5: Impact on earnings of main beneficiary

Dependent variable:	Number of years post reform			
MB: Yrly Earnings	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Treated: 5-8 years in DI - Control: 3 years in DI				
Treat × Post	-599.9			
	(411.7)			
Treat	-2126***			
	(264.8)			
Post	875.9**			
	(386.9)			
Mean (Treated)	6518			
N	68,903			
R^2	0.069			
Treated: 6-8 years in DI - Control: 3 years in DI				
Treat × Post	-600	-395.0		
	(418.5)	(356.1)		
Treat	-2376***	-2387***		
	(273.0)	(271.9)		
Post	878.7**	613.1*		
	(386.9)	(327.3)		
Mean (Treated)	6339			
N	54,475	79,491		
R^2	0.070	0.067		
Treated: 7-8 years in DI - Control: 3 years in DI				
Treat × Post	-547.3	-373.6	-557.3	
	(431.8)	(366.7)	(346.7)	
Treat	-2719***	-2714***	-2731***	
	(281.2)	(280.2)	(279.7)	
Post	885.7**	619.7***	790.7**	
	(386.7)	(327.3)	(309.9)	
Mean (Treated)	6116			
N	39,349	57,628	74,494	
R^2	0.073	0.070	0.070	
Treated: 8 years in DI - Control: 3 years in DI				
Treat × Post	-407.4	-289.5	-436.8	-349.1
	(473.0)	(403.0)	(380.2)	(364.1)
Treat	-3065***	-3048***	-3057***	-3064***
	(315.7)	(313.2)	(312.1)	(311.3)
Post	880.0**	615.4*	787.7**	575.8*
	(386.1)	(327.2)	(380.2)	(296.0)
Mean (Treated)	5918			
N	24,180	35,008	45,247	54,302
R^2	0.078	0.073	0.072	0.071

Notes: See Table 4.4. ***,**,*: statistically different at 1, 5 and 10 percent, respectively.

Table 4.A6: Impact on employment rate of spouse

Dependent variable:	Number of years post reform			
Spouse: Employed	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Treated: 5-8 years in DI - Control: 3 years in DI				
Treat × Post	0.011			
	(0.009)			
Treat	-0.117***			
	(0.006)			
Post	0.004			
	(0.009)			
Mean (Treated)	0.646			
N	68,903			
R^2	0.056			
Treated: 6-8 years in DI - Control: 3 years in DI				
Treat × Post	0.011	0.006		
	(0.010)	(0.008)		
Treat	-0.133***	-0.134***		
	(0.007)	(0.007)		
Post	0.004	0.011		
	(0.009)	(0.007)		
Mean (Treated)	0.631			
N	54,475	79,491		
R^2	0.059	0.060		
Treated: 7-8 years in DI - Control: 3 years in DI				
Treat × Post	0.011	0.008	0.012	
	(0.010)	(0.009)	(0.008)	
Treat	-0.151***	-0.152***	-0.152***	
	(0.007)	(0.007)	(0.007)	
Post	0.004	0.011	0.009	
	(0.009)	(0.007)	(0.007)	
Mean (Treated)	0.615			
N	39,349	57,628	74,494	
R^2	0.065	0.064	0.064	
Treated: 8 years in DI - Control: 3 years in DI				
Treat × Post	0.007	0.007	0.011	0.019**
	(0.012)	(0.010)	(0.010)	(0.009)
Treat	-0.162***	-0.163***	-0.164***	-0.164***
	(0.008)	(0.008)	(0.008)	(0.008)
Post	0.004	0.011	0.009	-0.000
	(0.009)	(0.008)	(0.007)	(0.007)
Mean (Treated)	0.605			
N	24,180	35,008	45,247	54,302
R^2	0.071	0.071	0.069	0.067

Notes: See Table 4.4. ***,**: statistically different at 1 and 5 percent, respectively.

Table 4.A7: Impact on earnings of spouse

Dependent variable:	Number of years post reform			
Spouse: Yrly Earnings	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Treated: 5-8 years in DI - Control: 3 years in DI				
Treat × Post	730.5			
	(906.7)			
Treat	-6207***			
	(591.0)			
Post	361.7			
	(830.5)			
Mean (Treated)	35473			
N	68,903			
R^2	0.130			
Treated: 6-8 years in DI - Control: 3 years in DI				
Treat × Post	522.2	-414.2		
	(916.0)	(797.9)		
Treat	-7240***	-7252***		
	(606.8)	(605.9)		
Post	352.0	1482**		
	(830.8)	(722.9)		
Mean (Treated)	34542			
N	54,475	79,491		
R^2	0.144	0.139		
Treated: 7-8 years in DI - Control: 3 years in DI				
Treat × Post	456.0	-569.1	-143.9	
	(954.7)	(829.7)	(777.1)	
Treat	-8111***	-8193***	-8209***	
	(637.9)	(636.3)	(635.6)	
Post	348.1	1486*	1360**	
	(829.6)	(722.5)	(672.5)	
Mean (Treated)	33641			
N	39,349	57,628	74,494	
R^2	0.148	0.143	0.135	
Treated: 8 years in DI - Control: 3 years in DI				
Treat × Post	776.9	-622.9	-238.5	227.9
	(1063)	(925.3)	(866.8)	(837.1)
Treat	-8771***	-8831***	-8928***	-8912***
	(730.6)	(727.2)	(725.3)	(724.2)
Post	365.5	1504**	1376**	980.4
	(826.5)	(719.7)	(670.2)	(644.5)
Mean (Treated)	32942			
N	24,180	35,008	45,247	54,302
R^2	0.155	0.151	0.148	0.144

Notes: See Table 4.4. ***,**,*: statistically different at 1, 5 and 10 percent, respectively.

Table 4.A8: Impact on share of supplementary beneficiaries

Dependent variable:	Number of years post reform			
Share SB	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Treated: 5-8 years in DI - Control: 3 years in DI				
Treat × Post	0.006			
	(0.008)			
Treat	0.034***			
	(0.005)			
Post	0.008			
	(0.007)			
Mean (Treated)	0.166			
N	68,903			
R^2	0.066			
Treated: 6-8 years in DI - Control: 3 years in DI				
Treat × Post	0.007	0.012*		
	(0.008)	(0.007)		
Treat	0.037***	0.036***		
	(0.005)	(0.005)		
Post	0.008	0.010*		
	(0.007)	(0.006)		
Mean (Treated)	0.168			
N	54,475	79,491		
R^2	0.067	0.068		
Treated: 7-8 years in DI - Control: 3 years in DI				
Treat × Post	0.009	0.015**	0.014**	
	(0.008)	(0.007)	(0.007)	
Treat	0.038***	0.037***	0.037***	
	(0.006)	(0.006)	(0.006)	
Post	0.009	0.011*	0.015***	
	(0.007)	(0.006)	(0.006)	
Mean (Treated)	0.168			
N	39,349	57,628	74,494	
R^2	0.069	0.069	0.070	
Treated: 8 years in DI - Control: 3 years in DI				
Treat × Post	0.003	0.011	0.011	0.007
	(0.009)	(0.008)	(0.008)	(0.007)
Treat	0.041***	0.041***	0.041***	0.041***
	(0.006)	(0.006)	(0.006)	(0.006)
Post	0.008	0.010*	0.014**	0.022***
	(0.007)	(0.006)	(0.006)	(0.005)
Mean (Treated)	0.172			
N	24,180	35,008	45,247	54,302
R^2	0.070	0.070	0.071	0.072

Notes: See Table 4.4. ***,**,*: statistically different at 1, 5 and 10 percent, respectively.

Table 4.A9: Effect heterogeneities with respect to foreigners

Dependent variable:	Number of years post reform			
HH: Share with SB	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
Foreigners				
Treated: 5-8 years				
Treat × Post	0.018			
	(0.013)			
Mean (Treated)	0.241			
N	29,754			
Treated: 6-8 years				
Treat × Post	0.020	0.027**		
	(0.014)	(0.012)		
Mean (Treated)	0.245			
N	23,482	32,932		
Treated: 7-8 years				
Treat × Post	0.023	0.030**	0.023**	
	(0.014)	(0.013)	(0.012)	
Mean (Treated)	0.248			
N	16,882	23,776	29,991	
Treated: 8 years				
Treat × Post	0.023	0.034**	0.028**	0.020
	(0.016)	(0.014)	(0.014)	(0.013)
Mean (Treated)	0.248			
N	10,337	14,407	18,167	21,441
Non-Foreigners				
Treated: 5-8 years				
Treat × Post	0.001			
	(0.009)			
Mean (Treated)	0.109			
N	39,149			
Treated: 6-8 years				
Treat × Post	0.002	0.005		
	(0.009)	(0.008)		
Mean (Treated)	0.110			
N	30,993	46,559		
Treated: 7-8 years				
Treat × Post	0.002	0.007	0.011	
	(0.009)	(0.008)	(0.007)	
Mean (Treated)	0.109			
N	22,467	33,852	44,503	
Treated: 8 years				
Treat × Post	-0.010	-0.003	0.004	0.002
	(0.010)	(0.009)	(0.008)	(0.008)
Mean (Treated)	0.117			
N	13,843	20,601	27,080	32,861

Notes: Every estimate represents the coefficient of Treat × Post from a separate regression including covariates. The control group always includes beneficiaries with 3 years in DI, so that the number of pre-reform years is always one (2007). Robust standard errors in parentheses. **,*: statistically different at 5 and 10 percent, respectively.

Table 4.A10: Sensitivity to different control groups: Employment MB

Dependent variable:		Number of years pre reform		
MB: Employed	1 (2007)	1 (2007)	2 (2006-07)	3 (2005-07)
Treated: 5-8 years in DI				
Control:	**3 years**	**2-3 years**		
Treat × Post	-0.007	-0.005	-0.005	
	(0.010)	(0.008)	(0.007)	
Treat	-0.072***	-0.093***	-0.092***	
	(0.007)	(0.005)	(0.004)	
Post	0.014	0.012	0.016***	
	(0.009)	(0.007)	(0.006)	
Mean (Treated)	0.277			
N	68,903	75,794	116,791	
R^2	0.090	0.093	0.092	
Control:		**1-3 years**		
Treat × Post		-0.012	-0.009	-0.006
		(0.007)	(0.006)	(0.006)
Treat		-0.109***	-0.112***	-0.115***
		(0.005)	(0.003)	(0.003)
Post		0.019***	0.020***	0.019***
		(0.007)	(0.006)	(0.005)
Mean (Treated)				
N		79,676	123,165	169,720
R^2		0.097	0.097	0.096

Notes: All regressions include covariates. Number of post-reform years is always one (2008). Robust standard errors in parentheses. ***: statistically different at 1 percent.

Table 4.A11: Sensitivity to different control groups: Supplementary Benefits

Dependent variable:		Number of years pre reform		
HH: Share SB	1 (2007)	1 (2007)	2 (2006-07)	3 (2005-07)
Treated: 5-8 years in DI				
Control:	**3 years**	**2-3 years**		
Treat × Post	0.006	0.011*	0.012**	
	(0.008)	(0.006)	(0.005)	
Treat	0.034***	0.046***	0.045***	
	(0.005)	(0.004)	(0.003)	
Post	0.013	0.004	0.005	
	(0.008)	(0.005)	(0.004)	
Mean (Treated)	0.166			
N	68,903	75,794	116,791	
R^2	0.066	0.065	0.064	
Control:		**1-3 years**		
Treat × Post		0.012**	0.016***	0.015***
		(0.005)	(0.005)	(0.004)
Treat		0.056***	0.052***	0.053***
		(0.004)	(0.003)	(0.002)
Post		0.002	0.001	0.006
		(0.004)	(0.004)	(0.004)
Mean (Treated)				
N		79,676	123,165	169,720
R^2		0.065	0.064	0.062

Notes: See Table 4.A10. ***,**,*: statistically different at 1, 5 and 10 percent, respectively.

Table 4.A12: Placebo tests with policy change in 2006

Treated	MB: Employed	Spouse: Employed	HH: Share with SB
5-8 years	-0.002	0.004	0.007
	(0.008)	(0.008)	(0.006)
6-8 years	-0.004	0.004	0.005
	(0.009)	(0.008)	(0.006)
7-8 years	-0.002	0.004	0.008
	(0.009)	(0.009)	(0.007)
8 years	-0.014	0.010	0.003
	(0.010)	(0.011)	(0.008)

Notes: Every estimate represents the coefficient of Treat × Post from a separate regression including covariates. The control group always includes beneficiaries with 3 years in DI, so that the number of pre-reform years is always one (2005). Number of post reform years is always one (2006). Robust standard errors in parentheses.

Table 4.A13: Placebo tests with divorced beneficiaries

Dependent Variable	Number of years post reform			
	1 (2008)	2 (2008-09)	3 (2008-10)	4 (2008-11)
MB: Employed				
5-8 years	0.015			
	(0.018)			
6-8 years	0.017	0.024		
	(0.019)	(0.016)		
7-8 years	0.011	0.020	0.020	
	(0.019)	(0.017)	(0.016)	
8 years	0.003	0.013	0.016	0.020
	(0.021)	(0.019)	(0.018)	(0.017)
HH: Share with SB				
5-8 years	-0.015			
	(0.018)			
6-8 years	-0.014	-0.024		
	(0.018)	(0.016)		
7-8 years	-0.021	-0.026	-0.029*	
	(0.019)	(0.017)	(0.016)	
8 years	-0.003	-0.018	-0.020	-0.023
	(0.022)	(0.019)	(0.018)	(0.017)

Notes: Every estimate represents the coefficient of Treat × Post from a separate regression including covariates. The control group always includes beneficiaries with 3 years in DI, so that the number of pre-reform years is always one (2007). Robust standard errors in parentheses. *: statistically different at 10 percent.

Table 4.A14: Background characteristics depending on time of outflow

Outflow after	5-8 years in DI	3 years in DI	Difference
Age	58.24	49.33	-8.91***
Female	0.444	0.423	-0.021
Foreigner	0.303	0.310	0.007
Number of minor children	0.313	0.741	0.427***
Mental illness	0.232	0.244	0.012
N	9784	1122	

Notes: ***: statistically different at 1 percent level.

5 No Disabled Student Left Behind? Evidence From a Social Field Experiment

joint with Eva Deuchert, Helge Liebert, Carl Wuppermann

Abstract

This paper uses a randomized field experiment to study if student counseling offices treat requests from disabled students systematically different. We send randomly manipulated emails from fictitious students to ask for information on the admission process and special accommodations to ease studying. Written responses are evaluated with the help of text mining tools. Our experiment provides empirical evidence for trivialization and stigmatization of certain disabilities. Students with depression or dyslexia are systematically less likely to receive formal information on special admission rules and are offered disability counseling services less often compared to students with chronic diseases. If this translates into lower access and a lower probability to graduate, already existing disadvantages on the labor market may be reinforced.

5.1 Introduction

People with disabilities face substantial disadvantages on the labor market even if they are equally qualified. This results in lower employment and a decline in earnings (see e.g. Ravaud et al., 1992; DeLeire, 2001). To make the situation worse, people with disabilities often have lower qualifications, particularly with respect to education. This is especially problematic since education serves as a buffer to (at least partly) protect against the negative effects of a disability (Ravaud et al., 1992; Hollenbeck and Kimmel, 2008). To ensure that people with disabilities are not excluded from education, the UN Convention on the Rights of Persons with Disabilities calls governments to provide access

to the general education system, including tertiary education, vocational training, adult education and lifelong learning (Article 24).

To understand if countries meet the requirements of the UN convention or if policy reforms are still needed to improve access to education, it is of primary importance to analyze the inclusiveness of the education system. The focus of the literature has been on equal access to lower level education (see e.g. Dhuey and Lipscomb, 2010; Hanushek et al., 2002). Very little is known about obstacles that disabled students face in higher education institutions. The available literature documents that disabled students perceive barriers with respect to the admission procedure (see e.g. Baron et al., 1996; Riddell, 1998; Pardeck, 2001), infrastructure (O'Connor and Robinson, 1999; Losinsky et al., 2003), examinations (Fuller et al., 2004; Madriaga et al., 2010), or attitudes, knowledge and awareness among lecturers, peers, library and administrative staff (Korinek and Prillaman, 1992; Riddell, 1998; Vogel et al., 2007; Shevlin et al., 2004). This literature is typically qualitative, represents the subjective assessment of a small number of students who have made it to the respective institution, and is based on few Anglo-Saxon higher education institutions.

This paper takes a different approach: Rather than relying on subjective assessments of students with disabilities, we perform an objective test of higher education institutions' quality. We send randomly manipulated emails from prospective applicants who ask for information about the admission process, consideration of illness-related absences during studies, and special accommodations for examinations to 204 German higher education institutions.[1] We analyze all written replies to these emails using text mining approaches.

Defining uniform quality indicators for services provided by a higher education institution is challenging since disability is a very heterogeneous phenomenon. The umbrella term comprises physical, sensory, mental, and cognitive or developmental impairments. These different types of disabilities have very different needs. People in wheelchairs, for example, may need reasonable accommodations with respect to infrastructure, while students with a learning disability (such as dyslexia) may need special rules for examinations. At the very least, however, all cases should be treated similarly with respect to access to basic information and services (our main outcome variables). In reality, this may not be the case if some types of disabilities are more prone to stigmatization (i.e. perceiving students as less capable or intelligent) or trivialization (i.e. not considering

[1] Germany is a very interesting case since the federal government ratified the UN Convention on the Rights of Persons with Disabilities in 2008, while the responsibility for higher education lies in the hands of the different states (Länder). With this division of responsibilities, it is unclear whether Germany currently meets the requirements of the UN convention.

the condition as a disability that hampers the student's ability to successfully enter and complete a study program). We therefore study if higher education institutions discriminate students with respect to disability types.

Our results provide empirical evidence for both stigmatization and trivialization of certain disabilities: Students suffering from depression or dyslexia are less likely to receive formal information on special admission rules and are offered disability counseling services less often compared to students with chronic diseases. Study choices of students with depression are more often questioned. These findings are consistent with the previous literature: Upton and Harper (2002), for example, show that university students without disabilities perceive educational accommodations for disabled students (such as testing alternatives, assistive technologies, or special administrative privileges) to be most needed for physical and visible disabilities (such as blindness, deafness, and cerebral palsy); invisible disabilities (such as ADHD, depression, and borderline personality disorder) are designated the least deserving of special accommodation. Bianco (2005) uses a vignette study to document that teachers are significantly influenced by disability labels when making referrals for students to enter in gifted programs. We confirm these results from hypothetic situations in a realistic setting. The main advantage of our design is, however, that we focus on revealed rather than stated attitudes.

The rest of the paper proceeds as following: Section 5.2 provides a brief overview on the institutional background of the higher education system in Germany. Section 5.3 describes the research method and section 5.4 discusses the data used in our empirical analysis. Results are presented in section 5.5. Section 5.6 concludes.

5.2 Institutional Background

Under German law, people with disabilities are entitled to assistance in order to ensure that they have access to and complete higher education. Under the Framework Act for Higher Education (HRG), higher education institutions (HEI) shall ensure that disabled students do not suffer any disadvantage in their studies, that they are able to use the facilities of the institution, if possible without assistance (§2 HRG) and that examination regulation considers the special needs of disabled students to preserve equality of opportunity (§16 HRG). Since higher education is the responsibility of the federal states, similar rules can be found in the state legislations governing higher education, where all states reinforce the right that disabled students are not disadvantaged and can take advantage of the offers without outside help. All states (except for Baden-Württemberg)

furthermore guarantee a compensation for disadvantages in examinations. Some states require in addition appointing a representative for the needs of disabled students (in Bavaria, Brandenburg, Hamburg, Mecklenburg-Western Pomerania, Rhineland-Palatinate, Saxony-Anhalt and Thuringia). Nevertheless, formal rules on the rights of disabled students are still relatively vague leaving substantial autonomy to the higher education institutions.

Of special interest for this paper are special admission rules since all fictitious students ask for information on the admission process (see also section 5.3). In general, the German constitution implies a general right of access to higher education. However, access to higher education can be restricted (*numerus clausus*) due to capacity constraints. Two different restrictions can apply: (1) University places for medicine, pharmaceutics, dentistry and veterinary medicine are nationwide restricted and are (at least partly) allocated by a federal agency. (2) Local admissions restrictions may apply for specific programs in some HEI, whereas the same program may be freely accessible at other HEIs. Places are typically allocated by average grades, waiting time and other relevant qualifications but some institutions also require letters of motivation, internships, tests or personal interviews.

For unrestricted study programs, no special rules for disabled students are necessary. All students that meet the minimum qualification (usually a secondary education degree) can enroll in a program. In case admission to a study program is restricted, disabled students can make special requests. These special requests rules are implemented for the admission process for generally restricted study programs, but they are not obligatory for locally restricted study programs. Nevertheless, many HEI implemented similar rules: (1) Students with a progressive illness or disability that makes a longer waiting time unreasonable, or those for whom the selected program promises successful professional (re-)integration, can apply for immediate admission (hardship application). (2) Students who can demonstrate that an illness or a disability resulted in a deterioration of school grades or delayed graduation can apply for an improvement in the average grade of the school certificate or can be credited waiting time (disadvantage compensation). For both types, a certificate of severe disability (*Schwerbehindertenausweis*) is usually necessary. Additionally, medical expert appraisals or school reports that document why a type of disability has led to specific disadvantages are often needed.

Due to the lack of harmonization and standardization of admission rules for disabled students and fields of study, affected students are forced to obtain information on the admission process directly from HEIs. In some cases, this information is made available on

HEI's web pages. In any case, student counseling offices are the first point of contact for prospective students seeking information. For this reason it is quite common that disabled students contact student counseling offices revealing their disabilities and requesting information on the admission process.

5.3 Research Method

We sent randomly manipulated emails to student counseling offices of German HEI to evaluate the quality of the provided services with respect to special needs of students with disabilities. We include three different disability types (physical, mental, and learning) and a non-disabled person as a control.[2] Each HEI thus receives four different emails. All emails describe the health condition and to which degree these conditions may hamper academic performance. In the emails the persons ask about the admission process, special accommodations during studies (i.e. exemption from compulsory attendance and special accommodations during exams), and whether the HEI provides additional support to students with disabilities.

The different disability types are selected so that they are relatively common in the population of students and comparable with respect to the special accommodations needed. A representative survey of students at German universities reveals that 45% of students with a disability have a mental illness (most often depression), 20% have a chronic somatic illness, 6% have a learning disorder (most often dyslexia), 5% have a visual impairment, and 4% have a mobility impairment (Unger et al., 2012). We choose the following conditions for our study:[3] (1) Chronic kidney disease which requires regular renal dialysis represents a chronic disease. This student misses classes for two days a week. (2) Depression represents a mental illness. This student is in therapy and misses two days of classes per week. (3) Dyslexia represents a learning impairment that requires special accommodations for exams.

Our prime interest is to study if student counselors reply to emails from disabled students in a systematically different way. There are two main reasons why this might be the case: (1) Different types of disabilities are associated with different stigmata. Dyslexic or depressed students, for example, may be perceived as less capable or intelligent (Bianco, 2005), even though they often do not differ in general cognitive ability from other students (Hatcher et al., 2002). (2) There is great variability in the degree to

[2]This person claims that he takes care of a sick family member and misses two days of classes per week.
[3]Our study does not include visual or mobility impairments because they require very different accommodations mostly with respect to infrastructure.

which individuals perceive certain health conditions as a disability (Kapteyn et al., 2007). Some of the described health conditions may therefore not be considered as a disability deserving special assistance (Upton and Harper, 2002), despite the fact that all chosen impairments are classified as a disability under German law (§2 SGB IX).

To detect causal effects with respect to the disability types, emails need to be as similar as possible to ensure that differences in the responses do not stem from different writing styles, but should be different enough to protect the experiment from detection. We achieve this in four steps: (1) Each email is sent from a different email account. The names and email accounts are random combinations of the most prevalent surnames and family names in Germany (Lukas Fischer, Julia Müller, Laura Schmidt, and Jan Schneider). It is unlikely that real counterparts would take harm from our experiment. The gender is balanced to exclude any potential bias. (2) The emails are sent in four waves between March and May 2012 with approximately four weeks between each wave. (3) Four different authors formulate four different emails (one email per disability type). Every HEI receives only one email from each author. (4) The student expresses interest in different study subjects (i.e. economics/management, technical studies such as engineering or computer science, medicine, and teaching with focus on Math and German). All four factors are randomized.[4] This has the advantage that we are able to study effect heterogeneities (see section 5.5.3) and to test for external validity of our findings particularly with respect to learning and framing effects (see section 5.5.4).

We use text mining tools to analyze written replies (see section 5.4). In this respect our design differs to standard applications of correspondence tests, which have been widely used to detect discrimination in the labor market.[5] In this setting, fictitious applications are sent to potential employers where the application differs in a particular characteristic (such as gender, age, race or sexual orientation). Lower response rates on applications with specific characteristics are interpreted as a sign for discrimination. In our study, in contrast, we expect response rates to be high (providing information to students is the prime task of the student counseling offices) and not to vary much by disability type.[6] Discrimination according to disability type is more likely to be reflected in different email contents.

[4]Details on the randomization design can be found in the appendix.

[5]A detailed review can be found in Riach and Rich (2002).

[6]See also table 5.A2 (appendix). In the following analysis, we assume that observations are missing at random.

5.4 Data

In total we sent 852 emails to 214 different higher education institutions (HEI). We did not include specialized schools such as schools of arts or music into our study because they do not offer academic programs in economics/management, technical studies, medicine or teaching. Ten HEIs (36 emails) had to be dropped from the sample because they share student counseling offices. Emails were sent directly to the general student counseling office. We did not send any email to the counseling office for disabled students because not all HEIs provide such a service. The majority of institutions are universities of applied sciences (N=118) and state-owned (N=161), with an average of 10,132 students (see Table 5.A3 in the appendix).

We use simple text-mining tools (i.e. content analysis and keyword searches) to analyze all written responses to our emails. These techniques are used to analyze unstructured data, such as journal/newspaper articles or books (e.g. van Driel et al., 2006; Talamini et al., 2012; Acerbi et al., 2013).

Measuring overall quality of the provided service is challenging due to the heterogeneity of special needs of students with disabilities. At the very least, however, students should receive the same kind of information on the admission process (measured by a dummy variable that is equal to one if the email includes a link or attachment *INF*), and should have equal access to special services for students with disabilities (measured by a dummy variable that indicates if the email was forwarded to a counseling office for disabled students *DIC*). Differences with respect to disability types are good indicators for trivilization: Student counseling offices may be less likely to perceive certain impairments as a disability that justifies special admission rules or assistance from counseling office for disabled students. We also test whether there are significant differences regarding the tone of the emails by including a dummy variable indicating that the counselor questions the student's study choice by suggesting a different HEI, a different study program or even to delay studying (*QDE*). Systematic differences with respect to disability type in this variable could be an indicator for stigmatization: Students with certain impairments may be perceived to be less intelligent or capable to graduate, or to be less likely to successfully exercise the profession after graduation.[7]

We also test for systematic differences in various additional outcome variables, such

[7] One HEI, for example, answered to a student with dyslexia who was interested in studying medicine that he should reconsider the study choice because there is a risk to misinterpret a diagnosis and then further mistreat a patient, or that a pharmacist or a nurse administers a wrong medication because it was misspelled.

as the information provided on the admission process,[8] special accommodations to ease studying and exams,[9] the emotional tone of the email,[10] and if the condition is perceived as a disability.[11] Descriptive statistics for all variables can be found in Table 5.A4 (appendix).

5.5 Results

5.5.1 Main effects

The results of our main specification are presented in Table 5.1. These results stem from probit regressions which additionally include responses to emails from students with no disabilities (proxying for overall quality or style of the HEI). Further control variables are indicators for the name of the student, author of the email, wave, and study subject.

The results document that student counseling offices discriminate disabled students by type: Students with depression and even more pronounced, students with dyslexia receive information via links and attachments less often compared to students with chronic kidney failure. In case of dyslexia, the difference at the means is 11 percentage points, which is significant at the 1% level. Students with mental or learning disabilities are offered significantly less access to disability counseling services (8 and 9 percentage points less, respectively). For the indicator measuring whether the study choice is questioned we find a significant increase in responses to emails from students with depression. Even though in only 5% of all responses student counseling services question the study choice, our results nevertheless demonstrate that stigmatization of certain disabilities is likely.

Non-disabled students who care for a sick family member receive information via links and attachments similarly often as students with chronic kidney failure. Their request is significantly less often forwarded to disability counseling services. This may not come as a surprise since they have no legal entitlement to compensations. Interestingly,

[8] This is measured by various dummy variables that indicate whether the response included the following keywords: special requests (SPR), hardship (HRS), disadvantage compensation (DAC), admission (ADM), or certificate (CER).

[9] This is captured by keywords for exam (EXA), attendance (ATT), study fees (FEE), and study contents (CON).

[10] We capture this by a dummy variable that offers personal advice via phone or a personal meeting (OCO), a dummy variable indicating that any kind of special requests are not possible (NSR), as well as indicators for the keywords happy (HAP) for positive feedbacks or unfortunately (UNF) for negative feedbacks.

[11] This is measured by a dummy variable indicating that the keywords disability, sickness or health were used in the response (DIA).

Table 5.1: Main results

Dependent variable	INF Link/Attachment	DIC Forwarded Counseling	QDE Questions Decision
Disability type (omitted category: chronic kidney failure)			
Depression	-0.045	-0.078***	0.041*
	(0.042)	(0.025)	(0.025)
Dyslexia	-0.110***	-0.093***	0.010
	(0.041)	(0.027)	(0.025)
Family	-0.013	-0.208***	0.046*
	(0.041)	(0.026)	(0.029)
Predicted probability at the mean	0.383	0.185	0.050
Pseudo R^2	0.029	0.068	0.022
N	692	692	692

Notes: The probit regressions additionally control for the name of the students, author of the email, wave, and study subject. Marginal effects at the means reported. Standard errors clustered by the Higher Education Institution. The outcome variables are coded as following: INF (Email includes links/attachments), DIC (Forwarded to disability counseling), QDE (Questions decision). * p<0.1; ** p<0.05; *** p<0.01

their decision is more often questioned as they are advised to study part-time or by correspondence course at other HEIs.

5.5.2 Additional results

A potential concern may be that student counseling offices simply substitute one transmission mode for information (i.e. links/attachments) with other transmission modes (i.e. providing the same information in the email or the intention to provide this information by personal conversation). Our additional results, however, reveal that this is not the case (see Table 5.A5 in the appendix). Responses to emails from students with depression or dyslexia include information on hardship application (*HRS*) less often compared to responses to emails from students with a chronic disease. Personal contacts (*OCO*) are offered less frequently as well.

We also confirm our results regarding trivialization, where depression and particularly dyslexia are less often labeled as a disability (*DIA*). These differences are systematic. Some HEIs sent almost identical responses to all students (showing that answers are very often based on text modules), but deleted all keywords related to disability in responses to students with dyslexia.

In addition, students with dyslexia or depression tend to receive less information on special accommodations during studies for example regarding study contents (CON). The only major difference is that students with dyslexia are more likely to receive information on special accommodations during exams (EXA), which is expected since questions regarding examination were raised in the initial request from these students.

5.5.3 Effect heterogeneity

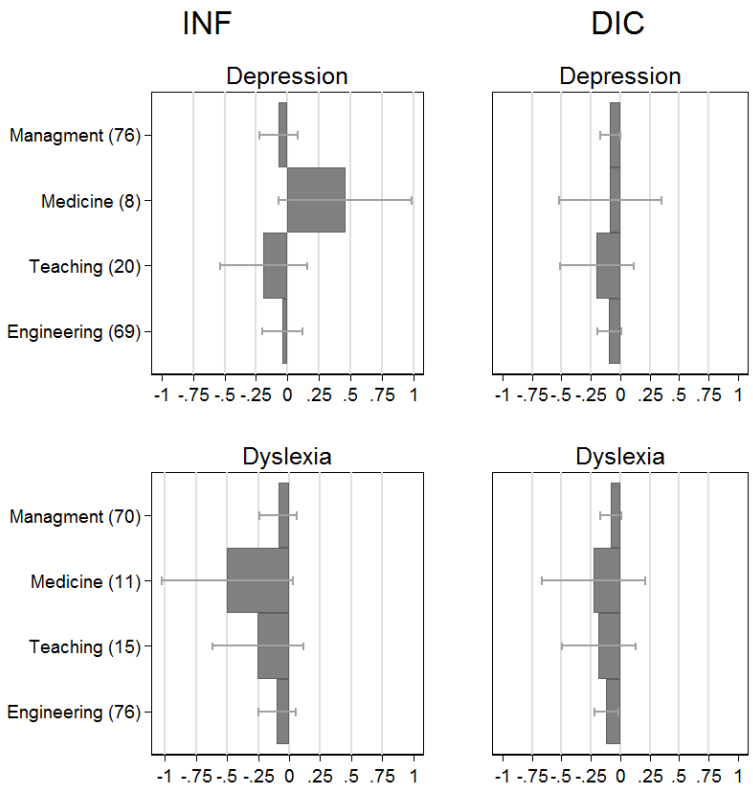
In the previous section we have seen systematic differences exist in the way student counseling offices reply to our fictitious emails. In the following we are interested if these differences are grounded by the choice of the study program.

Different study programs require different skills. Stratifying our results by study programs helps to better understand if the striking differences in responses are related to trivialization or stigmatization. Since all emails were sent to the general student counseling offices and not to the respective faculties, we do not expect systematic differences in results with respect to study program if observed differences in response behavior are grounded in ignorance or trivialization. If stigmatization is the driving reason for observed differences in response behavior, we expect significant effect heterogeneities for three main reasons: (1) Study counseling offices may perceive that students with learning or mental disability do not have the intellectual capacity to successfully complete certain study programs. One HEI, for example, replied to a student with dyslexia who asked for information on a study program for engineering that she would not recommend to study engineering if the student has problems to correctly write down formal contents, since both in mathematics and in code implementations, typos would immediately lead to wrong solutions.[12] (2) Student counselors may fear that the profession cannot be exercised after completing the study program. This is particularly likely for academic programs in teaching. One HEI, for example, replies to a student with dyslexia that error-free writing or reading is a key skill to practice the profession. (3) Study programs vary by their costs.[13] For public schools, these costs are predominantly borne by the society since the study fees for students are low or non-existing in Germany. However, HEIs may still fear that a disabled student drops out and that the investment is sunk. One HEI, for example, expressed this fear by claiming that they want to avoid in any case

[12]Note that a dyslexic person does not necessarily also suffer from dyscalculia (Lewis et al., 1994).

[13]While Economics or Management programs are relatively cheap to operate (on average the cost per graduate is about 27,100 €, the costs per student to graduate in Medicine can be substantial with 197,700 € (Brugger et al., 2012).

Figure 5.1: Effect heterogeneity with respect to study program



Notes: Bars display the marginal effects at the means for depression and dyslexia cases on stratified probit regressions that additionally control for no disability, sender, author, and wave. Range plots display 95% confidence intervals. The numbers in brackets are the numbers of responses on emails from students with depression or dyslexia who expressed interest in the respective study program. The outcome variables are coded as following: INF (Email includes links/attachments), DIC (Forwarded to disability counseling).

that the student invests money in his education and then runs the risk not to graduate. It may thus be the case that HEIs are unwilling to accommodate students with a mental or learning disability in study programs that are particularly expensive (such as Medicine for example).

Empirical evidence for effect heterogeneity with respect to study program is presented in Figure 5.1, where the bars represent the marginal effects at the means for depression and dyslexia from probit regressions in stratified samples. We do not find much evidence

for effect heterogeneity. All effects have the same sign and similar magnitude.[14,15] The observed differences in the provision of information and services are thus more likely to be grounded in ignorance and trivialization.

5.5.4 Potential frauds to external validity

All parameters of our experiment are fully randomized. Consequently, any difference in responses to requests from students with different disabilities can be fully attributed to the type of disability and our results are internally valid. The key question in any experimental study is, however, the extent to which results are also externally valid.

Our first concern regarding the external validity of our experiment relates to the fact that four different emails were sent to each HEI. This is typically considered to be an unnecessary procedure, given that our profiles are entirely randomized (Newman, 1980). Identification of our main effects therefore does not rely on a within-comparison. The reasoning for this procedure is a very simple power argument: There are not many higher education institutions in Germany and sending more than one email to each institution increases the number of observations. This procedure, however, bears the risk of arousing suspicion, leading to detection of the experiment (Riach and Rich, 2004b). This may particularly be the case if the emails do not differ enough or if HEIs do not often receive requests from students with disabilities. If this was the case, we could expect learning behavior in the sense that emails sent in later waves systematically differ to those sent in earlier waves. As shown in Figure 5.A2 (appendix), effects are mostly similar in sign and magnitude. The results for wave 3, however, are smaller or even go into the opposite direction. This could be a day of the week effect (emails were sent on a Monday as opposed to weekdays later in the week for other waves). Nevertheless, since these differences are not statistically significant and the results of wave 4 are qualitatively similar to the earlier waves, the results do not indicate a consistent time pattern which would suggest a learning effect.

Another threat to the external validity of our results are framing effects. Student counselors may not only react to the disability but also to other differences, such as the gender of the requesting person. If effect heterogeneities exist with respect to these differences (if for example emails by Lukas Fischer are assessed in a different way than

[14]Effects are not significant anymore due to smaller sample sizes.

[15]The only exception is medicine, where the effect for depression is reversed for the indicator that formal information is provided (*INF*). Note, however, that medicine is the only nationally restricted study program and most study counseling offices simply referred to the homepages of the national agency.

the ones by Julia Müller), our results are internally valid but not representative for other students with disabilities. We test if the name of the sender and the author are significantly associated with our main outcome variables and if the effect of different disability types is affected by these factors (see Figure 5.A3 and 5.A4 in the appendix). We do not find evidence for effect heterogeneities with respect to neither the author nor the sender of the emails. In most cases, the effects are similar in sign and magnitude.

5.6 Conclusion

This study uses a randomized field experiment to detect if student counseling offices treat requests from disabled students in a systematic different way. We sent randomly manipulated emails from fictitious students to ask for information on the admission process as well as special accommodations to ease studying. Responses to these emails are evaluated using text mining tools. We observe substantial differences in the way student counselors reply to our fictitious students: Students with dyslexia or depression receive systematically less information on special accommodations, and are provided access to disability counseling less often compared to students with chronic diseases. Students with depression additionally see their decision more often questioned.

Our results document stigmatization and trivialization of students with mental and cognitive disabilities. We are aware that these replies reflect student counselors' attitudes only, but it is very likely that other academic and administrative staff have similar attitudes. It is thus necessary to sensitize academic and administrative staff to the specific needs of students with disabilities. This implies to educate HEI staff that a disability does not only relate to physical handicaps, but that mental impairments commonly exist among students, and that these students have the same protection under the UN convention and other national laws as a student with a physical impairment. This is particularly important given that such impairments are considerably more prevalent.

5.7 Appendix

Randomization design Each Higher Education Institution (HEI) received in total four emails. Each email is a random combination of the disability type (no disability, chronic kidney disease, depression, and dyslexia), the field of study (economics/management, engineering/computer science, medicine, and teaching), the student's name (Jan Schneider, Julia Müller, Laura Schmidt, and Lukas Fischer), the authors who have written the emails, and the date on which the emails were sent. If the HEI does not offer all considered academic programs, the field of study was randomly chosen out of those considered fields that are offered. Randomization was based on a simultaneous random draw without replacement to ensure that each HEI receives every characteristic exactly once. The randomization design is visualized in Figure 5.A1. Each side of the matrix represents a specific characteristic, while the order of the characteristics is randomized for each HEI. The combination of characteristics represented by the diagonal is chosen to construct the emails.

Figure 5.A1: Visualization of the randomization design

Author of the email

		Author 1	Author 4	Author 2	Author 3		
Disability type	Dyslexia	x				Name 3	Name
	Kidney		x			Name 1	
	Depression			x		Name 4	
	No disability				x	Name 2	
		Subject 4	Subject 2	Subject 1	Subject 3		

Field of study

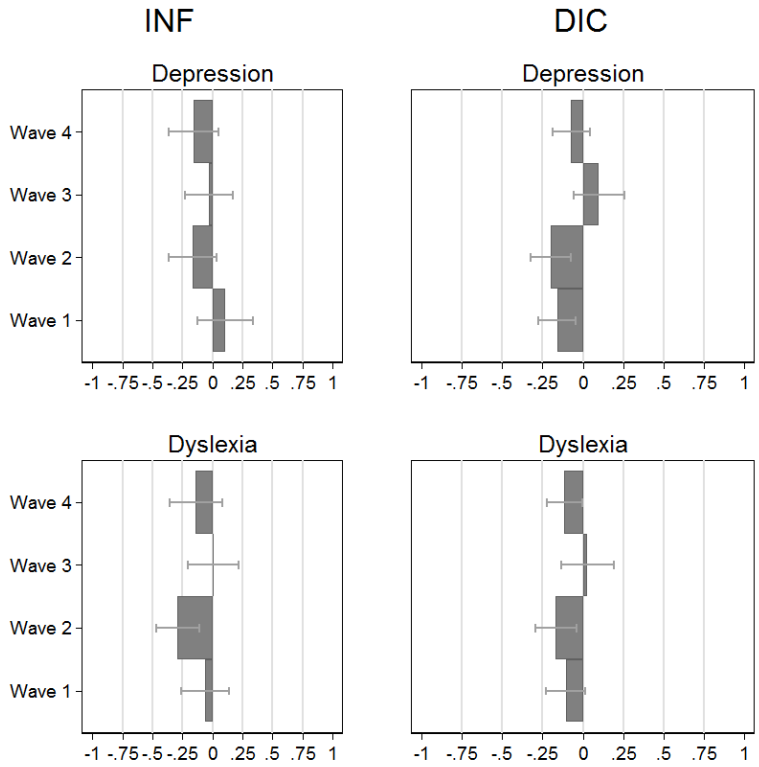
The randomization was generally successful. The null hypothesis that the randomization characteristics are independent (based on Pearson's Chi-square Test for Independence) cannot be rejected in almost all cases (see Table 5.A1). Only illness and wave are loosely associated.

Table 5.A1: Pearson chi-square test for independence

	Illness	Sender	Author	Subject
Sender	3.61			
Author	8.67	5.22		
Subject	5.75	5.16	3.89	
Wave	16.86*	9.29	12.43	5.13

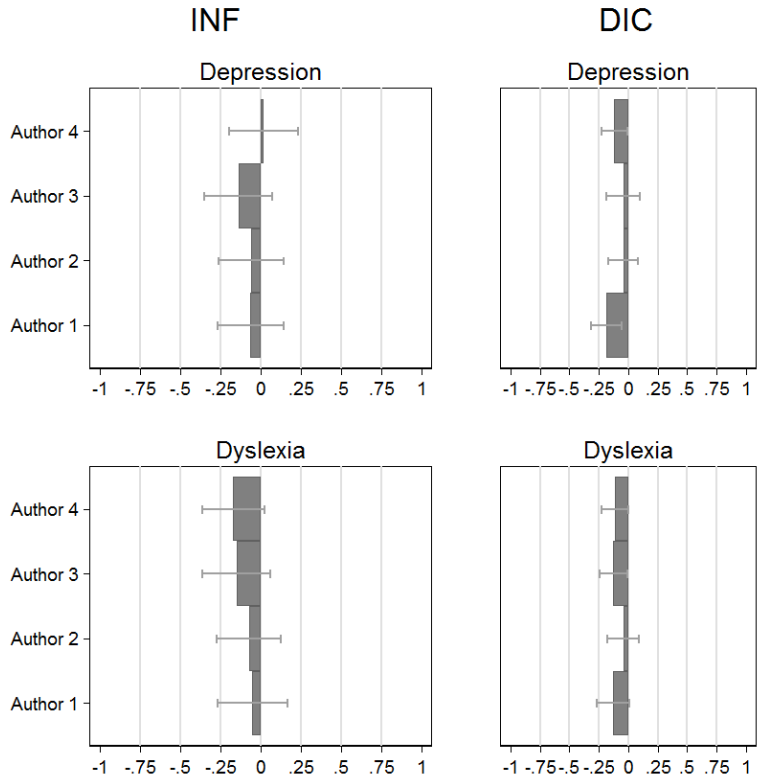
Notes: Degrees of freedom: 9. Significance level: * $p < 0.1$

Figure 5.A2: Learning effects



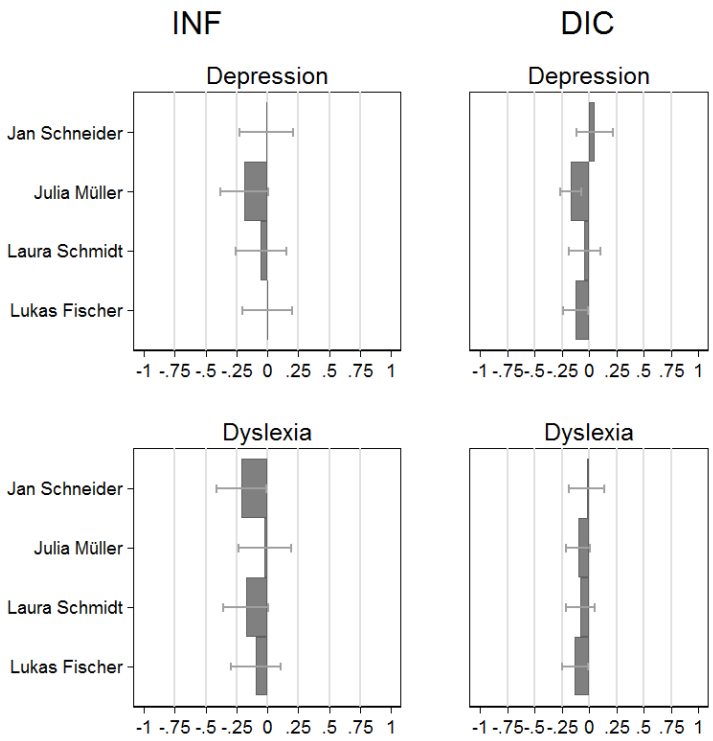
Notes: Bars display the marginal effects at the means for depression and dyslexia cases on stratified probit regressions that additionally control for no disability, study program, sender, author, and wave. Range plots display 95% confidence intervals. The outcome variables are coded as following: INF (Email includes links/attachments), DIC (Forwarded to disability counseling).

Figure 5.A3: Framing effects (author)



Notes: Bars display the marginal effects at the means for depression and dyslexia cases on stratified probit regressions that additionally control for no disability, study program, sender, author, and wave. Range plots display 95% confidence intervals. The outcome variables are coded as following: INF (Email includes links/attachments), DIC (Forwarded to disability counseling).

Figure 5.A4: Framing effects (sender)



Notes: Bars display the marginal effects at the means for depression and dyslexia cases on stratified probit regressions that additionally control for no disability, study program, sender, author, and wave. Range plots display 95% confidence intervals. The outcome variables are coded as following: INF (Email includes links/attachments), DIC (Forwarded to disability counseling).

Table 5.A2: Probit regression results for response

	Response (*RES*)
Disability type (omitted category: chronic kidney failure)	
Depression	-0.024
	(0.033)
Dyslexia	-0.035
	(0.035)
Family	-0.036
	(0.035)
Name of student (omitted category: Lukas Fischer)	
Laura Schmidt	-0.016
	(0.035)
Julia Müller	-0.003
	(0.030)
Jan Schneider	-0.040
	(0.034)
Author of Email (omitted category: Author 1)	
Author 2	0.054
	(0.027)*
Author 3	0.037
	(0.027)
Author 4	0.061
	(0.026)**
Wave (omitted category: Wave 1)	
Wave 2	0.071
	(0.026)**
Wave 3	-0.028
	(0.034)
Wave 4	-0.028
	(0.032)
Study subject	
(omitted category: Engineering/Computer Sciences)	
Teaching	0.048
	(0.042)
Medicine	0.050
	(0.052)
Management/Economics	0.024
	(0.029)
Predicted probability at the mean	0.855
Pseudo R^2	0.028
N	816

Notes: Marginal effects at the means. Standard errors are clustered by the Higher Education Institution. * p<0.1; ** p<0.05; *** p<0.01.

Table 5.A3: Background characteristics

Total number of institutions	204
Background information observed	200
Type: University of applied sciences (no doctoral granting)	118
Type: University (doctoral granting)	82
Type of control: state	161
Federal state: Baden-Wurttemberg	32
Federal state: Bavaria	29
Federal state: Berlin	10
Federal state: Brandenburg	5
Federal state: Bremen	4
Federal state: Hamburg	5
Federal state: Hesse	15
Federal state: Mecklenburg-Western Pomerania	4
Federal state: Lower Saxony	17
Federal state: North Rhine-Westphalia	38
Federal state: Rhineland-Palatinate	8
Federal state: Saarland	2
Federal state: Saxony	9
Federal state: Saxony-Anhalt	6
Federal state: Schleswig-Holstein	9
Federal state: Thuringia	7
Average number of students	10,132
Average year of foundation	1913

Table 5.A4: Descriptive statistics (outcome variables)

Variable	Description	Mean
Main outcome variables		
INF	Email includes links/attachments	0.39
DIC	Forwarded to disability counseling	0.20
QDE	Questions decision	0.05
Additional outcome variables		
Information on special accommodations		
SPR	Keyword: special request	0.04
HRS	Keyword: hardship	0.18
DAC	Keyword: disadvantage compensation	0.47
ADM	Keyword: admission (synonyms)	0.45
CER	Keyword: certificate (synonyms)	0.25
EXA	Keyword: exam (synonyms)	0.51
ATT	Keyword: attendance (synonyms)	0.18
FEE	Keyword: study fees (synonyms)	0.04
CON	Keyword: study contents (synonyms)	0.26
Tone of the email		
OCO	Offered contact	0.42
NSR	No special request possible	0.05
HAP	Keyword: happy ("we are happy to")	0.13
UNF	Keyword: unfortunately ("unfortunately it is not possible")	0.09
Disability		
DIA	Keyword: disabled, sick, health (synonyms)	0.39

Notes: Number of observations: 692.

Table 5.A5: Content measures

Keywords/ synonyms	SPR Special request	HRS Hardship	DAC Disadvantage compensation	ADM Admission (synonyms)	CER Certificate (synonyms)	EXA Exam (synonyms)	ATT Attendance (synonyms)
Disability type (omitted category: chronic kidney failure)							
Depression	0.019	-0.088***	0.124***	-0.067	0.018	-0.139***	-0.096**
	(0.015)	(0.031)	(0.047)	(0.049)	(0.044)	(0.048)	(0.039)
Dyslexia	-0.021	-0.100***	0.048	0.052	0.047	0.170***	-
	(0.013)	(0.028)	(0.048)	(0.051)	(0.045)	(0.046)	-
No disability	0.006	-0.123***	0.031	-0.019	-0.077*	-0.185***	0.004
	(0.017)	(0.029)	(0.050)	(0.050)	(0.042)	(0.049)	(0.042)
Predicted probability at the mean	0.025	0.176	0.470	0.447	0.244	0.515	0.216
Pseudo R^2	0.106	0.039	0.037	0.043	0.019	0.067	0.094
N	692	692	692	692	692	692	520

Keywords/ synonyms	FEE Study fee (synonyms)	CON Study contents (synonyms)	OCO Offered contact	NSR No special request possible	HAP Happy (synonyms)	UNF Unfortunately (synonyms)	DIA Recognized disability
Disability type (omitted category: chronic kidney failure)							
Depression	-0.007	-0.066*	-0.066	0.011	-0.021	-0.010	-0.120***
	(0.015)	(0.035)	(0.042)	(0.012)	(0.026)	(0.023)	(0.042)
Dyslexia	-0.016	-0.125***	-0.081*	0.009	0.020	-0.025	-0.268***
	(0.013)	(0.036)	(0.047)	(0.023)	(0.027)	(0.027)	(0.039)
No disability	0.006	-0.068*	-0.132***	0.046**	0.003	0.076***	-0.398***
	(0.015)	(0.037)	(0.044)	(0.023)	(0.029)	(0.032)	(0.033)
Predicted probability at the mean	0.036	0.233	0.414	0.042	0.123	0.079	0.370
Pseudo R^2	0.057	0.087	0.019	0.065	0.018	0.068	0.113
N	692	692	692	692	692	692	692

Notes: Marginal effects at the means. Additional controls include indicators for the name of the student, author of the email, wave, and study subject. Standard errors are clustered by the Higher Education Institution. *, **, *** denote significance at the 10%, 5% and 1% level, respectively.

Bibliography

Acerbi, A., V. Lampos, P. Garnett, and R. A. Bentley (2013): "The Expression of Emotions in 20th Century Books," *PLoS ONE*, 8, e59030.

Adam, S., A. Bozio, and C. Emmerson (2010): "Reforming Disability Insurance in the UK: Evaluation of the Pathways to Work Programme," London, Institute for Fiscal Studies.

Alexandre, P. K. and M. T. French (2001): "Labor Supply of Poor Residents in Metropolitan Miami, Florida: The Role of Depression and the Co-Morbid Effects of Substance Use," *Journal of Mental Health Policy and Economics*, 4, 161–173.

Autor, D. H. and M. G. Duggan (2003): "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics*, 118, 157–205.

——— (2007): "Distinguishing Income from Substitution Effects in Disability Insurance," *American Economic Review*, 97, 119–124.

——— (2010): "Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System," London, The Center for American Progress and The Hamilton Project.

Baron, S., R. Phillips, and K. Stalker (1996): "Barriers to Training for Disabled Social Work Students," *Disability & Society*, 11, 361–378.

Baumgartner, E., S. Greiwe, and T. Schwarb (2004): "Die berufliche Integration von behinderten Personen in der Schweiz - Studie zur Beschäftigungssituation und zu Eingliederungsbemühungen," Olten, Fachhochschule Solothurn Nordwestschweiz.

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

BBT (2010): "Vocational and professional education and training in Switzerland," Bern, Federal Office for Professional Education and Technology.

Bell, S. H. and L. L. Orr (1994): "Is Subsidized Employment Cost Effective for Welfare Recipients? Experimental Evidence from Seven State Demonstrations," *Journal of Human Resources*, 29, 42–61.

Bell, S. H. and L. R. Peck (2012): "Obstacles to and Limitations of Social Experiments: 15 False Alarms," Abt thought leadership paper, Abt Associates.

Bertrand, M. and S. Mullainathan (2004): "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94, 991–1013.

BFS (2009): "Regionale Disparitäten in der Schweiz," Neuchatel, Bundesamt für Statistik.

Bianco, M. (2005): "Winner of CLD'S 2004 Award for Outstanding Research: The Effects of Disability Labels on Special Education and General Education Teachers' Referrals for Gifted Programs," *Learning Disability Quarterly*, 28, 285–293.

Biesma, R. G., M. Pavlova, G. G. van Merode, and W. Groot (2007): "Using conjoint analysis to estimate employers preferences for key competencies of master level Dutch graduates entering the public health field," *Economics of Education Review*, 26, 375–386.

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

Blomquist, G., K. Blumenschein, and M. Johannesson (2009): "Eliciting Willingness to Pay without Bias using Follow-up Certainty Statements: Comparisons between Probably/Definitely and a 10-point Certainty Scale," *Environmental and Resource Economics*, 43, 473–502.

Blumenschein, K., G. C. Blomquist, M. Johannesson, N. Horn, and P. Freeman (2008): "Eliciting Willingness to Pay Without Bias: Evidence from a Field Experiment," *Economic Journal*, 118, 114–137.

Blundell, R., M. C. Dias, C. Meghir, and J. van Reenen (2004): "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2, 569–606.

Boockmann, B., T. Zwick, A. Ammermüller, and M. Maier (2012): "Do hiring subsidies reduce unemployment among the elderly? Evidence from two natural experiments," *Journal of the European Economic Association*, 10, 735–764.

Bound, J. (1989): "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review*, 79, 482–503.

——— (1991): "Self-reported versus objective measures of health in retirement models," *Journal of Human Resources*, 26, 106–138.

Bound, J. and R. V. Burkhauser (1999): *Economic Analysis of Transfer Programs Targeted on People with Disabilities*, Handbooks in Economics, vol. 5. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 3417–3528.

Bound, J. and T. Waidmann (2002): "Accounting for Recent Declines in Employment Rates among Working-Aged Men and Women with Disabilities," *Journal of Human Resources*, 37, 231–250.

Bricout, J. and K. Bentley (2000): "Disability status and perceptions of employability by employers," *Social Work Research*, 24, 87–95.

Brugger, P., M. Threin, and M. Wolters (2012): *Hochschulen auf einen Blick*, Wiesbaden: Statistisches Bundesamt.

BSV (2003): "Kreisschreiben über Invalidität und Hilflosigkeit in der Invalidenversicherung (KSIH)," Bern, Bundesamt für Sozialversicherungen.

——— (2008a): "IV-Statistik 2008," Bern, Bundesamt für Sozialversicherung.

——— (2008b): "Monatliche Vollrenten, Skala 44," *http://www.ahv-iv.info/andere/00194/index.html?lang=de*, [Date of access: July 25, 2013].

——— (2011): "IV-Statistik 2010," Bern, Bundesamt für Sozialversicherung.

Burtless, G. (1985): "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment," *Industrial and Labor Relations Review*, 39, 105–114.

Campolieti, M. (2004): "Disability Insurance Benefits and Labor Supply: Some Additional Evidence," *Journal of Labor Economics*, 22, 863–889.

Campolieti, M. and C. Riddell (2012): "Disability policy and the labor market: Evidence from a natural experiment in Canada, 1998-2006," *Journal of Public Economics*, 96, 306–316.

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

Carlsson, M. and D.-O. Rooth (2007): "Evidence of ethnic discrimination in the Swedish labor market using experimental data," *Labour Economics*, 14, 716–729.

Chatterji, P., M. Alegria, and D. Takeuchi (2009): "Racial/Ethnic Differences in the Effects of Psychiatric Disorders on Employment," *Atlantic Economic Journal*, 37, 243–257.

Chen, S. (2012): "Spousal Labor Supply Responses to Government Programs: Evidence from the Disability Insurance Program," *Michigan Retirement Research Center WP 2012-261*.

Chen, S. and W. van der Klaauw (2008): "The work disincentive effects of the disability insurance program in the 1990s," *Journal of Econometrics*, 142, 757–784.

Chiteji, N. (2010): "Time Preference, Noncognitive Skills and Well Being across the Life Course: Do Noncognitive Skills Encourage Healthy Behavior?" *American Economic Review*, 100, 200–204.

Cimera, R. E. (2007): "The cost-effectiveness of supported employment and sheltered workshops in Wisconsin (FY 2002-FY 2005)," *Journal of Vocational Rehabilitation*, 26, 153–158.

Clark, R. E., B. J. Dain, H. Xie, D. R. Becker, and R. E. Drake (1998): "The Economic Benefits of Supported Employment for Persons with Mental Illness," *Journal of Mental Health Policy and Economics*, 1, 63–71.

Clayton, S., C. Bambra, R. Gosling, S. Povall, K. Misso, and M. Whitehead (2011): "Assembling the evidence jigsaw: insights from a systematic review of UK studies of individual-focused return to work initiatives for disabled and long-term ill people," *BMC Public Health*, 11, 170.

Cornwell, K., C. Forbes, B. Inder, and G. Meadows (2009): "Mental illness and its effects on labour market outcomes," *Journal of Mental Health Policy and Economics*, 12, 107–118.

Corrigan, P. W. and S. G. McCracken (2005): "Place First, Then Train: An Alternative to the Medical Model of Psychiatric Rehabilitation," *Social Work*, 50, 31–39.

Crowther, R. E., M. Marshall, G. R. Bond, and P. Huxley (2001): "Helping people with severe mental illness to obtain work: systematic review," *BMJ: British Medical Journal*, 322, 204–208.

Cullen, J. B. and J. Gruber (2000): "Does Unemployment Insurance Crowd out Spousal Labor Supply?" *Journal of Labor Economics*, 18, 546–572.

de Bekker-Grob, E. W., M. Ryan, and K. Gerard (2012): "Discrete choice experiments in health economics: a review of the literature," *Health Economics*, 21, 145–172.

DeLeire, T. (2001): "Changes in Wage Discrimination against People with Disabilities: 1984-93," *Journal of Human Resources*, 36, 144–158.

Dhuey, E. and S. Lipscomb (2010): "Disabled or young? Relative age and special education diagnoses in schools," *Economics of Education Review*, 29, 857–872.

Dorsett, R. (2006): "The new deal for young people: effect on the labour market status of young men," *Labour Economics*, 13, 405–422.

Drydakis, N. (2009): "Sexual orientation discrimination in the labour market," *Labour Economics*, 16, 364–372.

Dubin, J. A. and D. Rivers (1993): "Experimental estimates of the impact of wage subsidies," *Journal of Econometrics*, 56, 219–242.

Duggan, M., R. Rosenheck, and P. Singleton (2010): "Federal Policy and the Rise in Disability Enrollment: Evidence for the Veterans Affairs' Disability Compensation Program," *Journal of Law and Economics*, 53, 379–398.

Eriksson, S., P. Johansson, and S. Langenskiöld (2012): "What is the Right Profile for Getting a Job? A Stated Choice Experiment of the Recruitment Process," *IZA Discussion Paper No. 6691.*

Ettner, S. L., R. G. Frank, and R. C. Kessler (1997): "The Impact of Psychiatric Disorders on Labor Market Outcomes," *Industrial and Labor Relations Review*, 51, 64–81.

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

Forslund, A., P. Johansson, and L. Lindqvist (2004): "Employment subsidies - a fast lane from unemployment to work?" *IFAU Working Paper 2004:18*.

Fuller, M., A. Bradley, and M. Healey (2004): "Incorporating disabled students within an inclusive higher education environment," *Disability & Society*, 19, 455–468.

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

Gettens, J. W. (2009): "Medicaid Expansions: The Work and Program Participation of People with Disabilities," Ph.D. thesis, The Heller School for Social Policy and Management: Brandeis University.

Giertz, S. and J. Kubik (2011): "The Disability Screening Process and the Labor Market Behavior of Accepted and Rejected Applicants: Evidence from the Health and Retirement Study," *Journal of Labor Research*, 32, 237–253.

Goetzel, R. Z., S. R. Long, R. J. Ozminkowski, K. Hawkins, S. Wang, and W. Lynch (2004): "Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers," *Journal of Occupational and Environmental Medicine*, 46, 398–412.

Grob, A., R. Lüthi, F. G. Kaiser, A. Flammer, A. Mackinnon, and A. J. Wearing (1991): "Berner Fragebogen zum Wohlbefinden Jugendlicher (BFW)," *Diagnostica*, 37, 66–75.

Gruber, J. (2000): "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy*, 108, 1162–1183.

Gruber, J. and J. D. Kubik (1997): "Disability insurance rejection rates and the labor supply of older workers," *Journal of Public Economics*, 64, 1–23.

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*.

——— (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*.

Halpern, J. and J. A. Hausman (1986): "Choice under uncertainty: A model of applications for the social security disability insurance program," *Journal of Public Economics*, 31, 131–161.

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*, 27, 498–520.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2002): "Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities?" *The Review of Economics and Statistics*, 84, 584–599.

Hatcher, J., M. J. Snowling, and Y. M. Griffiths (2002): "Cognitive assessment of dyslexic students in higher education," *British Journal of Educational Psychology*, 72, 119–133.

Heckman, J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24, 411–482.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999): *The Economics and Econometrics of Active Labor Market Programs*, Amsterdam: North-Holland, vol. 3, 1865–2097.

Heckman, J. J. and Y. Rubinstein (2001): "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 91, 145–149.

Hollenbeck, K. and J. Kimmel (2008): "Differences in the Returns to Education for Males by Disability Status and Age of Disability Onset," *Southern Economic Journal*, 74, 707–724.

Hoyos, D. (2010): "The state of the art of environmental valuation with discrete choice experiments," *Ecological Economics*, 69, 1595–1603.

Huttunen, K., J. Pirttilä, and R. Uusitalo (2013): "The employment effects of low-wage subsidies," *Journal of Public Economics*, 97, 49–60.

Imdorf, C. (2009): "Die betriebliche Verwertung von Schulzeugnissen bei der Ausbildungsstellenvergabe," *Empirische Pädagogik*, 23, 392–409.

Jaenichen, U. and G. Stephan (2009): "The effectiveness of targeted wage subsidies for hard-to-place workers," *Applied Economics*, 43, 1209–1225.

Johannesson, M., G. C. Blomquist, K. Blumenschein, P.-o. Johansson, B. Liljas, and R. M. O'Conor (1999): "Calibrating Hypothetical Willingness to Pay Responses," *Journal of Risk and Uncertainty*, 18, 21–32.

Jones, M. K., P. L. Latreille, and P. J. Sloane (2006): "Disability, gender, and the British labour market," *Oxford Economic Papers*, 58, 407–449.

Kapteyn, A., J. P. Smith, and A. van Soest (2007): "Vignettes and Self-Reports of Work Disability in the United States and the Netherlands," *American Economic Review*, 97, 461–473.

——— (2009): "Life Satisfaction," *IZA Discussion Paper No. 4015*.

Karlström, A., M. Palme, and I. Svensson (2008): "The employment effect of stricter rules for eligibility for DI: Evidence from a natural experiment in Sweden," *Journal of Public Economics*, 92, 2071–2082.

Katz, L. (1996): "Wage Subsidies for the Disadvantaged," *National Bureau of Economic Research Working Paper Series*.

Korinek, L. and D. Prillaman (1992): "Counselors and Exceptional Students: Preparation Versus Practice," *Counselor Education and Supervision*, 32, 3–11.

Kornfeld, R. and K. Rupp (2000): "The Net Effects of the Project NetWork Return-to-Work Case Management Experiment on Participant Earnings, Benefit Receipt, and Other Outcomes," *Social Security Bulletin*, 63, 12–33.

Kreider, B. and J. V. Pepper (2007): "Disability and Employment," *Journal of The American Statistical Association*, 102, 432–441.

Lechner, M. (2011): "The Estimation of Causal Effects by Difference-in-Difference Methods," *University of St. Gallen Discussion Paper*, 2010-28.

Lewis, C., G. Hitch, and P. Walker (1994): "The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls," *Journal of Child Psychology and Psychiatry*, 35, 283–292.

Linden, M., S. Baron, and B. Muschalla (2009): *Mini-ICF-APP - Mini-ICF-Rating für Aktivitäts- und Partizipationsstörungen bei psychischen Erkrankungen - Ein Kurzinstrument zur Fremdbeurteilung von Aktivitäts- und Partizipationsstörungen bei psychischen Erkrankungen in Anlehnung an die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) der Weltgesundheitsorganisation*, Bern: Verlag Hans Huber.

——— (2010): "Relationship between work-related attitudes, performance and capacities according to the ICF in patients with mental disorders." *Psychopathology*, 43, 262–267.

Losinsky, L.-O., T. Levi, K. Saffey, and J. Jelsma (2003): "An investigation into the physical accessibility to wheelchair bound students of an Institution of Higher Education in South Africa," *Disability and Rehabilitation*, 25, 305–308.

Louviere, J. J., D. A. Hensher, J. D. Swait, and W. Adamowicz (2000): *Stated Choice Methods: Analysis and Applications*, Cambridge: Cambridge University Press.

Madriaga, M., K. Hanson, C. Heaton, H. Kay, S. Newitt, and A. Walker (2010): "Confronting similar challenges? Disabled and non-disabled students' learning and assessment experiences," *Studies in Higher Education*, 35, 647–658.

Marie, O. and J. Vall Castello (2012): "Measuring the (income) effect of disability insurance generosity on labour market participation," *Journal of Public Economics*, 96, 198–210.

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

Meisen Zannol, F. (2010): "Wie Melanie M. den Einstieg in die Arbeitswelt schaffte," *Schweizerischer Arbeitgeber*, 105, 18–19.

Mitra, S. (2009): "Disability Screening and Labor Supply: Evidence from South Africa," *American Economic Review*, 99, 512–516.

Moore, T. J. (2011): "The Employment Effects of Terminating Disability Benefits: Insights from Removing Drug Addictions as Disabling Conditions," College park, University of Maryland.

Morrison, M. and T. Brown (2009): "Testing the Effectiveness of Certainty Scales, Cheap Talk, and Dissonance-Minimization in Reducing Hypothetical Bias in Contingent Valuation Studies," *Environmental and Resource Economics*, 44, 307–326.

Mowbray, C. T., M. E. Collins, C. D. Bellamy, D. A. Megivern, D. Bybee, and S. Szilvagyi (2005): "Supported education for adults with psychiatric disabilities: an innovation for social work and psychosocial rehabilitation practice," *Social Work*, 50, 7–20.

Murphy, J., P. Allen, T. Stevens, and D. Weatherhead (2005): "A Meta-analysis of Hypothetical Bias in Stated Preference Valuation," *Environmental and Resource Economics*, 30, 313–325.

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*, NBER Working Paper No. 16866.

Newman, J. M. (1980): "[Discrimination in Recruitment: An Empirical Analysis]: Reply," *Industrial and Labor Relations Review*, 33, 547–550.

Norwood, F. B. and S. R. Henneberry (2006): "Show Me the Money! The Value of College Graduate Attributes as Expressed by Employers and Perceived by Students," *American Journal of Agricultural Economics*, 88, 484–498.

O'Connor, U. and A. Robinson (1999): "Accession or Exclusion? University and the Disabled Student: A Case Study of Policy and Practice," *Higher Education Quarterly*, 53, 88–103.

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

——— (2009): "Sickness, Disability and Work: Keeping on track in the economic downturn," Stockholm, High-Level Forum.

——— (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*, 3, 148–171.

Özdemir, S., F. R. Johnson, and A. B. Hauber (2009): "Hypothetical bias, cheap talk, and stated willingness to pay for health care," *Journal of Health Economics*, 28, 894–901.

Pardeck, J. T. (2001): "Social Work Admissions and Academic Accommodations for Students With Disabilities: An Exploratory Study," *International Journal of Adolescence and Youth*, 9, 147–158.

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*, 41, 203–212.

Ravaud, J.-F., B. Madiot, and I. Ville (1992): "Discrimination towards disabled people seeking employment," *Social Science & Medicine*, 35, 951–958.

Riach, P. A. and J. Rich (2002): "Field Experiments of Discrimination in the Market Place," *Economic Journal*, 112 483, F480–F518.

——— (2004a): "Deceptive Field Experiments of Discrimination: Are they Ethical?" *Kyklos*, 57, 457–470.

——— (2004b): "Fishing for Discrimination," *Review of Social Economy*, 62, 465–486.

Riddell, S. (1998): "Chipping away at the mountain: Disabled students' experience of higher education," *International Studies in Sociology of Education*, 8, 203–222.

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

Sainani, K. (2010): "The Importance of Accounting for Correlated Observations," *PM&R*, 2, 858–861.

Savoca, E. and R. Rosenheck (2000): "The Civilian Labor Market Experiences of Vietnam-Era Veterans: The Influence of Psychiatric Disorders," *Journal of Mental Health Policy and Economics*, 3, 199–207.

Schmid, M. and M. Storni (2004): "Transition - Zur Qualifikation und Rekrutierung von Lehrlingen im Kanton Basel-Stadt," Basel, ecco Gemeinschaft für Sozialforschung.

Schulz, U. and R. Schwarzer (2003): "Soziale Unterstützung bei der Krankheitsbewältigung: Die Berliner Social Support Skalen (BSSS)," *Diagnostica*, 49, 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," *http://userpage.fuberlin.de/∼health/world14.htm*, [Date of access: April, 21, 2011].

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

Seyd, W., K. Schulz, and B. Vollmers (2007): "Verzahnte Ausbildung METRO Group mit Berufsbildungswerken (VAMB) - Abschlussbericht der wissenschaftlichen Begleitung," Hamburg, Universität Hamburg.

Shevlin, M., M. Kenny, and E. Mcneela (2004): "Participation in higher education for students with disabilities: an Irish perspective," *Disability & Society*, 19, 15–30.

Sianesi, B. (2008): "Differential effects of active labour market programs for the unemployed," *Labour Economics*, 15, 370–399.

Singleton, P. (2012): "Earnings of rejected applicants to the Social Security Disability Insurance program," *Economics Letters*, 116, 147–150.

Staubli, S. (2011): "The impact of stricter criteria for disability insurance on labor force participation," *Journal of Public Economics*, 95, 1223–1235.

Talamini, E., C. E. Caldarelli, E. F. Wubben, and H. Dewes (2012): "The composition and impact of stakeholders' agendas on US ethanol production," *Energy Policy*, 50, 647–658.

Thornton, C., G. Livermore, T. Fraker, D. Stapleton, B. O'Day, D. Wittenburg, R. Weathers, N. Goodman, T. Silva, E. Sama Martin, J. Gregory, D. Wright, and A. Mamun (2007): "Evaluation of the Ticket to Work Program," Washington, Mathematica Policy Research.

UN (2006): "International Convention on the Rights of Persons with Disabilities - Some Facts about Persons with Disabilities," http://www.un.org/disabilities/convention/pdfs/factsheet.pdf, United Nations, [Date of access: July 25, 2013].

Unger, M., P. Wejwar, S. Zaussinger, and A. Laimer (2012): *Studying with impairments in Germany: A 2011 survey of the situation for students with a disability or chronic illness*, Berlin: German National Association for Student Affairs.

Upton, T. D. and D. C. Harper (2002): "Multidimensional disability attitudes and equitable evaluation of educational accommodations by college students without disabilities," *Journal of Postsecondary Education and Disability*, 15, 115–130.

van Beek, K. W. H., C. C. Koopmans, and B. M. S. van Praag (1997): "Shopping at the Labour Market: A Real Tale of Fiction," *European Economic Review*, 41, 295–317.

van Driel, M. A., J. Bruggeman, G. Vriend, H. G. Brunner, and J. A. M. Leunissen (2006): "A text-mining analysis of the human phenome," *European Journal of Human Genetics*, 14, 535–542.

Vogel, G., B. Fresko, and C. Wertheim (2007): "Peer Tutoring for College Students With Learning Disabilities: Perceptions of Tutors and Tutees," *Journal of Learning Disabilities*, 40, 485–493.

von Wachter, T., J. Song, and J. Manchester (2011): "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program," *American Economic Review*, 101, 3308–3329.

WHO (2002): "Towards a Common Language for Functioning, Disability and Health," Geneva, World Health Organization.

Woodbury, S. A. and R. G. Spiegelman (1987): "Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois," *American Economic Review*, 77, 513–530.

Curriculum vitae

Born April 11, 1982 in Affoltern am Albis, Switzerland

Education

2010 -	University of St. Gallen, Switzerland *PhD Programme in Economics and Finance*
2003 - 2007	University of Zurich, Switzerland *Licentiate in economics (lic.oec.publ.)*
2006	Humboldt-University Berlin, Germany *Exchange semester (Erasmus)*

Professional experience

2010 -	Center for Disability and Integration, University of St. Gallen, St. Gallen, Switzerland *Research assistant*
2008 - 2009	Winterthur Institute of Health Economics, Zurich University of Applied Sciences, Winterthur, Switzerland *Research assistant*
2006 - 2007	KOF Swiss Economic Institute, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland *Junior research assistant*