Institute for the Study of Social Change

ISSC DISCUSSION PAPER SERIES

THE EFFECT OF DISABILITY ON LABOUR MARKET OUTCOMES IN GERMANY: EVIDENCE FROM MATCHING

Dr. Rosalia Vasquez Alvarez
Dr Michael Lechner

Rosalia Vasquez is a member of the Economics Dept. in UCD and an associate of the Institute for the Study of Social Change (ISSC). Michael Lechner is professor of Econometrics at the University of St Gallen, Switzerland.

ISSC WP 2003 / 03

The Effect of Disability on Labour Market Outcomes in Germany: Evidence from Matching

Michael Lechner[a], Rosalia Vazquez-Alvarez[b]

Swiss Institute for International Economics and Applied Economic Research (SIAW), University of St.Gallen, Switzerland

This version: February 2003[*]

Abstract

If labour market policies aimed at people with disabilities are effective, we should observe no significant difference in labour market outcomes between disable and non-disable individuals. This paper examines the impact of disability status on labour market outcomes using matching methods associated with treatment effect techniques for program evaluation. Such techniques avoid model misspecification and account for the common support problem, thus improving the identification strategy of alternative techniques that also select on observables. Using several waves from the German Socio-Economic Panel (GSOEP, 1994-2000) we estimate the impact of disability on both labour market participation and labour earnings. We find no significant difference in either of these two measures of labour market outcomes between disable and non-disable. Due to the construction of the treated and comparison groups, our results imply that (in Germany) disability labour market policies are effective at achieving their aim.

JEL – Classification: C13, C14, I12, I18, J23

Keywords: Treatment effect, evaluation of disability policies, health status, causality, matching on the propensity score, labour market outcomes.

---

[a] Michael Lechner is also affiliated with CEPR, London, ZEW, Mannhein, and IZA, Bonn. Email: Michale.Lechner@unisg.ch, http://www.siaw.unisg.ch/lechner

[b] Rosalia Vazquez-Alvarez is also affiliated to ISSC-UCD Ireland, and the Health Policy Research Division at the ESRI, Dublin-2, Ireland. Email: Rosalia.Vazquez-Alvarez@ucd.ie, http://www.ucd.ie/~economic/Staff.htm.

[*] Address for correspondence: SIAW, University of St.Gallen, Dufourstrasse, 48. St.Gallen, CH-9000, Switzerland.

1 Introduction

Most industrialized economies recognize the need for effective policies and practices in support of workers whose prospects of either remaining or (re-)integrating in employment are jeopardised by work injury, long term illness and/ or disability. For example, in Germany the Severely Disabled Persons Act of 1974 (*Schewerbehindertengesetz Schwbg*) – further amended in 1986–, or SDPA for short, sets forth the obligation of an statutory quota of a minimum of 6% disable employees on employers with a workforce of 16 employees or more, with such quota applying equally to both the private and the public sector.[1] The same Act obligates the employer to adjust their premises in order to accommodate disable workers, provides legislation which protects the disable against discrimination in recruitment, employment, and unfair dismissal, as well as setting down fines for those who fail to fulfil their quotas, along side a variety of generous subsidies to facilitate employers to adjust to such policies and practices. Likewise, the SDPA provides a wide range of advantages to encourage participation in paid labour market activities of disable individuals who are able to participate, for example, tax benefits, subsidized transport cost, re-training programs and the legal right to longer holidays per year, among others. Countries such as the UK, the USA and Australia, follow practices similar to those in Germany.[2]

Research focusing on the effect of disability on labour market outcome is identical in nature to empirical studies which focus on labour market outcome differentials between genders or due to racial difference. Nevertheless, studies of the effect of disability on labour market outcome is by no means as prolific, specially in Europe. In the United States many studies have focused on the importance of health status (i.e., disability status) on labour supply behaviour, but have centred attention on the population nearing retirement age (for example, see Kreider and Pepper (2002) and references therein, and

1 The Act was further amended in July 2001. Given that our empirical section looks at data between 1994 and 2000, the relevant Act for our purpose is that of 1974/1986.

2 For example, in the UK, the 1944 Disable Persons Employment Act (further amended in 1996), imposes an statutory quota of 3% of disable persons in the workforce for employers with 20 or more workers, imposing fines on those whose quotas are not met. The same Act defines the obligations on behalf of employers to adjust their premises in order to accommodate disable individuals, as well as legislation for the protection of disable employees with respect to discrimination in recruitment, employment, or dismissal for reasons which relates to disability.

Williamson and McNamara (2002)). With respect to Europe, Kidd, Sloane and Ferko (1998) provide an example on the effect of disability on both wages and participation-rates using data on males from the UK 1996 Labour Force Survey. Their study estimates the participation rates of disable and non-disable individuals using independent probit models for each of the two sub-populations. Following Even and McPherson (1991), they decompose the difference between the two estimated participation rates between explained and unexplained components. They find a 50% participation rate differential, and suggest that only half of this estimate can be explained by productivity related characteristics, thus providing some evidence on the ineffectiveness of labour market policies which aim at integrating the disable into the labour force.

The key econometric difficulty in the aforementioned literature results from the non-random selection of individuals into different status with respect to disability, i.e., workers in sectors with higher occupational hazard, individuals with a taste for sports with high risk, living on a highly urbanized metropolitan area, are all factors which increase the chances of an individual to become disable. For example, in the Kidd et al (1998) exercise, identification of the effect of disability in the presence of such non-random selection comes from conditioning on pre-determined observed characteristics of both participants and non-participants, hoping that such observable characteristics will account for any bias which might result from differentials in the chances of becoming disable.

There are two potential problems with this approach. The first problem is that for any given set of conditioning variables, we might fail to observe persons in each of the two disability status we seek to compare. This problem is known as the failure of the common support condition. To show this point, assume the extreme case where all the disable in the sample are associated with high risk pre-disability occupational activities, while all non-disable observations show occupational activities with zero risk throughout all their working lives. In this case our data cannot identify the effects of disability on labour market outcome, as we cannot separate out the effects of disability and occupational sector.

The second problem is that, even when the support problem is not an impediment in identifying disability effects, the choice of model (parametric, semi-parametric) is often based on strong data and/or functional form assumptions, to the extend that model misspecification might also lead to a second source of bias.

The two above mention problems are typical when evaluating the effect of a particular treatment (in our case disability) using non-experimental data. However, recently micro-econometricians have adopted the techniques of epidemiologist based on studying the effect of an intervention (or treatment) to evaluate non-experimental rather than laboratory data. [3] By using nonparametric techniques, such as matching procedures, it is possible to address both the common support issue and problems associated with model specification. As with linear (and non-linear) specifications, matching also assumes selection on observable characteristics. The ideas is to think that there exist a set of observed variables such that conditional on these, the impact of the treatment is independent of the untreated outcome. Such assumption is know as the Conditional Independence Assumption, or CIA for short. In our context, matching methods allow us to assume that given a set of X variables, within subgroups defined by X, and for any given individual, becoming disable is unrelated to what the participation outcome or wage outcome would be if she had not become disable. Thus, conditional on X, we can find a counterfactual outcome to each treated observation and estimate the impact of the treatment.

Matching techniques may overcome the non-random selection problem (with respect to observables) and highlight the common support problem. Matching techniques are nonparametric in nature, and although one needs a large number of informative covariates to identify the causal effect, if the dimension of X is very large, the estimation process can suffer from the curse of dimensionality. In order to reduce such problem, this paper employs propensity score matching methods, thus matching on a function of X rather than X itself (see Rosenbaum and Rubin, 1983). Because of the binary nature of our treatment[4] (disability versus non-disability), our estimates of the propensity score are based on probit estimates, where $\hat{P}(x) = P_n(disable \mid X)$ for $X = x \in \mathbb{R}^p$, and a sample size n. The outcomes of non-treatment observations $j's$ who are "close" to the i^{th} observation from the treatment group in terms of $\hat{P}_j(x)$ relative to $\hat{P}_i(x)$, become the counterfactual outcomes. We use various matching methods, namely, nearest neighbour – with and without calliper –, Gaussian Kernel and Epanechnikov Kernel. The reason for using various methods is because in finite samples each method weights the distance between $\hat{P}_j(x)$ and $\hat{P}_i(x)$ differently. Likewise, each method differs with respect to the treatment of the common support problem.

3 See Angrist (1991), Heckman and Horz (1989), Ichimura and Todd (1997), Lechner (1995, 1996, 1997), Smith and Todd(2000), but to mention a few.

4 For an example on how to deal with the problem in case of multiple treatments see Lechner (2001)

The empirical results of this paper are based on data taken from the German Socio-Economic Panel (GSOEP, 1994-2000).[5] This annual survey is very informative with respect to labour market outcomes as well as on social, economic and living conditions in Germany. The panel dates from 1984 but it was only in 1994 when health related variables became homogenous for all sub-samples in the surveyed population, specially with reference to the variable identifying the degree of disability of each surveyed individual.[6] We believe that the richness of the data allows us to make the assumption that outcomes (participation, earnings) and disability status are independent conditional on observed attributes, thus solving the identification problem inherent in causal analysis. The need for different waves is due to our definition of treated versus untreated observations. An individual who was non-disable at t_1 (e.g., 1994), but become disable at t_2 (i.e., 1995) and remained disabled at t_3 (i.e., 1996)[7] is defined as an ADD (or treated) individual, that is, such individual receives the impact of the policies aimed at the disable at t_2 and we can observe the effect of such policies on labour market outcome at t_3. The control (or untreated – non-disable –) group, referred to as AAA, are individuals who declare non-disability over the same three year period. With seven waves (1994 to 2000) we can define 5 sequence of three years each, thus allowing us to increase our sample size, specially with respect to the treatment group. Our analysis consists on using matching techniques to compare the labour market outcomes at t_3 of ADD versus AAA. One of the consequences of constructing the control and treatment groups in such way is that the set X in the matching process is defined with respect to pre-disability period t_1, and, therefore, our empirical analysis is completely void of endogeneity problems between outcome (defined at t_3) and the conditioning set at t_1 (see Section 4 for further details).

Besides our contribution to a growing body of econometric literature using treatment effect techniques for program evaluation, this paper makes very important contributions to the understanding of how disable individuals fair in the labour market. Our results indicate different empirical conclusions. First, our results show that for this particular application the common support is not a problem. We have

5 Based on the English based general public release version, which includes 95% of the surveyed population.

6 See Section 2 for the legal definition of disability in Germany and Section 3 for a description of this variable with reference to the GSOEP data set.

7 This paper would not be completed without analysing the data allowing for attrition. Work in progress accounts for this in the form of Manski type of bounds on the measure of average treatment effect on the treated (for example, see Manski (1994)).

examined the impact of disability labour policies on both participation rates and labour earnings. The magnitude of our results show that non-disable have an slightly higher participation rate than disable persons, while this later group have slightly higher labour earnings relative to non-disable individuals. However, the magnitude of the difference are almost negligible (with participation rate differentials between 0.1-4.3% and only up to DM. 3,100 earning differentials) and statistically insignificant. Our results suggest that in Germany, the impact of disability policies on the disable are effective at reducing their participation cost into competitive labour market activities . These results are consistent for all matching methods employed in the empirical section.

The remainder of this paper is organized as follows. Section 2 provides a concise explanation of the labour market policies and legislation in Germany with respect to disable persons. Section 3 describes the GSOEP data used in the empirical section of the paper. Section 4 defines the econometric methodology, identifying conditions and matching methods employed in the empirical section. Section 5 presents the estimated impact of disability on two different labour market outcomes using the techniques of Section 4. Section 6 concludes.

2 Policies and Practices in Germany for labour market participants with disabilities

The main legislation concerning disable persons in Germany is the "Severely Disabled Persons Act (1974)" – *Schewerbehindertengesetz Schwbg* – which was further amended in 1986, and issued by the Federal Ministry for Labour and Social Affairs. In short, we refer to this Act as the SDPA. Although the SDPA does not adhere to one exact definition of disability, in its broader terms it takes up the three tiered definition proposed by the World Health Organization (WHO), where disable persons are defined as those who suffer from the consequences of the effects of a physical, mental or psychological condition which is not typical for the respective age, and where the consequences are not merely of a temporary nature. The definition covers the terms handicap, disability and impairment.[8] With such definition as a benchmark, each individual who wishes – voluntarily – to be assessed in terms of

8 The definition varies according to additional requirements for the application to specific situations, and with regards to the assistance required by different circumstances and institutions (*Bundesministerium fuer Arbeit und Socialordnung (BMA, 1996 Publication, p.11*).

disability has to go through a formal medical procedure conducted by a special independent institution (*Versorgungsamt*), where he or she is identified with a particular degree of disability. The degree of disability is express in percentage increments from 0 to 100% (total disability). The degree of disability is given to each person independently from his or her fitness to work in his or her present occupation or in future view of desired occupation. Once an individual is assigned a particular degree of disability, the public welfare authorities (*Hauptfuersorgestellen*) decides if the legislation as set in the SDPA is applicable to that person. Two possibilities exist. First, legislation as set in the Act covers all individuals with a degree of disability greater or equal to 50%. Second, individuals with a degree between 30 and 50% are also covered if the *Hauptfuersorgestellen* considers that the disability is the reason why the individual cannot find or hold an existing job. The SDPA prescribes and legislates for both sides of the labour market, namely the employer and the employee. Whereas the SDPA provides legislation, prescriptions, penalties and benefits for the employer, legislation with respect to employees are penalty free and only with the voluntary consent of the disable person.

The SDPA legislates that employers with a workforce greater or equal to 16 are legally obliged to employ a minimum of 6% disable workers. Furthermore, employers subject to the legislation have to provide adequate workspace for disable employees, according to their skills and capabilities, as well as appointing a representative inside the workplace who will look after the disable person's interest. Employers who do not fulfil the quota have to pay a levy of 200 DM (or the equivalent in Euros) per month for unfulfilled compulsory placements.[9] This revenue is used fully to finance national measures for the integration of severely disable persons. Since the quota system was introduced in 1974, the fulfilment of the quota has steadily declined over the years; while the 6% target has never been achieved, the highest percentage was in 1982 with 5.9%, with the latest figures showing an average of 4.2%.[10] One could think of such figures as a measure that the policies are not working, and consequently disable are less likely to be employed than non-disable. However, other evidence suggest that what such figures

9 An alternative to paying the full levy, enterprises can see their levy reduced if they award contracts to sheltered workshops. This workshops are places where severely disable individuals participate on paid labour market activities while sheltered from the competitiveness of the labour market. It is often the case that mentally handicap individuals, e.g., Down Syndrome persons, will work in such shelters.

10 In general the public sector is better at fulfilling its quotas – e.g., the federal government has to report to parliament every year on such quotas, so it makes an effort to employ at least up to the minimum of 6%–, while the West Germany does better on average (4.2%) than the East Germany who average around 2.9% (Zentras, 1997).

show is a badly designed quotas system. In 1995, and according to the quota requirement, there should have been 397,700 vacancies allocated to the disable in West Germany, but during that year only 155,500 severely disable people (at least 30% degree of disability) were registered as unemployed. Similar figures for East Germany show 20,000 registered disable persons versus 107,000 quota required vacancies. In 1996 the figures for West Germany were 513,187 required vacancies versus 181,200 registered unemployed with disabilities, while there was an almost balance with respect to the number of disable and vacancies offered to them with a ratio of 112:100 (although this does not indicate the ratio of match vacancies). Furthermore, the quota system does not take into account the number of disable employees who are employed beyond the required quota and companies who, without an obligation, still employ disable individuals (Albrecht and Braun, 1998).

The quota legislation comes along with other financial mechanisms that benefit the employer, with an aim to retain and or (re-)employ disable people. Examples of these are subsidizing the creation of new vacancies, wage subsidies (this can be up to 80% of gross wage for first year, 70% for second year and 60% for third year), financial support for workplace adaptation and creation (loans and subsidies of up to 100%), and financial support for special employee training and vocational rehabilitation which can cover up to 100% of the cost. However, according to Thornton and Lunt (1997) the reason why this financial mechanism are rarely taken up by employers is because of lack of information, specially for small enterprises, as well as too much bureaucratised procedures which discourages many small and medium size employers. Furthermore, in the case where benefits might only extend to workspace adaptation, perhaps this is not needed, at least not for existing employees. One further possibility for the failure of enterprise to take advantage of such benefits is because of the double role of the Hauptfuersorgestellen; while employers might take up some of the benefits, they also become fully subject to the sovereignty of the authorities, and this might make employers to be reserved (Albrecht and Braun, 1998). On the other hand, it is often the case that such subsidies might end up having a dead weight effect with respect to promoting additional disability employment, since employers who receive the subsidy might have employed (or continued to employ) the disable individual anyway. In Germany, it is generally accepted that financial incentives do not promote and/ or maintain employment of disabled people, but rather they reinforce a willingness to do so for the already existing disable workforce (Oyen, 1989).

The SDPA also sets legislation for the protection of disable employees making dismissal of such workers a very difficult task. If an employer decides to dismiss a disable individual, the representative of the disable in the workplace has to be informed, and such dismissal has to be approved by the welfare authorities (*Hauptfuersorgestellen)*. Such protective measures apply also to individuals whose disability degree is been ascertained (e.g., those who become disable with respect to or outside work, are given protection as if severely disable, at least until their disability degree is been assessed). The decision of the *Hauptfuersorgestellen* is mandatory, unless there is some outside agreement on behalf of the employer and employee which satisfies both parties. The basic guidance is that the dismissal will be approved if the employer can proof that the employee stands against the interest of the enterprise. If the dismissal is not approved the employer can appeal to a labour court. In 1995, 35% of such dismissals resulted in job retention (with 15.5% been in disapproval with the employer) while 46.9% resulted in job loss without the consent of the disable employee. The remainder (18.1%) also resulted in job loss but with consent of the employee (e.g., early retirement).

Besides legal protection, disable are also offered financial incentives to encourage them into paid working activities. These include financial support of vocational rehabilitation measures, reimbursement of the cost resulting from job search activities (e.g., application forms, travelling expenses), financial assistance to set up self-employment, purchase of working aids, subsidizing public and private transport, and subsidizing expenses associated with promoting mobility (e.g., subsidize adaptation of a new house if reallocating for work reasons).

All the above legislations and prescriptions should motivate profit maximising employers to employ a percentage of disable at least up to the minimum quota. Likewise, such policies should increase the motivation of disable persons who are capable to enter a competitive labour market, since the aim of such policies is to lower the entry cost of participation.

Overall, if such policies work, we should observe no differential between disable and non-disable participation rate. Wage subsidizing and tax incentives should also lead to zero labour income differential between the two groups, since such measures should account for disability related productivity differentials. Social scientist suggest otherwise, and focus attention on macroeconomic figures as a way to back up the argument that persons with disability fair worst in the labour market than non-disable. A set of figures often mentioned is the overall unemployment rate. For example, Albrecht and Braun (1998) compare the 1996 unemployment rate of officially unemployed disable persons in West Germany (15.9%) to that of the non-disable population (9.1%) – in East Germany the figures where 18.9 and 15.7%, respectively – and suggest this figures as evidence that the policies do not work. However, this figures compares groups without telling us about the causal relation between disability status and employment status. It might be that the disable who are registered unemployed are associated with occupational sectors that suffers from higher unemployment rates than the non-disable in the population, thus the above snap-shot provides a distorted comparison between the two sub-populations. Table 1 shows the distribution of disable employees among economic sectors using data from 1995 and 2000[11]. This evidence shows that disable employees are more likely to be associated with blue collar occupations (manufacturing, transportation, production and related) than any of the other economic sectors. Furthermore, the differences in share of blue collar workers between disable and non-disable is positive (13.2% in 1995) and significantly different than zero (t-value = 2.72). The 2000 figure suggest that such difference does not change over time. It is therefore not sufficient to make inference in the overall unemployment rate of Germany, since unemployment rates might differ within occupational categories. A more unbiased inference would result if we compared the overall unemployment rate within cells defined by attributes such as occupation, but also with reference to other characteristics (e.g., ability, motivation, vacancy matching, etc.). Examination of micro-economic survey data overtime, together with the appropriate statistical tools, might provide a more robust set of conclusions.

11 Table 1 looks at 1995 because the Albrecht and Braun (1998) study picks such year to show the brake-down between occupational sectors, using the national estimates based on Zentras (1997). The problem is that neither Albrecht and Braun (1998) or Zentras (1997) provide comparative figures for non-disable employees. The GSOEP data set used in Table 1 is able to provide similar statistics based on the one-digit ISCO classification, with the added advantage that we further compare between able and non-disable. We are confident that our weighted estimates in Table 1 are representative of the German population since they compare very well to the Zentras (1997) estimates in Albrecht and Braun (1998). For example, in their studies the share of disable employees in manufacturing, transportation, building and construction is equal to 45.9%, while the share of such employees in sales and services (trade, banking and insurance) is 21.9%. Our estimates are 42.8% and 20.2%, respectively.

Table 1: Distribution of disable among economic sectors, 1995 and 2000. (weighted population estimates for West and East Germany).

	1995		2000	
Occupational Category	Non-Disable	Disable	Non-Disable	Disable
Sample size	7,446	105	5,692	168
Professional, technical and related	19.2	15.1	22.7	8.2
Administration and Managerial work	3.8	1.5	4.3	1.7
Clerical and related work	21.0	18.3	22.3	22.7
Sales worker	9.0	4.5	8.4	12.8
Service Worker	10.7	15.7	10.9	10.8
Agricultural, animals, forestry, fishery.	1.7	0.6	1.7	1.0
Production, manufacturing, transport and related.	29.6	42.8	27.8	41.1
Others	4.9	1.5	2.0	1.6

Source: Weighted working sample, GSOEP 1995, 2000

Note: Table 1 is based on the weighted sample of disable and non-disable who declare to be active participants at the time of the survey. Although the sample size of disable is small relative to that on non-disable, the weighted percentage with respect to the population in the sample is well in line with estimates of the population. For example, according to the *Mikrozensus* statistical survey in 1995 (Statistiches Bundersamt Deutschland, see www.destatis.de), there were 930,600 disabled in employment out of 38.9 million actively engaged in labour market activities (see www.laborsta.ilo.org) therefore disable individuals counted as 2.4% of the total working population. Our sample of disable in employment in the 1995 GSOEP account for 2% of the working population (where our sample selection criteria is very much in line with the definition of prime age population – see Section 3 for more detail.)

3 The GSOEP Data

The data used in this study is based on seven waves from the German Socio Economic Panel (GSOEP, 1994-2000). The GSOEP is an annual micro-economic panel with the first wave starting in 1984. In 1990 the panel was extended to cover the new adhered East German states. The aim of the panel is to provide data for the analysis of social, economic and living conditions in Germany, with data representative of the German population at individual, household level and family level. The core questions cover demographics, education, labour market status and history, earnings, housing, health, household production and a section on subjective valuations (e.g., satisfaction with work, life, etc.). Apart from the core sample representing the full German population, the panel also contains specific sub-samples representative of minority groups, for example, migration workers (those who are German resident but of Spanish, Turkish, Italian or Yugoslav origin), and immigrants (of any origin) who have settled in Germany since 1984.[12]

Interviews are carried out face to face, with each household member age 16 or over counting as an individual observation. Questions referring to household issues are answered by an appointed household representative. In 1994 the survey format changed so that for the first time since unification East and West German households received identical questions homogenized into one single questionnaire. This implied that only from 1994 onward the question which objectively identifies if a person is legally classified as disable was identical for both East and West Germany. It is for this reason that we select our waves from 1994 to the most recent wave (2000).[13] In the same section of the questionnaire a second question identifies the degree of disability of individuals who have had a disability assessment. This and related disability questions in the survey are replicated in Appendix A. According to Section 2, an individuals can benefit from policies on disability with respect to labour market outcomes if they are assigned a degree of disability of 50% or greater. However, those with a degree between 30 and 50% also fall within the benefits of the policy, and therefore, in our empirical section we identify individuals as disable if they declare a degree of disability equal or greater to 30%.[14]

Our target population is the permanent inhabitants of Germany after the process of unification, thus we draw from the Samples A (West Germany), B (working immigrants), C (East Germany) and D (new immigrants since 1984). Combining these four different samples implies the use of weights for our sample to be representative of the population.[15]

---

12 For a more detail account of the structure and contents of the panel visit www.diw.de .

13 At the time of editing 2000 was the most recent wave available.

14 The fact that there is a clear cut distinction between those in the 30% – 50% group and those in the 50% – 100% does not imply that the second group are the only ones to benefit mostly from the SDPA act. In fact, falling within the Act is discretionary and depends very much as the labour offices as the implementing institution. Semlinger (1995) shows that it is sufficient to show *some* kind of disability as permanently reducing the chances of integration into working life to benefit fully from legislation in the SDPA. Intuitively, if an individual *voluntarily* submits for an assessment on the degree of disability, we would expect that he or she is already aware of the benefits, if only because of tax incentives and disability allowances. It is therefore very plausible to assume that anyone who has been diagnosed with a degree of disability of 30% or above is treated equally as anyone with a degree of 50% or above.

15 Anticipating the text below, the empirical analysis does not make use of the time-series structure of the data. Therefore, we weight our sample using cross sectional weighs, as provided by the GSOEP.

Our selection criteria is based on individual respondents, of ages between 17 and 60 years old, excluding those in full time education and individuals performing military or civil service. This criteria leads to 12,757 unique observations over the seven year period. With our selection criteria, our definition of non-participant on labour market activities for both, the disable and the non-disable sample, includes registered unemployed, not employed who declare housework as main activity and those who declare to be on early retirement. Factors such as inadequate information channels, motivation (e.g., inadequate policies do not provide enough motivation for the disable to participate), etc., might result on non-participant disable persons opting not to register as unemployed. Thus, to avoid selecting on characteristics correlated to the efficacy of the policies on disability/labour, we focus on non-participation rates rather than unemployment rate. Besides participation versus non-participation in paid labour market activities, we also look at yearly labour earnings as a second measure of labour market outcomes.

3.1 Constructing the control and treatment groups

One way to observe the effect of disability (i.e., the effect of policies on disability/labour), is to examine the labour market outcomes at time t for individuals who became disable at time $t-s$, where s is a sufficiently large elapse of time to justify the adaptation of such individuals to the new health status, the workings of disability policies that help disable back into paid labour market activities, and/or a combination of the two. It is to this aim that we use several waves of the panel. We define an individual $i \in n$ (where n is the sample size) as a treatment unit, if such person is non-disable at time t_1, becomes disable at t_2 and remains classified as disable at t_3. We identify such unit with the mnemonic ADD_i. Individuals in the ADD group can receive the treatment of the policies in the second time period t_2, but we can assume that it is only in the third time period t_3 (and beyond) when both the policies and the individual's adaptation to the new status will have had an impact on their labour market outcome. With annual data we require at least three waves to construct the treatment group. The use of seven waves from the GSOEP (1994 to 2000) allows the formation of 5 sequences (S_1 to S_5) of 3 years each. Having more than one sequence increases the number of observed treatment units, thus increasing the precision with which we estimate the impact of disability on labour market outcomes. The control (or untreated) group is defines by individuals who declare themselves as non-disable at t_1, t_2 and t_3 at any given sequence, and therefore do not receive the impact of the policy. We define these control individuals with the mnemonic AAA. The idea is to use adequate statistical tools to perform an appropriate comparison of the labour market outcomes of individuals in AAA versus those in ADD at t_3.

Table 2 shows the dynamics of the data, the formation of the sequences and the possible combinations between the treated and the untreated samples over time. This table shows that an individual who is a control in S_j for j = [1,5], can also be a control unit for any of the other four sequences. For example, an individual who is a control in S_1, who is further observed as been non-disable in 1996, is also counted as a control unit at S_2. However, individuals who are observed as controls in various sequences, count as independent observations for each of the different sequences. The reason for this is twofold. First, at S_1 we are interested on labour market outcomes in 1996, whereas at S_2 the labour outcomes of interest are those observed in 1997; a similar argument applies to the comparison of any other two of the 5 three years sequences. Second, comparing labour market outcomes at t_3 between control and treatment units can only be done within each sequence. The reason is that macroeconomic conditions might change over time, thus affecting the outcome variables at t_3. If so, comparing the outcomes of treated units in S_j at t_3 with the outcomes of control units in $S_{k\neq j}$ at t_3, might result in biased estimates due to within year changes in the economy as a whole. With these arguments, it becomes clear that at the point of estimation, and due to the construction of the control and treatment units, the estimation needs to be conducted independently for each sequence, while the final estimate of the effect of disability on labour market outcomes is based on the average over all the (sequence based) independent estimates.[16]

Although it is possible for an individual to contribute as a control unit at each an every one of the five sequences considered, this is not the case for the treatment units. For example, by construction, a treatment unit at S_1, cannot be a treatment unit at S_2. In theory, it is possible to observe treatment unites at S_1 further participating as treatment units at S_4 or S_5 and likewise for treatment units at S_2 with respect to S_5. In practice, after constructing our samples of controls and treatments, we observe a combination

16 See Sections 4 and 5 for further details on the estimation techniques, and Appendix B for the algorithm followed in the estimation procedure.

of these possibilities in 6 occasions. We assume these are data coding errors and do not use them in our empirical analysis.

Table 2: Definition of Treatment and Comparison Groups.

TREATMENT SAMPLE: ADD							
	1994	1995	1996	1997	1998	1999	2000
S_1 [1994-1996]	$A(t_1)$	$D(t_2)$	$D(t_3)$				
S_2 [1995-1997]		$A(t_1)$	$D(t_2)$	$D(t_3)$			
S_3 [1996-1998]			$A(t_1)$	$D(t_2)$	$D(t_3)$		
S_4 [1997-1999]				$A(t_1)$	$D(t_2)$	$D(t_3)$	
S_5 [1998-2000]					$A(t_1)$	$D(t_2)$	$D(t_3)$
COMPARISON (untreated) SAMPLE: AAA							
	1994	1995	1996	1997	1998	1999	2000
S_1 [1994-1996]	$A(t_1)$	$A(t_2)$	$A(t_3)$				
S_2 [1995-1997]		$A(t_1)$	$A(t_2)$	$A(t_3)$			
S_3 [1996-1998]			$A(t_1)$	$A(t_2)$	$A(t_3)$		
S_4 [1997-1999]				$A(t_1)$	$A(t_2)$	$A(t_3)$	
S_5 [1998-2000]					$A(t_1)$	$A(t_2)$	$A(t_3)$

It is clear from Table 2 that anyone in our selected sample who, over the seven year period, shows a pattern between non-disability (A) or disability (D) which does not allow for either sequence AAA or ADD – or a combination of the two – at least once, will not be used in the empirical analysis.[17] Thus of the 12,757 unique individuals who entered the sample at some point between 1994 and 2000, only 10,589 individuals contribute to the formation of the control and treatment groups. Table 3 shows the distribution of the 10,589 observations according to year of entry, attrition, net cumulative number of observations per year, and the population analogue.

17 For example, an individual can show a pattern DDDDDDD over the seven year period. This individual, who is disable throughout the period under study, would be taken out of our sample. It is not only that we cannot identify this individual as either control or treatment, but because of the lack of pre-disability data we would be unable to

Table 3: Distribution of observation according to time of entry and exit from the survey.

	New Entries	Old (new) Entries	Attrition	Total cumulative number of observations	Population analogue.
1994	8619	-	-	8619	39.4 mill.
1995	904	-	133	9390	41.8 mill.
1996	428	102	670	9250	42.2 mill.
1997	329	171	786	8964	42.1 mill.
1998	309	171	825	8619	42.0 mill.
1999	0	176	767	8028	41.1 mill.
2000	0	121	615	7534	39.9 mill.

Note 1: New entries refer to individuals who enter the panel for the first time a that year, assuming for our purpose that the first year of the panel is 1994.

Note 2: Old (new) entries refer to individuals who where part of the panel at some point since 1994, drop the sample for one or more waves, and re-enter the sample at some point in the future.

Note 3: The cumulative number of observations is read as the total number in the previous year, minus attrition, plus new and old(new) entries.

Note 4: The population estimates are based on the weighted sample using yearly cross-sectional weights provided by the GSOEP.

The last column in Table 3 shows the weighted sample using cross section weights. Comparing these estimates to the estimates of labour force participation in Germany provided by the International Labour Organization (ILO) suggest that our selection criteria produces a sample representative of the active labour force in Germany for the period under consideration.[18] Table 4 examines the distribution of the net entries per year between states of health (disability versus non-disability) and labour market participation. The percentage of disable in the sample seem to increase steadily over time. In the year 2000 the figures shows how disable are three times more likely to be represented in the sample than in 1994. This change might be partly explained due to a change in wording of the question from 1998 onwards (see Appendix A). Other reasons are that the natural overtime erosion of households in the panel might not have a greater (relative) affected on non-disable individuals rather than disable ones. What is clear from Table 4 is that non-participation in labour market activities is more prominent for the disable, relative to non-disable, with such differentials maintained over time. However, although Table 4 shows that disable are double more likely to be outside the workplace than non-disable, these figures do

identify the effect of the policies on her outcome in the labour market.
18 See www.laborstat.ilo.org for further details.

not show a causal relation between disability and labour market outcome, i.e., the between group's comparisons are not adjusted for different distributions of background characteristics.[19]

Table 4: Non-participation, disability and interaction between health and labour market status.

Year	Total number of observations per year	Share of disable in %	Share of non-participants in %		
	Full sample	Full sample	Full sample	Disable sample	Non-disable sample
1994	8619	1.1 (0.12)	20.0 (0.43)	37.6 (5.2)	19.8 (0.43)
1995	9390	1.9 (0.14)	19.2 (0.41)	31.2 (3.7)	19.0 (0.41)
1996	9250	2.6 (0.17)	19.8 (0.41)	38.2 (3.4)	19.3 (0.32)
1997	8964	3.0 (0.18)	20.2 (0.42)	44.8 (3.2)	19.4 (0.42)
1998	8619	3.5 (0.20)	21.0 (0.44)	43.5 (3.0)	20.0 (0.44)
1999	8028	4.2 (0.22)	22.0 (0.46)	44.4 (2.9)	21.0 (0.46)
2000	7534	4.5 (0.24)	21.0 (0.47)	45.4 (2.8)	20.0 (0.47)

Note: Not-employed are those defined as either officially registered unemployed, those declaring full time housework and early retired.

Our second measure of labour market outcomes is annual labour income. Such measure is generated by the GSOEP data management by working on calendar data as provided by respondents. Non-participating individuals are assigned zero labour income. Table 5 shows the distribution of this variable over time and between samples (disable and non-disable). We take into account all the sample (participants and non-participants), such that averages reflect the overall distribution with non-participation rate acting as a penalizing weight for within sample labour income distribution. The variable income is expressed in Deutsch Marks with base 1999. First, column 4 shows that, for all the sample, real labour income has experienced very little change over time. However, comparison of this trend between disable and non-disable shows that contrary to the trend in the overall population, the disable have experience a drop in real income; by inspection, the drop seems to be not significantly different than zero. Second, a year per year comparison between the two health status shows that non-disable

19 Table 4 present only part of the picture. If anything, the figures with respect to the population of disables is a lower bound. The reason is that Table 4 is based on observations net from attrition. Appendix C , Table C1 shows similar estimates for attrition units with information based on pre-attrition year. This table shows that attrition is not random with respect to disability status, i.e., for any of the years considered, the percentage of disable who leave the sample is at least 1.5 times greater than the percentage of disable remaining in the sample. A future extension of our work is based on taking into account attrition in the sample, thus estimating bounds on the impact of disability on

are significantly higher earners than disable persons, with an average annual labour income for non-disable that can be up to as much as 25% higher than the average earnings of disable individuals. Third, comparing the median to the average the overall picture shows that the labour income distribution of disable persons is skew to the left, relative to non-disable, thus illustrating higher degree of earnings inequality among the disable than non-disable (estimates of the relative inter-quartile range confirms this fact).

Table 5: Annual labour income (waves 1994-2000)

	Full Sample	Non-Disable	Disable
1994			
Average (s.d)	38,820 (36,840)	38,840 (36,830)	37,420 (37,650)
Median	36,300	36,360	27,930
IQR/Q50	1.28	1.28	2.13
1995			
Average (s.d)	38,110 (36,250)	38,170 (36,300)	35,270 (33,340)
Median	35,950	36,000	31,670
IQR/Q50	1.33	1.32	1.83
1996			
Average (s.d)	39,750 (37,020)	39,870 (37,090)	35,540 (33,970)
Median	37,350	37,350	30,820
IQR/Q50	1.33	1.33	1.88
1997			
Average (s.d)	39,860 (36,510)	40,130 (36,610)	31,030 (31,920)
Median	37,360	37,560	27,430
IQR/Q50	1.36	1.35	1.85
1998			
Average (s.d)	39,260 (37,320)	39,570 (37,460)	30,600 (32,170)
Median	36,320	36,360	29,060
IQR/Q50	1.40	1.41	1.75
1999			
Average (s.d)	40,040 (37,770)	40,400 (37,850)	31,590 (34,820)
Median	36,500	36,690	21,470
IQR/Q50	1.42	1.40	2.45
2000			
Average (s.d)	42,700 (39,550)	43,080 (39,500)	34,840 (39,950)
Median	39,730	40,040	27,010
IQR/Q50	1.30	1.27	2.29

Note 1: IQR/Q50 refers to the relative Inter-Quartile Range, weighted by the median Q50. If IQR/Q50=1 it reflects perfect equality, with IQR/Q50 $\in [1,\infty)$.

Note 2: All number are weighted estimates in Deutsch Mark units, with base year 1999.

As was the case when interpreting Table 4, Table 5 does not provide a causal relation between disability and labour market outcome (in this case labour earnings), but rather it summarizes evidence of possible differentials between disable and non-disable persons, which needs to be studied with

labour market outcomes, where such bounds will account for attrition error.

appropriate statistical tools.[20]

The sample of 10,589 form the basis for the construction of the treatment and control (untreated) groups. Table 6 shows how the dynamic changes between health states (non-disability and disability) over the seven years leads to the distribution of controls and treatment units with respect to the five three-year sequences.

Table 6: Distribution of sequential valid observations between treated and untreated (control) groups, according to sequences.

Sequences	[t_1, t_2, t_3]	Total number of sample points	Total in the untreated sample, i.e., controls [AAA]	Total in the treated sample, i.e., treatments [ADD]
S1	[1994, 1995, 1996]	7,666	7,581	85
S2	[1995, 1996, 1997]	7,727	7,659	68
S3	[1996, 1997, 1998]	7,420	7,351	69
S4	[1997, 1998, 1999]	7,131	7,074	57
S5	[1998, 1999, 2000]	6,964	6,904	60
Total		**36,908**	**36,569**	**339**

As expected, the number of controls far outweighs the number of treated units. The situation can only be improved as more waves become available. Nevertheless we believe that 339 observations who receive the treatment (the impact of the policies) over a two year period, might be sufficient to make inference on the impact of such treatment with a significant degree of confidence. The vast number of controls simply reflect the overlapping possibility for control individuals to be counted as multiple units.

Table 7 shows some comparative statistics for a selected set of covariates comparing the 36,569 controls to the 339 treatment units as defined in Table 6.[21] Instead of comparing them in a given point in time, we compare these two sets of individuals with respect to changes over time, that is, the average changes between t_1 and t_3 over the five sequences, for each sample.

20 Appendix C, Table C2 shows a table similar to Table 5 but accounting only the labour income of those who are assigned a positive amount; by construction these are the sample of employed (participants) in the population.

21 The comparison is based on the average of the per sequence estimates, thus treating each sequence as an independent set of information, as previously suggested in page 16.

Table 7: Change over time (between t_1 and t_3, average over all sequences), for outcomes and a selected set of control variables.

		All observations	Comparison Group [$A(t_1)A(t_2)A(t_3)$]	Treatment Group [$A(t_1)D(t_2)D(t_3)$]
Annual Labour income				
	Increased	0.483	0.483	0.542
	Stayed the same	0.068	0.069	0.039
	Decreased	0.448	0.448	0.419
Number of hours working.	Increased	0.298	0.298	0.339
	Stayed the same	0.360	0.360	0.406
	Decreased	0.342	0.343	0.255
Employment status (E= Employed) (U=Not-Employed)	From E to U	0.078	0.077	0.138
	Stayed E over time	0.125	0.124	0.239
	Stayed U over time	0.737	0.739	0.590
	From U to E	0.060	0.060	0.033
Satisfaction with work (Subjective)	Increased	0.315	0.316	0.244
	Stayed the same	0.318	0.317	0.379
	Decreased	0.367	0.367	0.378
Satisfaction with health (Subjective)	Increased	0.320	0.319	0.373
	Stayed the same	0.303	0.304	0.199
	Decreased	0.377	0.376	0.429
Hunting for work behaviour	Hunt at t_1, but no hunt at t_3	0.044	0.044	0.029
	No Hunt at t_1, No hunt at t_3	0.905	0.905	0.916
	Hunt at t_1, Hunt at t_3	0.015	0.015	0.019
	No Hunt at t_1, but hunt at t_3	0.036	0.036	0.036
Household owner (Acquire can be either inherited or bought)	Sold between t_1 and t_3	0.170	0.171	0.134
	Remained not owner	0.369	0.368	0.448
	Remained owner	0.216	0.217	0.166
	Acquire between t_1 and t_3	0.244	0.244	0.252

To some extend, Table 7 provides some evidence on the impact of health status over time, for such selection of characteristics (labour income, hours worked per week, employment status, two measures of subjective evaluation, job searching behaviour and wealth – in the form of household ownership).

When making inference from Table 7 we bear in mind that these estimates do not control for factors that might have a simultaneous affect on both health and a given (variable) characteristic. Estimates in this table suggest that a larger percentage of people in the treatment group see their labour income increase relative to the non-disable group. Likewise, the percentage of individuals who experience an increase in the number of working hours per week is greater in the disable than for the non-disable sample. In both cases (for income and hours work), the difference in the increase percentages between the two groups relatively small and insignificant. Flows in and outside employment for each of the two groups shows that those who become disable over time are more likely to flow into unemployment (following a period of employment), as well as half as likely to leave unemployment status (towards employment) than their non-disable counterpart. This results in a larger net stock of unemployed (over the three year period) for those in the treatment group, relative to the control group. In this case differentials between the two groups are significantly different than zero, thus, a treatment unit has a significantly larger probability to go from employment towards unemployment, to remain unemployed over time, and is significantly less likely to either stay employed over time or to find employment after a period of unemployment. Treatment units show a significant decrease in their satisfaction with health than the control group, whereas changes with respect to satisfaction at work are not dissimilar between the two samples. The same is true with respect to job hunting behaviour: this does not differ significantly between the two groups. As a proxy for wealth we take ownership of household. This estimate shows that the treatment sample are more likely to remain not-owners, while the selling and acquiring behaviour is not dissimilar among the two groups.

An additional characteristic for the treatment group is the distribution of their degree of disability.[22] Table 8 shows how the distribution of the 329 treatment units with reference to 10%

22 Notice that the control group are defined as individuals who do not receive the impact of the policies as defined by the SDPA, which means that they might answer 'no' to the first question in Section A, but it is also possible for these individuals to answer 'yes', and declare a degree of disability below 30%. However, the number of individuals who declare some degree of disability in this group is very low. For example, for the control sample in the sequence S_1, (1994-1996), at t_3 only 48 out of 7,666 individuals declare a degree of disability, accounting for 0.8% of the

increments in the degree of disability.[23] In Table 8, the second column shows the distribution among degree categories at the point at which these individuals became disable (at t_2) and column 3 shows the distribution of the degree over at t_3. Comparing this two columns indicates that the diagnosis (assigned degree of disability) changes over time with a rightwards shift in the upper tail, reflecting that those diagnosed as disable at t_2 have their diagnosis upgraded in subsequent years.

Table 8: Distribution of degree of disabilities

Degree of Disability	**Distribution between degree brackets at t_2**	**Distribution between degree brackets at t_3**
30 # degree < 40	42.3	40.3
40 # degree < 50	14.6	11.5
50 # degree < 60	16.8	22.7
60 # degree < 70	11.0	9.6
70 # degree < 80	8.8	1.0
80 # degree < 90	4.1	3.5
Equal to 90	1.6	1.6
Equal to 100	0.8	0.8

Note: Estimates based on 329 treatment units.

Whether we look at t_2 or t_3, it is clear that the majority concentrate between 30% and 70%. We might think of individuals with a degree of 90 or 100% as problematic, since it would be difficult to justify their competitiveness within the labour market. In our weighted sample of 329 individuals, there are 3 individuals with a degree equal to 90% and 4 individuals with 100% (total disability). Table C4, Appendix C, shows some of the characteristics for each of these individuals. Only two of these are females. Their ages show that only one approximates retirement age, and while all declare to be employed, this is always in the manufacturing/production industry (code 7 in the 1-digit ISCO

(weighted) sample within such control group, with a (weighted) mean degree of disability equal to 18.7%. Table C3 in Appendix C shows similar estimates for all the five 3-year sequences. Although we could think of these individuals as a particular type of controls (thus allowing for three different states, namely controls, controls with disability and no policies, treatment units with disability who receive the policies), in this firs t version of the paper we consider controls with disability in the (0,30) window to be units with similar labour market conditions as controls with zero disability.

23 Notice that although Table 6 shows a total number of treatments as 339, ten of these individuals are given a weight zero with reference to the cross-section weights in their corresponding t_3 year.

classification), with one exception (who is a sales worker). All of them show earnings well above the average of their group (compare column 9 in Table C5 to corresponding years in Column 4, Table 5). However, these incomes might be artificially inflated with associated benefits and special needs in accordance to their disabilities. Nevertheless, their contribution with respect to hours per week at work suggest that these are mostly full time workers (estimates based on annual hours minus 6 weeks holidays). Their characteristics do not suggest that we should take them out of the sample, as we have no indication on the qualitative characteristics of their disability.

3.2 Control variables

The GSOEP provides rich quality data at the individual level. The panel contains information on key variables that can drive the chances of an individual becoming disable while having an effect on individual's labour market outcome. These variables are needed because it is conditioning on them which allows for the assumption of independence between labour market outcome in the control group and disability treatment.[24] The selection of such variables could be done by following some international guidance on the classification of causes or underlying conditions on disability, for example, "*The WHO international Classification of Disease and Related Health Problems (ICD-10)*". This report has been used by national health surveys and census (e.g., 1993 Australian Survey on Disability, 1996 household disability survey in New Zealand or 1998 Netherlands Health Interview Survey, but to mention a few) include a question, very similar in all countries, which allows individuals to classify their impairment/limitation among a set of categories.[25] Based on examination of health surveys in various countries and over time, the United Nations Statistical Division[26] has proposed a short list for classifying causes of disablement which includes three categories relating to genetics and acquire diseases (infectious and parasitic diseases, congenital anomalies and perinatal conditions, other diseases related

24 See Section 4.

25 The question differs among countries with respect to the detail given to each category. For example, in the 1998 Netherlands survey, individuals are asked to classify disability via illness between congenital or occurring at birth, illness of childhood, or illness of old age; the 1996 New Zealand survey is similar but provides further categories by age groups, distinguishing with reference to illness due to either psychological or physical abuse.

26 "Guidance and Principles for the development of disability Statistics"(2001?), UN Department of Economics and Social Affairs, UN Publications, Statistical Division, http://www.un.org/depts.unsd

conditions), four categories with reference to external injuries (motor vehicle accidents, other transport accidents, accidental poisoning, injuries from activities - falls, fires and wars), and a category which includes all disability causes related to environmental factors.

We have followed such guidelines in order to select variables related to the UN guidelines. In particular, we think of five categories of variables which might be important in our analysis, named as 'Traffic', 'Genetics and objective health', 'Labour market classification', 'Leisure' and 'Demographic and social-economic status (SES) variables'. In terms of traffic we think that the degree of urbanization is positively correlated with the chances of becoming disable, while urbanization also has an effect on overall participation rate and labour income. By 'Genetics' we mean any endowment which parents can pass on to their children that might affect both the chances of the child to become disable as an adult, but also their work status. Variables such as parental education would enter this category. One could assume that, on average, parents with more education are better at transferring information on safety to their children (e.g., using seatbelts when driving) that will effect the probability of the child on becoming disable once the child becomes an adult. Another example is that parent with more education are better at processing information, including the importance of nutritional needs of growing children, so that parental education may have a direct impact on the child's capacity to avoid illness in adulthood which are associate with poor environmental growing-up conditions. At the same time, parental education has a direct impact on the child's ability and education level, therefore directly affecting the child's work status once they reach adulthood. Genetics would also include either objective or subjective measures of health, since these are also important determinants of both the labour market outcome and health status. The category 'labour market status' includes variables such as occupational category (e.g., blue collar), as well as variables specifically related to work and work activities (e.g., job search behaviour, number of hours at work). The category 'Leisure' refers to activities such as sports or any other variable that indicates activities outside the workplace; we can think that such activities are often associated with the status of individuals (i.e., income, availability of time), while the risk element in leisure activities can also affect disability status. Finally, demographics and SES variables, for example, family size, age, education, marital status, etc. are also important controls for our analysis. Table 9 list the available information, organizing variables according to categories while Appendix D defines such variables as well as providing summary statistics for all samples across time.

Table 9: Distribution of covariates within 5 different categories.

	Traffic	**Health measures & Genetics (parental background)**	**Labour market classification**	**Leisure**	**Household and Socio-Economic Status**
	-Degree Urbanization [1]	-Subjective evaluation of Satisfaction with health -Number days sick off work -Visit doctor due to work injury -Gender -Father's level of education (low, medium or high)[2] -Mother's level of education (low, medium or high)[2]	-Satisfaction with work. -Blue collar classification (low skill, medium skill, high skill). -Civil servant classification (low skill, medium skill, high skill). -White collar classification (low skill, medium skill, high skill). -Job search behaviour -ISCO one digit code (Risk of injury at work) -Size of company (no. of employees) -Years working for company. -Number of hours worked (annual)	-Sports practice behaviour (sport)	-Holder of personal private health insurance -A, B, C, D sample -Country of origin (Germany, E.C, World) -Education -Family Size -Number of kids -Age at time of survey -Marital Status -Labour income, annual. -Per capita household income (pre and post government) -Owner of household at time of survey

Note 1: Drawn from the German Version of the GSOEP data set, to be included in future versions of the paper.

The GSOEP is very good for detailed information with respect to the category 'Labour market status', demographics and SES. The section on health in the questionnaire provides objective information on disability for us to identify legally disable individuals with zero ambiguity.[27] At the same

27 It improves many studies of disability where the degree of disability is often not well defined and subjectively interpreted by either researcher, data collection techniques or subjective evaluation of surveyed units. See, for example, Kreider and Pepper (2002) which provides a very good discussion on how to analyse disability when "true" disability is in fact unobserved, and how estimates can change with changing interpretations of subjective valuation of health status.

time this section has information on individual's demand for health and some other subjective evaluation on their health status. In terms of parental information, it is possible to construct a categorical index for the parental level of education combining four variables (two for the mother and two for the father) which identify either the highest level of education of the each parent or their training within the workplace (see Appendix D). Unfortunately the panel provide information on habits (i.e., drinking, smoking and diet) for only two waves (1998 and 1999). Although these health habits might certainly have an impact on health, as well as been related to labour market outcomes, due to the way in which we construct the control and treatment groups, the information in the panel at this point is insufficient to enter our analysis. We do however have a variable with respect to sport practice behaviour.

4 Identification issues and the Parameter of interest

The question we aim to answer is "What is the effect of becoming disable, for those who become disable, on their labour market outcomes, compared to the hypothetical state of not having received the impact of disability?" This question targets the causal relation from disability to outcome (i.e., labour earnings, participation), and can be answered using Rubin (1974) potential outcome approach to causality.

The units of interest are individuals $i \in n$ observed over a sequence of three consecutive years, where n is the sample size for a given number of sequences. For each unit we have health status information (either disable or non-disable) at each of the three years of the sequence. The dynamic (un-) change between non-disability and disability defines two possible states of the world, namely the treated and the control state. Let S_i be a binary assignment indicator that determines whether unit i gets the treatment ($S_i = d$) or not ($S_i = a$), and let Y_i^d and Y_i^a be the potential labour market outcomes associated with the treated and untreated (or control) states, respectively. The notion "potential" is used to emphasis that only one of $\left(Y^d, Y^a\right)$ is observed for every unit in the sample. Each individual $i \in n$ is identified as non-disable in period 1 (t_1). The sample is constructed such that an individual observed to be disable at t_2 is also observed as disable at t_3: Y_i^d is the actual (observed) outcome (Y_i) at t_3

associated with an individual $i \in n$ with such health pattern. Likewise, the actual outcome (Y_i) at t_3 is Y_i^a for a unit observed to be non-disable at t_2 who, also by construction, is observed to be non-disable at t_3.

Our parameter of interest, J^0, is the mean effect (at t_3) of receiving the impact of a disability shock, rather than not receiving the shock, on those individuals who having become disable at t_2 do in fact receive the impact of such an status thereafter (e.g., receiving the impact of policies aimed a the disable, modification of behaviour with respect to labour market activities, etc). This parameter is known in the literature as the average effect of the "treatment on the treated" (ATET), and can be expressed as:

$$J^0 = E[Y^d - Y^a \mid S = d] = E[Y^d \mid S = d] - E[Y^a \mid S = d] \tag{1}$$

Clearly, J^0 is not identified by the data, since identification of the causal effect would require the observation of the counterfactual outcome Y_i^a to Y_i^d for each i unit in the treated sub-sample, that is, the counterfactual $E[Y^a \mid S = d]$. Assuming that the probability of becoming disable is the same for those observed as becoming disable and those who did not become disable, (i.e., $E[Y^d \mid S = 1] = E[Y^d \mid S = 0]$), would solve the problem, since the untreated sample would be used as counterfactual for the treated units. However, in light of the evidence discussed in the previous section, the rather strong exogeneity assumption (or random selection) is violated in our context, specially with characteristics such as occupational sector, which are key determinants of labour market outcome, clearly differing with respect to incidence of disability.[28]

Section 3 suggested that our data was very informative with reference to observed characteristics which might determined both health state and outcomes. Assume there exist such set of observed characteristics given by the vector X which affects both health status (i.e., the chance to become disable) and labour market outcomes (earning and participation status). If X is both sufficiently informative and unaffected by the treatment itself, identification of J^0 is possible since conditioning on X implies that within sub-groups (as defined by X), being a control (or not) is unrelated to what the

28 See Table 1, Section 2.

outcome would had been if you had become disable (or not). This assumption, weaker than exogeneity, is known as the Conditional Independence Assumption (CIA) and is formally given by:

$$(Y^a, Y^d) \perp S \mid X = x; \quad \forall x \in c; \quad c \subset \mathbb{R}^p \tag{2}$$

Therefore, $E[Y^a \mid S = d, X] = E[Y^a \mid S = a, X]$ and J^0 is identified such that,

$$J^0 = E[Y^d \mid S = d] - \underset{X \mid S=d}{E}[E[Y^a \mid S = a, X = x] \mid S = d] \tag{3}$$

The CIA is a workable assumption as long as it holds for the available X set, but does not account for unobserved characteristics that may also play a role in selection.[29] As previously suggested, one further condition for the implementation of the CIA is that all characteristics in the set X has to be unaffected by the treatment itself; a violation of these would lead to endogeneity between control and outcome variables. For example, an individual who becomes disable at t_2 may decide to engage in further education, a decision that might not have come about without the disability shock, while such decision will probably affect her labour market outcome at t_3. In this case, if we allowed for years of education to enter the conditioning set X in (3) we would end up with endogeneity problems because the treatment, which affects the outcome, determines the controls. To avoid this problem we need to use a set X which is not influence by the treatment. In our case, this is already given when we construct our treatment and control units, i.e., by construction all individuals (treated and control) are non-disable in period t_1, and, therefore, conditioning on X at t_1 (i.e., condition on X_1) implies that the exogeneity condition necessary for the implementation of the CIA is fulfilled. With this, a more complete version of (3) is given by:

$$J^0 = E[Y_3^d \mid S = d] - \underset{X_1 \mid S=d}{E}[E[Y_3^a \mid S = a, X_1 = x_1] \mid S = d] \tag{4}$$

29 See Heckman and Siegelman (1993) for a view of the effect of unobserved on matching methods.

where (4) implies that we are interested on comparing outcomes at t_3 between treatment and comparison units, given that they shared similar characteristics at the point where both control and treatments where in one single state of the world (non-disability).

The parameter J^0 can be estimated using the sample analogue, provided that for every treated unit there is a comparison unit in the control sample with similar X_1 characteristics. This is known as the common support condition, which for ATET in our particular application is defined as $P(S=d \mid X_1 = x_1) < 1; \quad \forall x_1 \in c_1 \subset \mathbb{R}^p$. The implication is that there is a common overlap between the distributions of the set X_1 in the two states.

When X_1 is of high dimension, estimates of $E[Y^a \mid S=a, X_1]$ using distribution free techniques such as Kernel based nonparametric methods, raises the problem of the curse of dimensionality, (i.e., very low density per cell), therefore increasing imprecision on estimates, specially at the tails of the distribution. To overcome this problem Rosenbaum and Rubin (1983) show that it is not necessary to compare observations with the same value of X_1, but it is sufficient to compare observations having the same conditional treatment probability, $P(S=d \mid X_1 = x_1) = p(x_1)$, where $p(x_1)$ is also known in the literature as "the propensity score". Conditioning on $p(x_1)$ rather than X_1 itself reduces the problem to a one dimensional problem so that estimate of J^0 will be based on:

$$J^0 = E[Y_3^d \mid S=d] - E_{p(x)|S=d}\left[E[Y_3^a \mid S=a, p(X_1) = p(x_1)] \mid S=d \right] \tag{5}$$

4.1 Matching Methods

Matching on the propensity score can lead to different estimates according to which matching method is used. The key difference among these methods is the weight each method assign to each observation in the control (or comparison group). All matching methods are based on the following form:

$$E_n[Y^a \mid P_n(X_i)] = \sum_{j=1}^{J} w\left(p_{ni}(x_i),\ p_{nj}(x_j) \right) Y_j^a \tag{6}$$

where $j = 1,...,J$ is the index for the control group and $i = 1,....,I$ is the treatment group. The expectation in (6) is taken over all J for each i^{th} individual in the treatment group, therefore the counterfactual outcome for each treated unit is a weighted average of the outcome of the untreated group. Different weighting methods (i.e., matching methods) implies different ways to weight the potential counterfactual observations, but also different treatment of the common support problem.

The empirical section in this paper uses alternative variations of two different matching methods, leading to six estimates on the average treatment effect on the treated for each of two outcomes (labour market participation and labour income).

The first matching methods consists on assigning a weight of one to the observation in the non-disable group with the closest propensity score to each treated observation, and zero to all other observations in the non-disable group. This method is known as matching by the nearest neighbour, with $w(.)$ in (6) expressed as:

$$w(p_n(x_i), p_n(x_j)) = \begin{cases} 1\, if \;\; j = \underset{k \in \{S=0\}}{\text{argmin}} \{ | \, p_n(x_i) - p_n(x_k) \, | \} \\ \quad\quad 0 \quad otherwise. \end{cases} \tag{7}$$

In cases where there is a large overlap between the distribution of the two estimated propensity scores, it is common practice to follow the statistical literature and match without replacement; each observation in the control group is used at most once. In cases where the overlap is small, throwing away observations might lead to the violation of the CIA, thus matching with replacement might be a more appropriate practice (see Black and Smith (2002) for an example). Weighting using the matching technique as given by (7) leads to different estimates according to how the absolute distance (between estimated proportions) treats the common support. As it stands, estimates based on (7) does not specify what constitutes an appropriate distance before an observation in the control group becomes a counterfactual for an observation in the treatment group, and therefore does not take the common support problem into account. For example, if the absolute distance between estimated propensities equals 0.90, and such happens to be the minimum distance between an i^{th} treated unit and all the control units, matching by (7) will assign that counterfactual to such i^{th} observation, even if by such distance it is likely to be an bad match. To solve this problem, an alternative based on (7) is to match such that $j = \text{argmin}_{k \in \{S=0\}} \{| p_i(x) - p_j(x)| \leq c\}$, where c is defined as a "caliper". For example, if $c = 0.10$, all observations in the treatment group without a comparison observation in the control group which is plus or minus 0.10 distance (with respect to the estimated propensity score) is not accounted for in the final estimate. Where the common support is not well supported by the data, many treated units would need to be thrown away, with and estimate that might differ considerably from the original estimate based on (7), since instead of estimating the average treatment effect on the treated, the estimate is the average treatment effect on the treated who are confined to a particular common support relative to the control group (see Black and Smith (2002) for an empirical example of this).

The second matching method considered in the empirical section is kernel matching, with analogue to (7) given by:

$$w(p_n(x_i), p_n(x_j)) = \frac{K\left[\dfrac{p_n(x_i) - p_n(x_j)}{h_n}\right]}{\sum_{k=1}^{J}\left[\dfrac{p_n(x_i) - p_n(x_j)}{h_n}\right]} \tag{8}$$

where $K[.]$ is the kernel function and h_n stands for the bandwidth. Relative to the nearest neighbour, kernel may assign a non-zero weight to more than one observation. As with kernel regression, each observation in the control group is weighted according to distance between estimated propensity scores, where the weights are determined by distance within a subgroup of observations defined by the bandwidth. In practice, we choose the bandwidth small enough with respect to the variation in the density of the propensity score. In the case of matching perhaps what is more crucial is the choice of functional form for $K[.]$. The choice of Gaussian Kernel for $K[.]$ leads to a similar treatment of the common support as with the nearest neighbour method without caliper, since all observations in the control group are potential candidates to obtain a particular weight without restriction with reference to the relative distance between estimated propensities. On the other hand, choosing $K[.]$ as the trimmed quadratic kernel (Epanechnikov Kernel), such that $K[.] = (3/4)(1-z^2)I[|z|<1]$, where z is a function of the bandwidth, accounts for the common support problem. In this case observations at the tails of the distribution are symmetrically trimmed, according to the distance between the treated and the control group (i.e., a distance as defined by z). In the empirical section we estimate both a Gaussian and trimmed quadratic kernel, and each of this kernels is estimated twice, once using a bandwidth by rule of thumb (Silverman, 1986) and again using the bandwidth by cross-validation (Haerdle and Marron, 1985).

When using kernels, in theory all observations in the untreated group could be given a non-zero weight. The counterfactual assigned to each treated unit is the sum of the weighted observations in the control group. Because more observations are used to construct the counterfactual, the variance of the estimate is less than in the nearest neighbour case, but given that kernel uses more observations to construct the counterfactual, the distance (with respect to X) between the treated unit and final counterfactual outcome will increase and, therefore, so will the bias (relative to the nearest neighbour).

No particular matching estimator can be thought as been superior to the other. In the empirical section we answer our question (the effect of disability/disability policies on labour market outcomes), comparing estimates based on various matching methods: nearest neighbour without caliper and with two caliper (0.05 and 0.10), Gaussian kernels with two bandwidths (inspection and cross validation), and trimmed quadratic kernels with a similar choice of bandwidth. The precision of the estimated parameter is obtained by means of bootstrap, that is, the original data is re-sampled with replacement 500 times, to attain an empirical distribution of the error term. Section 5 presents a summary of these results.

5 Results

In this section we present estimates of the impact of becoming disable on labour market participation and labour earnings. These estimates are obtained by matching on the propensity score.

Because we have a binary set up (two possible states, *ADD* and *AAA*), the propensity score $P(S = d \mid X_1)$ is estimated using a Probit model, where S=d indicates ADD=1 – i.e., an non-disable individual at t_1 became disable at t_2 and is observed disable at t_3. The conditional set is indexed with *1* to indicate that we condition on pre-disability variables, thus solving the question of endogeneity between state dependent and explanatory variables (see Section 4 for a detail account). Section 3 showed that we have 5 different sequences of observations, with a set of control and treatment units in each sequence. We treat our data as having 5 independent sets of information. Therefore, we estimate the propensity score for both treatment and controls by estimating a Probit model on each of the five independent sequences, and joining the final estimates in two vectors $p_i(x)$ and $p_j(x)$ of estimates for treatment and control groups, respectively.[30] The specification of our Probit model is based on all variables as defined in Table 9, allowing for various square terms and interactions. As example, Appendix E shows the result of estimating the model for one particular sequence (1994,1995, 1996). Score test and goodness of fit test are used to find the appropriate variable specification, for each of the five sequences of observations. The choice of variables is identical for each of the sequences and corresponds to the original unrestricted model as shown in Table E1, (Appendix E).

Using the estimated Probit coefficient for each of the sequence, the propensity score estimates are based on $\hat{P}(S=1|X_{1,i})=\Phi(x_{1,i}\hat{b}_s)$ and $\hat{P}(S=1|X_{1,j})=\Phi(x_{1,j}\hat{b}_s)$ for treatment and control groups, respectively, where s indicates 'sequence'. Table 10 shows the cumulative distribution – in percentiles – for the treated and control groups.

Table 10. Distribution of Propensity Score for subgroups

	Propensity Score for the Treatment Group (ADD)	**Propensity Score for the Control Group (AAA)**
5TH Percentile	0.00251	0.00000
10TH Percentile	0.00541	0.00001
15TH Percentile	0.00870	0.00002
20TH Percentile	0.01300	0.00006
25TH Percentile	0.01833	0.00012
30TH Percentile	0.02967	0.00021
35TH Percentile	0.03483	0.00035
40TH Percentile	0.05254	0.00056
45TH Percentile	0.06124	0.00083
50TH Percentile	0.07434	0.00120
55TH Percentile	0.09461	0.00170
60TH Percentile	0.09935	0.00242
65TH Percentile	0.11413	0.00341
70TH Percentile	0.13797	0.00470
75TH Percentile	0.14556	0.00642
80TH Percentile	0.16971	0.00917
85TH Percentile	0.20817	0.01390
90TH Percentile	0.24855	0.02225
95TH Percentile	0.31680	0.04133
100TH Percentile	0.98857	0.94215

30 We could think that we have stratified the full sample according to characteristics as, for example, in Mueser et al. (2003). Our stratification is with respect to sequential time. See Appendix B for a detail account of the algorithm leading to the propensity score estimate.

Comparing the two distribution shows that, for our particular selection of covariates, the common support is well defined for the treatment group with respect to the comparison group. We need to think that matching is with respect to similarities on the p-score from the control to the treatment group. Table 10 shows that from the 60th percentile of the comparison group, the common support mimics that of the treatment group. This means that 60% of the control group are poor matches for the treatment units. However, with 35,569 units in the control group – Table 6, column 4 – for 329 treatment observations, dropping 60% will not create a problem. In fact, the only important issues is to make sure that treatment units have comparable control units, and this is the case even after deleting all controls with a propensity score in the first 6 percentiles of this group's propensity score distribution.

We use a variety of two alternative matching methods – namely, nearest neighbour and kernels –, to estimate the impact of disability (disability policies) on both participation rate and labour earnings. In all cases, and for each sequence, we are interested on testing the impact on *the outcome at* t_3. For example, if a unit ADD is define in Sequence 1, this means that we observe such individual characteristics and labour market outcomes in 1994, 1995 and 1996. Given that such person become disable in 1995, and remains so in 1996, we are interested on $E[Y_{1996}|p(X_{1994})]$, that is, the outcome Y in 1996 given pre-disability characteristics in 1994.

We match the treatment group to the control (comparison) group using the nearest neighbour method with two alternative calliper (0.05 and 0.10), and without calliper. Likewise, we estimate treatment impacts using both Gaussian and Epannechnikov Kernels, providing two alternative bandwidth, namely a normal approximation (Silverman, 1986) and an optimal bandwidth by cross-validation (H@dle and Marron,1985).[31] Table 11 summarizes the results. Columns 4 and 7 show the estimated impact of disability (disability policies) on participation rate and income, respectively. For each of the estimated impacts, the bracketed numbers show the estimate of the standard error, where these later are based on naïve bootstrap procedure which consists on re-sampling the original data 500 times with replacement (see Section 4).

Comparison of the three different nearest neighbour estimates suggest almost no difference in the

31 The bandwidth estimates are the same for either Kernel method. Whereas for the normal approximation h=1.096*F[p(xj)]*n^(-0.2), the bandwidth by cross validation is set to h=(ho)*F[p(xj)]*n^(-0.2), where ho is the optimal bandwidth which minimises 3(y-y(h))^2, with y(h) as the non-parametric estimate of the dependent variable y.

magnitude of the estimated impact, for either of the two labour market outcomes considered. Any potential difference between these comes from dropping treatment observations according to the value of the calliper. Whereas in the case of no calliper all treatment observations are used, a calliper of 0.05 implies discarding three treatment units (thus estimates are based on 326 units), whereas a calliper of 0.10 results on dropping one treatment unit only. These results simply reflect that the common support is very well defined for our data.

Comparing matching estimates between Gaussian and Epanechnikov suggest greater variance between the magnitudes of the estimates; whereas a Gaussian kernel suggest negative estimates for the impact of disability on participation rate, the analogue Epanechnikov estimates show positive magnitudes. Comparing estimates shows that the difference is due to overall lower magnitude for estimates using Kernels rather than nearest neighbour. Perhaps such difference reflect that Kernel estimates use more than one observation, thus allowing for bias in the estimates which might not be present when the counterfactual match is based on one single observation in the comparison group.

Estimates of the impact of disability on the outcome "participation rate" based on nearest neighbour and Gaussian Kernels suggest that disable persons have a participation rate between 0.1 and 4.1% lower than non-disable persons, whereas estimates based on Epanechnikov Kernel suggest that disable participation rate is slightly higher than non-disable by about 0.3%. In all cases, the difference is almost negligent, and moreover, all estimates are statistically insignificant. The suggestion, therefore, is that becoming disable has no effect on participation rate in paid labour market activities, relative to non-disable persons. Estimates of the impact of disability on the second outcome considered, i.e., annual labour income, shows a similar story. In this case, there seems to be some discrepancy with respect to the magnitude of the impact between the matching methods used. Overall, these estimates show a positive impact, thus suggesting that disable earn higher labour earnings than non-disable, ranging between DM 990 and DM 3,704. The only case where the magnitude of the impact suggest a higher labour income for non-disables is when matching is done by Gaussian Kernel with normalized bandwidth. However, comparing the normalized bandwidth (0.005 over the five sequences) with the optimal bandwidth using cross validation (0.001 average), suggests a normalized bandwidth which is too large for our particular data set, and this might lead to an increase in the bias of the Kernel estimates

---

Thus, the bandwidth differs between outcomes of interest.

(over-smoothing of the kernel density estimate). Thus, if we focus on the impact of disability on labour earnings, using nearest neighbour estimates and kernel estimates based on optimal bandwidths, our results suggest that disable can earn between DM 1,374 and DM 3,100 more than non-disable in annual labour income. The range in itself is not very large, and again, as in the case of participation rates, for any of the matching methods, the impact is statistically insignificant.

Table 11: Estimates of Impact of disability on labour market outcomes (ATET).

	OUTCOME: PARTICIPATION (PROBABILITY)			OUTCOME: ANNUAL LABOUR INCOME (DM,1999)		
NEAREST NEIGBOUR	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**
No Calliper	0.623	0.664	-0.041 (0.031)	34,540	32,470	2,070 (3,050)
Calliper=0.05	0.623	0.667	-0.044 (0.028)	34,580	32,590	1,990 (2,510)
Calliper=0.10	0.622	0.667	-0.044 (0.029)	34,480	32,600	1,920 (2,300)
GAUSSIAN KERNEL	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**
Bandwidth by normal approximation $\hat{h}=(1/s)\sum h_{s\in[1,5]}=0.005$	0.623	0.645	-0.0216 (0.025)	34,540	34,900	-365.6 (730)
Bandwidth by Cross Validation $\hat{h}=(1/s)\sum h_{s\in[1,5]}=0.001$	0.623	0.625	-0.0013 (0.010)	34,540	33,160	1,374 (890)
EPAN'KOV KERNEL	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**	**$E[Y^{ADD}\|S=d]$**	**$E[Y^{AAA}\|S=d]$**	**$\hat{J}$ (impact estimate)**
Bandwidth by normal approximation $\hat{h}=(1/s)\sum h_{s\in[1,5]}=0.005$	0.623	0.620	0.0033 (0.022)	34,540	33,550	990 (820)
Bandwidth by Cross Validation $\hat{h}=(1/s)\sum h_{s\in[1,5]}=0.001$	0.623	0.597	0.0261 (0.018)	34,540	30,830	3,710 (1,090)

Note 1: Estimates of the impact refer to estimates of expression (7) in Section 4.

Note 2: Each of the 5 sequence of information leads to a different estimate of the bandwidth. Such estimate is the one used to find the appropriate matching group. The bandwidths expressed in Table 11 refers to the average over the 5 bandwidths for each the normal and cross validation.

7 Conclusions

In this paper we have estimated the impact of disability status on labour market participation and labour earnings, on disable persons.

Our empirical section makes use of matching methods to allow for the counterfactual approach associated with treatment effect techniques for program evaluation. In particular, we estimate by matching on the propensity score. Such method improves on other parametric and semi-parametric approaches to program evaluation because it avoids potential bias due to model misspecification. At the same time, matching on the propensity score allows to compare the outcome of sub-groups in the same support as defined by a set of observed characteristics, thus accounting for the common support problem.

Our empirical study draws data from seven waves of the German Socio Economic Panel (GSOEP, 1994-2000), thus using all years for which the panel provides identical health information for all regions of Unified Germany, in particular, information on disability status as legally defined by German laws. The use of several waves allows us to construct two groups of individuals defined as treatment individuals and control (or comparison) individuals. Those in the treatment group are individuals who been non-disable at a particular year, become disable and remain so in the consecutive second and third year. According to German law, from the moment a person becomes (legally) disable she is entitled to advantages (e.g., particular re-training, free rehabilitation, subsided wages for employers, etc.) which should help her to lower the cost of engaging in paid labour market activities. Thus, we assume this policies to have some impact on the observed labour market outcome of treatment units, given that they have been disable for at most two years. The control group are individual who declare non-disability status over a given set of three years, and, therefore, do not receive the impact of policies which are built specifically for disable persons.

We estimate the propensity score using variables grouped according to categories of observed characteristics which may have an effect on both labour market outcomes and the probability of becoming disable. These categories are occupational sector, objective health measures, genetics and/or parental background, demographics, leisure activities and degree of urbanization. We control for these variables at a determined pre-disability period, thus avoiding any possible endogeneity problems between the outcome of interest an the conditioning set. Comparing estimates of the propensity score for both treatment and control groups shows that observed variables solve the common support problem; there seems to be no sorting with respect to any particular variable and, therefore, matching provides an adequate set of tools to identify the treatment effect on the treated.

Matching on the propensity score can lead to different estimates according to which matching method is used. Our empirical section compares a variety of estimates using two different matching techniques, namely, nearest neighbour with various callipers, and kernel estimation (both Gaussian and Epanechnikov) with both a normal bandwidth and a bandwidth based on cross validation. The magnitude of the estimates does not differ significantly between methods. However, Gaussian kernels, specially those estimates based on the normal-approximation bandwidth, suggest a possible bias relative to nearest neighbour estimation. This is expected since nearest neighbour is based on drawing one observation from the control group to serve as counterfactual, whereas the counterfactual kernel estimates is a weighted sum of many, if not all, observations in the comparison group.

With respect to the outcome “labour market participation”, our results show an almost negligent difference in the participation rate between disable and non-disable, with non-disable persons showing a 0.1 to 4.3 higher participation rate than disable individuals. In all matching methods considered, such differences are statistically insignificant. A similar story emerges with reference to the outcome labour earnings. Estimates suggest that disable earn somehow slightly more than non-disable, with at most DM. 3,070 per year difference. However, as with participation rates, the difference is statistically insignificant.

Besides our contribution to the growing area in applied econometrics that uses treatment effect techniques for program evaluation, we believe that this paper makes an important contribution to the understanding of how disable individuals fair in the labour market. Many studies suggest an adverse participation and wage differential for disable individuals who are able and willing to participate in paid labour market activities. For example, in the case of Germany, Albrecht and Braun (1998), suggest large differentials in unemployment rate, whereas in the UK Kidd et al. (2001) finds a 50% gap in the participation rate, where only half of it can be explained by productivity differentials. Such studies employ methods which might lead to both bias due to model misspecification and the failure of the common support. Our method is flexible, intuitive and it avoids bias associated with alternative estimation techniques. In our study, we find no significant difference in terms of labour market outcomes between non-disable and disable groups. Because of our particular way of constructing the treatment and control units, a direct implication of our results is that policies which aim at helping disable individuals into paid labour market activities are effective at achieving their aim. It is often thought that the quota system in Germany is wrongly defined. Such system obliges employers with a workforce of 16 or more to employ disable employees at least up to a minimum of 6% of the total workforce. The un-fulfilment of such quotas are often named to suggest that disable fair worst in the labour market, with less chance of been unemployed than non-disable. Nevertheless, it is often the case that the number of disable seeking vacancies is far below the number of legally allocated vacancies by the quota system. Our results show evidence to suggest that disable and non-disable are very similar with respect to labour market outcomes, thus providing some empirical evidence to back up the argument that the outcome of the quota system is not a good indicator to judge the labour market outcome of disable individuals.

References

Albrecht, M. and H. Braun (1998), "International research project on job retention and return to work strategies for disable workers. Study Report: Germany, ed. *International Labour Organization*, Geneva.

Black, D. and J. Smith (2002), "How Robust is the evidence on the effects of college quality? Evidence from matching", *University of Maryland working paper*.

Even, W.E., and D.A. MacPherson (1990), "Plant size in the decline of unionism", *Economic letters*, 32, 392-398.

Haerdle, W. and J.S. Marron (1985), "Optimal bandwidth selection in nonparametric regression functions estimation", *Annals of Statistics*, 13, 1465-1481.

Heckman, J. and J. Hotz (1989), "Choosing among alternative methods for estimating the impact of social programs: The case of manpower training", *Journal of the American Statistical Association*, 84, 862-874.

Heckman, J., H. Ichimura, and P. Todd (1997), "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme", *Review of Economics Studies*, 64, 605-654.

Kidd, M., P. Sloane and I. Ferko (2000), "Disability and labour market: an analysis of British males", *Journal of Health Economics*, 19, 961-981.

Kreider, B. and J. Pepper (2002), "Disability and employment: re-evaluating the evidence in light of reporting errors", *CRR WP 2002-6*, Boston College.

Lechner, M. (1995), "Continuous off-the job training in east Germany after Unification", *ZEW-discussion paper*, 95-27.

Lechner, M. (2001), ""Identification and estimation of causal effects of multiple treatments under the conditional independence assumption", *Econometric Evaluation of Labour Market Policies*, ed. Michael Lechner and Friedhelm Pfeiffer, 43-58. Heidelberg: Physica.

Manski, C., (1994), "The selection problem", in: *Advances in Econometrics*, ed. C. Sims, Cambridge University Press, 143-170.

Mueser, P., K. Troske and A. Gorislavsky (2003), "Using state administrative data to measure program performance", *Working paper*, University of Missouri-Columbia.

Oyen, R. (1989), “Vocational training, labour market opportunities, and internal company integration of disable persons. Survey of the empirical research of the 1980s”, *Mitteilungen aus der Arbeitsmrket – und Berufsforschung (MittAB),* ed. Stuttgart et al., 4, 507-522.

Rosenbaum, P., and D. Rubin (1983), “The central role of the propensity score in observational studies for causal effects”, *Biometrika*, 70(1), 41-55.

Rubin, D. (1974), “Estimating casual effects of treatments in randomised and non-randomised studies”, *Journal of Educational Psychology*, 66, 688-701.

Silverman, R. (1986), “Density estimation”, *London: Chapman and Hall*.

Smith, J. and P. Todd (2002), “Does matching overcome LaLonde’s critique of non-experimental methods?”, *Journal of Econometrics*, forthcoming.

Thornton, P. (1998), “International research project on job retention and return to work strategies for disabled workers. Key issues”. *International Labour Organization Publications*, Geneva.

Thornton, P. and N. Lunt (1997), “Employment policies for disabled people in eighteen countries: a review”, *York Social Policy Research Unit WP series*, University of York.

Williamson, J. and T. McNamara, “The effects of unplanned changes in marital and disability status: interrupted trajectories and labour force participation”, *CRR WP 2002-6*, Boston College.

Appendix A: Survey questions on disability

In 1994, the GSOEP questionnaire became identical for all surveyed household in Unified Germany, and for the first time all questions in the survey were equal for both East and West Germany. The health section of the questionnaire includes two questions to elicit information on the legal position of individuals with respect to disability status, and, if any, the extend of disability in the form of a "percentage of disability". The two questions are reproduced below as given in the original questionnaire (English translation):

"Are you officially registered as having a reduced capacity for work or of being severely disabled? (If you are receiving disability benefits, then enter "yes.")"

yes _____

no _____

"If yes, what is the degree of your disability?"

percent of disability _____

The above refer to years 1994 to 1997. In 1998 and thereafter the wording of both questions changed slightly, although the intended information remained equivalent. The new wording is reproduce below:

"Are you legally classified as handicapped or capable of gainful employment only to a reduced extent due to medical reasons?(If you receive social security due to illness or disability, etc., please answer "yes".)

Yes ____

No _____

" What is the extent of this capability reduction or handicap according to the most recent diagnosis?"

____%

Appendix B: Algorithm for estimating the propensity score with 5 independent sequence of information.

We are interested on estimating the impact of disability on labour market outcomes at t_3 of individuals who been non-disable at t_1, become disable at t_2 and remain classified as disable at t_3. Our data is based on an annual survey, which means that we require individuals who are observed for at least for 3 consecutive years. For example, if the first wave is for the year 1994, we need to observe both the labour market outcome Y_{t3}, and the health pattern – non-disability or disability – of the sample consecutively for the years 1994, 1995 and 1996. However, from 1994 onwards the survey provides a total of 7 years of information (1994 to 2000), such that 5 different sequence of three years each can be used to estimate the parameter of interest. The following shows the algorithm to estimate the final ATET:

Step 1: Starting with the first three year sequence $S_j;\ j=1$, select individuals from the original N sample who are, first, observed consistently over the three year period, and secondly, have a health pattern either defined as AAA (or a for short) – the controls or comparison group – or ADD (or d for short) – the treated group. Disregard any other units in the sample. The sample n of controls and unit form a mutually exclusive binary outcome, with $n = n_a + n_d$. Our assumption is that the original N sample is a representative sample of the target population.

Step 2: Select any variable in the information set at t_1 within the survey that might be thought to have an effect on both the treatment and outcome of interest. Let the (k,n) matrix $X_1 = [X'_{a,1} : X'_{d,1}]$ identify these variables allowing for any properly justified interaction between them. With this, estimate the propensity scores $p_j(x_{a,1}) = \Phi_j(x'_{a,1}\hat{b}_n)$ and $p_j(x_{d,1}) = \Phi_j(x'_{d,1}\hat{b}_n)$ for comparison and treatment groups, respectively, where Φ stands for the cumulative normal distribution, and $\hat{b}_n$ is the parameter estimate of a binary model (e.g., Probit) such that $P(ADD=1 \mid X_1 = x_1; b)$.

Step 3: With an appropriate matching method, compare the distance of the ith element of the estimated propensity score vector for treated units $p_j(x_{d,1})$, to all elements in $p_j(x_{a,1})$, the estimated propensity score vector for the comparison units. The ith element in the treated group receives the counterfactual outcome Y_i^c, where Y_i^c is the labour market outcome belonging to the comparison unit that minimises $\left|p_j^i(x_{d,1}) - p_j(x_{a,1})\right|$. Repeat the process for each i unit in n_d to end up with a vector of counterfactual outcomes $y_j^c = (y_1^c,..., y_{nd}^c)'$.

Step 4: Repeat step 1 to step 3 for each of the available 3 year sequences. In our case we end up with 5 vectors of counterfactuals $y_1^c, y_2^c,..., y_5^c$, one for each of our constructed 3 year sequence.

Step 5: Estimate the expected value of the counterfactual outcome with the sample average such that $\hat{E}[Y^a \mid ADD] = (1/n_d)\sum_{j=1}^{5}\sum_{l} y_{l\in j}^c$, where l is the number of treatment units in sequence j. Do the same with respect to the expected value of the actual outcome for the treated units, such that $\hat{E}[Y^d \mid ADD] = (1/n_d)\sum_{j=1}^{5}\sum_{l} y_{l\in j}^d$. The average treatment effect on the treated, or ATET, is given by $\hat{E}[Y^a \mid ADD] - \hat{E}[Y^d \mid ADD]$.

To estimate standard errors for ATET, repeat steps 1 to 5 an appropriate number of times (for example, 500), each time re-sampling with replacement from the original N in the survey. This process will give a vector of ATET estimates, $(ATET_1,..., ATET_{500})$. The standard error for the actual estimate ATET is obtained by estimating the standard error of $(ATET_1,..., ATET_{500})$.

Appendix C: Complementary tables with summary statistics for Section 3.

Table C1 shows the analogue to Table 4, Section 3, for attrition units, with summary statistics based on pre-attrition year. For example, the row 1994 contains no information because this is the first year we consider in our empirical analysis. There are 133 individuals observed in 1994 who are no longer observed in 1995: for this row (1995) any summary statistics are based on 1994 information.

Table C1: Non-participation, disability and interaction between health and labour market status for attrition units.

Year	Total number of observations per year	Share of disable in %	Share of non-participants in %		
	Full sample	Full sample	Full sample	Disable sample	Non-disable sample
1994					
1995	133	3.0 (0.2)	18.1 (0.4)	-	17.2 (4.1)
1996	670	4.1 (0.1)	25.0 (0.2)	46.6 (7.1)	24.1 (3.7)
1997	786	4.6 (0.01)	27.7 (0.3)	39.1 (5.1)	27.4 (2.8)
1998	825	3.7 (0.01)	23.1 (0.2)	64.5 (4.0)	21.5 (4.0)
1999	767	5.2 (0.01)	24.8 (0.2)	68.7 (5.5)	22.4 (5.2)
2000	615	6.3 (0.01)	20.4 (0.2)	66.3 (6.2)	17.3 (5.0)

Note: Not-employed are those defined as either officially registered unemployed, those declaring full time housework and early retired.

Table C2 shows the analogue to Table 8, but with the distribution of degree of disability for those who are not considered as "severely" disable, that is, for non-disable individuals who have been assessed a degree of disability between [0,30). Very few individuals who comprise the control sample show some degree of disability.

Table C2: Distribution of degree of disabilities for non-disable

Degree of Disability	**Distribution between degree brackets at t_1**	**Distribution between degree brackets at t_3**
0 degree	99%	99%
0 < degree < 5	0.0053%	0.0052%
5 # degree < 10	0.011	0.016%
10 # degree < 15	0.031%	0.032%
15 # degree < 20	0.016%	0.042%
20 # degree < 30	0.56%	0.75%

Note 1: Estimates based on individuals forming the comparison group.

Note 2: Numbers expressed up to 4 significant figures.

Table C3: Summary statistics for individuals in the treatment group with a degree of disability of 90 or 100%. These 7 individuals show characteristics not dissimilar to individuals in the general population.

Table C3: Summary characteristics for 7 individual with extreme disability.

Individuals persnr	t_3	**degree**	**Age/ gender**	**Employed (E), or Unemployed (U)**	**Years of education**	**1-Digit ISCO category**	**Labour income**	**Hours work per year**
474003	1996	100	33 (f)	E	9	7	43,058	1800
169001	1997	90	47 (m)	E	11	7	95,972	2100
292902	1997	90	60 (f)	E	11	7	41,616	2200
7008901	1997	100	39 (m)	E	12	7	41,398	2300
79905	1997	100	36 (m)	E	9	7	50,009	2100
322901	1998	100	55 (m)	E	11	7	51,657	2100
278101	1999	90	51 (m)	E	11	4	71,100	2200

Note: Income based on annual labour income in Deutsch Marks with Base 1999. Annual hours worked are rounded to the nearest hundred.

Appendix D: Description of conditioning variables and summary statistics of variables by health status for each of the waves considered.

Description of variables used in the Conditioning set X_1

Variable	Description	Category
Satih	Satisfaction with health. Categorical, with 1=lowest and 10=highest	Genetics
Satiw	Satisfaction with work. Categorical, with 1=lowest and 10=highest	Labour
Sport	Practice sport regularly =1, 0 otherwise	Leisure
Hunt	Shows hunting for job behaviour in the past three months=1, 0 otherwise	Labour
Blow	Blue collar worker, unskilled or semi-skilled=1, 0 otherwise	Labour
Bmed	Blue collar worker, skilled=1, 0 otherwise	Labour
Bhigh	Blue collar worker, highly skilled=1, 0 otherwise.	Labour
farmer	Farmer and self-employed =1, 0 otherwise	Labour
Self	Other type of self-employed=1, 0 otherwise	Labour
Wlow	White collar worker, low skilled=1, 0 otherwise	Labour
Wmed	White collar worker, clerical and medium skilled=1, 0 otherwise	Labour
whigh	White collar worker, professional and managerial=1, 0 otherwise.	Labour
Clow	Civil servant, low skilled=1, 0 otherwise	Labour
Cmed	Civil servant, medium skilled =1, 0 otherwise	Labour
Chigh	Civil servant, managerial and professional=1, 0 otherwise.	Labour
insure	Has private insurance for medical health=1, 0 otherwise	SES
avg.degree	Average degree of disability	Genetics
sick	Number of days sick per year (off work), average	Genetics
sickw	Visited doctor for work related injuries=1, 0 otherwise	Labour
aa	Average in the A-sample of the GSOEP (Original West Sample)	SES
bb	Average in the B-sample of the GSOEP (Working immigrants)	SES
cc	Average in the C-sample of the GSOEP (Original East Sample)	SES
dd	Average in the D-sample of the GSOEP (immigrants since 1984 in Germany)	SES
german	If of German origin=1, 0 otherwise	SES
europe	If of European Union origin=1, 0 otherwise	SES
world	If of any other place in the world =1, 0 otherwise	SES
edu	Number of years of education	SES
fsize	Number of individuals permanent living in the household	SES
kids	Number of dependent children	SES
age	Age of individual a time of survey	SES
hours	Average weekly hours at paid work	Labour
gender	Gender with 1=male, 0= female	Genetics
partner	If permanent partner=1, 0 otherwise	SES
Dadlow	Father's education is low (relative to medium – dadmed - or high - dadhigh)	Genetics
Motlow	Mother's education is low (relative to medium – dadmed - or high - dadhigh)	Genetics
own	If household owner=1, 0 otherwise	SES (Wealth)

Summary Statistics Covariates. Full weighted sample.

	1994	1995	1996	1997	1998	1999	2000
Size	8267	9084	8933	8674	8343	7782	7301
Satih	6.90 (2.11)	6.89 (2.11)	6.85 (2.10)	6.80 (2.07)	6.80 (2.06)	6.74 (2.09)	6.65 (2.14)
Satiw	5.24 (3.91)	5.31 (3.88)	5.28 (3.82)	5.20 (3.83)	5.12 (3.85)	5.13 (3.89)	4.96 (3.88)
Sport	25.91 (0.48)	32.37 (0.49)	27.29 (0.47)	27.49 (0.48)	32.10 (0.61)	30.16 (0.52)	29.35 (0.53)
Hunt	5.81 (0.26)	5.16 (0.23)	5.04 (0.23)	5.88 (0.25)	6.54 (0.27)	4.65 (0.24)	3.79 (0.22)
Blow	11.49 (0.35)	12.03 (0.34)	11.38 (0.34)	10.63 (0.33)	11.84 (0.35)	12.13 (0.37)	12.98 (0.39)
Bmed	14.13 (0.38)	12.88 (0.35)	13.74 (0.36)	12.70 (0.36)	12.35 (0.36)	12.54 (0.38)	11.77 (0.38)
Bhigh	0.69 (0.09)	0.67 (0.09)	0.59 (0.08)	0.58 (0.08)	0.56 (0.08)	0.71 (0.09)	1.14 (0.12)
farmer	0.37 (0.07)	0.42 (0.07)	0.37 (0.06)	0.34 (0.06)	0.33 (0.06)	0.36 (0.07)	0.36 (0.07)
Self	6.60 (0.27)	6.55 (0.26)	7.00 (0.27)	6.91 (0.27)	6.45 (0.27)	7.44 (0.30)	6.62 (0.29)
Wlow	3.69 (0.21)	3.83 (0.20)	3.79 (0.20)	4.12 (0.21)	3.55 (0.20)	3.72 (0.21)	3.58 (0.22)
Wmed	5.60 (0.25)	5.47 (0.24)	5.78 (0.25)	4.91 (0.23)	5.59 (0.25)	6.63 (0.28)	5.53 (0.27)
whigh	25.45 (0.48)	26.38 (0.46)	25.70 (0.46)	26.85 (0.48)	25.48 (0.48)	26.46 (0.50)	28.94 (0.53)
Clow	0.27 (0.06)	0.18 (0.04)	0.16 (0.04)	0.33 (0.06)	0.30 (0.06)	0.12 (0.04)	0.16 (0.05)
Cmed	2.45 (0.17)	2.03 (0.15)	2.00 (0.15)	1.79 (0.14)	1.72 (0.14)	1.60 (0.14)	1.55 (0.14)
Chigh	3.38 (0.20)	3.71 (0.20)	3.33 (0.19)	3.57 (0.20)	3.59 (0.20)	3.41 (0.21)	4.16 (0.23)
insure	10.61 (0.34)	10.62 (0.32)	11.87 (0.34)	10.29 (0.33)	11.44 (0.35)	8.47 (0.32)	8.7 (0.33)
avg.degree	0.68 (0.01)	0.75 (0.00)	1.34 (8.05)	1.57 (8.78)	1.82 (9.37)	2.15 (10.41)	2.32 (10.74)
sick	7.40 (20.44)	7.52 (20.64)	7.97 (21.03)	7.75 (22.8)	7.04 (19.20)	8.41 (24.49)	7.89 (25.51)
sickw	4.27 (0.22)	4.54 (0.22)	4.38 (0.22)	4.23 (0.22)	4.44 (0.23)	3.89 (0.22)	n.a
aa	69.95 (0.50)	65.66 (0.50)	65.92 (0.50)	66.05 (0.51)	65.14 (0.52)	64.29 (0.54)	64.39 (0.56)
bb	8.69 (0.31)	7.36 (0.27)	7.41 (0.28)	7.95 (0.29)	8.79 (0.31)	9.21 (0.33)	9.32 (0.34)
cc	21.36 (0.45)	21.13 (0.43)	21.00 (0.43)	20.57 (0.43)	20.87 (0.44)	21.45 (0.47)	21.19 (0.48)
dd	0.00 (0.00)	5.85 (0.25)	5.66 (0.24)	5.43 (0.24)	5.20 (0.24)	5.04 (0.25)	5.10 (0.26)
german	89.32 (0.34)	86.05 (0.36)	86.37 (0.36)	86.35 (0.37)	86.50 (0.37)	86.41 (0.39)	86.35 (0.40)
europe	2.54 (0.17)	2.77 (0.17)	2.71 (0.17)	2.55 (0.17)	2.57 (0.17)	2.47 (0.18)	2.48 (0.18)
world	8.13 (0.30)	11.18 (0.33)	10.92 (0.33)	11.09 (0.34)	10.93 (0.34)	11.13 (0.36)	11.16 (0.37)
edu	11.85 (2.57)	5.10 (9.01)	11.93 (2.55)	11.94 (2.55)	11.91 (2.53)	11.88 (2.61)	12.35 (2.45)
fsize	2.95 (1.35)	2.91 (1.35)	2.87 (1.34)	2.86 (1.35)	2.85 (1.36)	2.86 (1.40)	2.84 (1.38)
kids	0.74 (0.00)	0.73 (0.00)	0.70 (0.00)	0.69 (0.00)	0.68 (0.01)	0.67 (0.01)	0.69 (0.01)
age	38.40 (11.15)	38.78 (11.07)	39.09 (11.28)	39.37 (11.21)	39.70 (11.26)	40.30 (11.00)	40.81 (10.56)
hours	26.66 (20.47)	26.88 (20.71)	26.04 (20.72)	26.88 (21.39)	25.77 (21.04)	26.47 (21.14)	27.43 (20.79)
gender	49.45 (0.55)	49.73 (0.52)	49.84 (0.53)	50.72 (0.54)	50.02 (0.55)	50.16 (0.57)	49.35 (0.59)
partner	56.72 (0.54)	58.00 (0.52)	61.28 (0.52)	60.55 (0.52)	59.52 (0.54)	60.10 (0.56)	61.92 (0.57)
dadlow	12.52 (0.36)	13.29 (0.36)	13.55 (0.36)	13.32 (0.36)	13.64 (0.38)	13.86 (0.39)	14.05 (0.41)
dadmed	71.11 (0.50)	77.70 (0.44)	65.57 (0.50)	65.84 (0.51)	66.02 (0.52)	64.72 (0.54)	64.54 (0.56)
dadhigh	16.37 (0.41)	9.00 (0.30)	20.88 (0.43)	20.84 (0.44)	20.33 (0.44)	21.42 (0.47)	21.41 (0.48)
mothlow	12.57 (0.36)	13.47 (0.36)	13.77 (0.36)	13.62 (0.37)	14.08 (0.38)	14.15 (0.40)	14.55 (0.41)
mothmed	81.44 (0.43)	84.53 (0.38)	80.01 (0.42)	79.98 (0.43)	79.38 (0.44)	79.22 (0.46)	78.87 (0.48)
mothhigh	6.00 (0.26)	2.00 (0.15)	6.22 (0.26)	6.41 (0.26)	6.54 (0.27)	6.63 (0.28)	6.57 (0.29)
own	37.57 (0.53)	37.39 (0.51)	37.94 (0.51)	39.69 (0.53)	39.65 (0.54)	55.69 (0.56)	53.51 (0.58)

Summary Statistics Covariates. Non-disability weighted sample.

	1994	1995	1996	1997	1998	1999	2000
N	8183	8928	8735	8444	8083	7494	6998
Satih	6.93 (2.10)	6.92 (2.10)	6.91 (2.06)	6.87 (2.02)	6.87 (2.02)	6.84 (2.02)	6.75 (2.08)
Satiw	5.26 (3.90)	5.32 (3.88)	5.34 (3.79)	5.28 (3.79)	5.20 (3.81)	5.24 (3.83)	5.06 (3.82)
Sport	26.03 (0.49)	32.49 (0.50)	27.43 (0.48)	28.00 (0.49)	33.07 (0.52)	30.54 (0.53)	29.83 (0.55)
Hunt	5.80 (0.26)	5.20 (0.23)	4.99 (0.23)	5.87 (0.26)	6.48 (0.27)	4.63 (0.24)	3.86 (0.23)
Blow	11.42 (0.35)	11.86 (0.34)	11.25 (0.34)	10.52 (0.33)	11.66 (0.36)	11.98 (0.38)	13.05 (0.40)
Bmed	14.19 (0.39)	12.87 (0.35)	13.58 (0.37)	12.57 (0.36)	12.31 (0.37)	12.66 (0.38)	11.80 (0.39)
Bhigh	0.65 (0.09)	0.67 (0.09)	0.60 (0.08)	0.59 (0.08)	0.59 (0.08)	0.74 (0.10)	1.14 (0.13)
Farmer	0.38 (0.07)	0.42 (0.07)	0.35 (0.06)	0.35 (0.06)	0.34 (0.07)	0.38 (0.07)	0.37 (0.07)
Self	6.65 (0.28)	6.63 (0.26)	7.14 (0.28)	7.06 (0.28)	6.57 (0.28)	7.65 (0.31)	6.82 (0.30)
Wlow	3.70 (0.21)	3.86 (0.20)	3.82 (0.21)	4.15 (0.22)	3.58 (0.21)	3.76 (0.22)	3.57 (0.22)
Wmed	5.61 (0.25)	5.50 (0.24)	5.79 (0.25)	5.01 (0.24)	5.67 (0.26)	6.75 (0.29)	5.62 (0.28)
Whigh	25.55 (0.48)	26.47 (0.47)	26.09 (0.47)	27.27 (0.48)	25.93 (0.49)	26.83 (0.51)	29.54 (0.55)
Clow	0.27 (0.06)	0.17 (0.04)	0.15 (0.04)	0.31 (0.06)	0.30 (0.06)	0.11 (0.04)	0.15 (0.05)
Cmed	2.48 (0.17)	2.07 (0.15)	2.04 (0.15)	1.82 (0.15)	1.78 (0.15)	1.66 (0.15)	1.58 (0.15)
Chigh	3.36 (0.20)	3.69 (0.20)	3.35 (0.19)	3.62 (0.20)	3.65 (0.21)	3.46 (0.21)	4.19 (0.24)
Insure	10.61 (0.34)	10.70 (0.33)	12.08 (0.35)	10.43 (0.33)	11.64 (0.36)	8.70 (0.33)	8.85 (0.34)
avg.degree	0.13 (0.00)	0.01 (0.00)	0.15 (0.00)	0.15 (0.00)	0.16 (0.00)	0.14 (0.000	0.14 (0.00)
Sick	7.23 (19.30)	7.21 (19.24)	7.44 (18.66)	6.95 (18.9)	6.61 (18.12)	7.63 (21.49)	6.94 (21.24)
Sickw	4.26 (0.22)	4.55 (0.22)	4.35 (0.22)	4.16 (0.22)	4.31 (0.23)	3.91 (0.22)	n.a
Aa	69.97 (0.51)	65.59 (0.50)	65.74 (0.51)	65.91 (0.52)	64.84 (0.53)	64.17 (0.55)	64.33 (0.57)
Bb	8.55 (0.31)	7.28 (0.27)	7.42 (0.28)	7.95 (0.29)	8.77 (0.31)	9.22 (0.33)	9.24 (0.35)
Cc	21.48 (0.45)	21.18 (0.43)	21.14 (0.44)	20.62 (0.44)	21.02 (0.45)	21.46 (0.47)	21.24 (0.49)
Dd	0.00 (0.00)	5.95 (0.25)	5.70 (0.25)	5.52 (0.25)	5.37 (0.25)	5.14 (0.26)	5.20 (0.27)
German	89.44 (0.34)	86.04 (0.37)	86.32 (0.37)	86.22 (0.38)	86.38 (0.38)	86.22 (0.40)	86.23 (0.41)
Europe	2.55 (0.17)	2.76 (0.17)	2.73 (0.17)	2.58 (0.17)	2.57 (0.18)	2.50 (0.18)	2.51 (0.19)
World	8.01 (0.30)	11.20 (0.33)	10.95 (0.33)	11.20 (0.34)	11.05 (0.35)	11.29 (0.37)	11.26 (0.380)
Edu	11.86 (2.57)	5.04 (8.96)	11.96 (2.56)	11.96 (2.56)	11.94 (2.55)	11.91 (2.63)	12.40 (2.46)
Fsize	2.95 (1.35)	2.91 (1.35)	2.88 (1.34)	2.88 (1.35)	2.87 (1.36)	2.88 (1.40)	2.86 (1.38)
Kids	0.74 (0.00)	0.73 (0.00)	0.71 (0.00)	0.70 (0.00)	0.69 (0.01)	0.69 (0.01)	0.72 (0.01)
Age	38.32 (11.14)	38.62 (11.03)	38.80 (11.19)	39.02 (11.08)	39.32 (11.13)	39.84 (10.85)	40.36 (10.42)
Hours	26.70 (20.47)	26.88 (20.72)	26.11 (20.70)	27.03 (21.34)	26.04 (21.00)	26.73 (21.08)	27.84 (20.71)
Gender	49.26 (0.55)	49.70 (0.53)	49.59 (0.53)	50.44 (0.54)	49.70 (0.56)	49.70 (0.58)	48.79 (0.60)
Partner	56.71 (0.55)	57.88 (0.52)	60.98 (0.52)	60.28 (0.53)	59.29 (0.55)	59.89 (0.57)	61.83 (0.58)
Dadlow	12.46 (0.37)	13.29 (0.36)	13.62 (0.37)	13.43 (0.37)	13.59 (0.38)	13.86 (0.40)	14.03 (0.42)
Dadmed	71.07 (0.50)	77.69 (0.44)	65.25 (0.51)	65.47 (0.52)	65.81 (0.53)	64.41 (0.55)	64.22 (0.57)
Dadhigh	16.47 (0.41)	9.01 (0.30)	21.13 (0.44)	21.10 (0.44)	20.60 (0.45)	21.73 (0.48)	21.75 (0.49)
Mothlow	12.48 (0.37)	13.46 (0.36)	13.81 (0.37)	13.62 (0.37)	14.09 (0.39)	14.08 (0.40)	14.48 (0.42)
Mothmed	81.50 (0.43)	84.51 (0.38)	79.83 (0.43)	79.80 (0.44)	79.21 (0.45)	79.05 (0.47)	78.66 (0.49)
mothhigh	6.02 (0.26)	2.03 (0.15)	6.36 (0.26)	6.58 (0.27)	6.70 (0.28)	6.87 (0.29)	6.86 (0.30)
Own	37.57 (0.54)	37.58 (0.51)	38.20 (0.52)	39.81 (0.53)	39.75 (0.54)	55.32 (0.57	53.14 (0.60

Summary Statistics Covariates. Disability weighted sample.

	1994	1995	1996	1997	1998	1999	2000
N	84	156	198	230	260	288	303
satih	5.02 (2.49)	4.99 (2.01)	4.48 (2.44)	4.40 (2.05)	5.01 (2.31)	4.54 (2.36)	4.49 (2.27)
satiw	3.02 (4.20)	4.31 (4.12)	3.09 (4.22)	2.70 (4.21)	2.77 (4.13)	2.58 (4.37)	2.90 (4.46)
sport	15.35 (3.93)	25.70 (3.50)	21.84 (2.94)	11.12 (2.07)	16.23 (2.29)	21.41 (2.42)	19.24 (2.26)
hunt	6.10 (2.61)	2.95 (1.36)	6.78 (1.79)	5.95 (1.56)	8.14 (1.70)	4.96 (1.28)	2.23 (0.85)
blow	17.35 (4.13)	21.51 (3.29)	16.53 (2.64)	14.22 (2.30)	16.79 (2.32)	15.55 (2.14)	11.55 (1.84)
bmed	9.01 (3.12)	13.32 (2.72)	19.81 (2.83)	16.87 (2.47)	13.43 (2.11)	9.77 (1.75)	11.15 (1.81)
bhigh	4.69 (2.31)	0.47 (0.55)	0.00 (0.00)	0.38 (0.40)	0.00 (0.00)	0.00 (0.00)	1.16 (0.61)
farmer	0.00 (0.00)	0.00 (0.00)	0.98 (0.70)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
self	2.30 (1.64)	1.74 (1.05)	1.57 (0.88)	1.76 (0.87)	2.99 (1.06)	2.55 (0.93)	2.29 (0.86)
wlow	2.42 (1.68)	2.66 (1.29)	2.77 (1.17)	3.38 (1.19)	2.89 (1.04)	2.95 (1.00)	3.85 (1.11)
wmed	3.90 (2.11)	3.98 (1.57)	5.31 (1.59)	1.66 (0.84)	3.40 (1.12)	3.74 (1.12)	3.53 (1.06)
whigh	16.52 (4.05)	21.61 (3.30)	10.96 (2.22)	13.02 (2.22)	13.11 (2.09)	18.15 (2.27)	16.32 (2.12)
clow	0.00 (0.00)	0.65 (0.65)	0.42 (0.46)	0.95 (0.64)	0.29 (0.33)	0.24 (0.29)	0.23 (0.28)
Cmed	0.00 (0.00)	0.00 (0.00)	0.59 (0.54)	0.66 (0.54)	0.27 (0.32)	0.22 (0.28)	0.86 (0.53)
Chigh	4.40 (2.24)	4.51 (1.66)	2.32 (1.07)	1.89 (0.90)	2.05 (0.88)	2.29 (0.88)	3.44 (1.05)
Insure	10.70 (3.37)	6.19 (1.93)	3.92 (1.38)	5.85 (1.55)	6.23 (1.50)	3.34 (1.06)	4.91 (1.24)
avg.degree	48.96 (20.49)	41.70 (15.84)	46.67 (17.54)	47.79 (16.65)	47.41 (15.87)	48.17 (17.64)	48.45 (16.30)
Sick	22.92 (64.84)	24.52 (56.80)	28.05 (60.15)	33.8 (70.9)	18.76 (36.52)	26.31 (58.41)	28.02 (66.74)
Sickw	4.75 (2.32)	3.93 (1.56)	5.59 (1.63)	6.51 (1.63)	8.15 (1.70)	3.29 (1.05)	n.a
Aa	68.05 (5.09)	69.38 (3.69)	72.77 (3.16)	70.65 (3.00)	73.32 (2.74)	67.19 (2.77)	65.64 (2.73)
Bb	21.33 (4.47)	11.72 (2.58)	7.22 (1.84)	8.11 (1.80)	9.52 (1.82)	8.99 (1.69)	11.03 (1.80)
Cc	10.62 (3.36)	18.28 (3.09)	15.88 (2.60)	18.86 (2.58)	16.59 (2.31)	21.17 (2.41)	20.21 (2.31)
Dd	0.00 (0.00)	0.62 (0.63)	4.13 (1.41)	2.38 (1.01)	0.57 (0.47)	2.65 (0.95)	3.12 (1.00)
German	78.45 (4.49)	86.91 (2.70)	88.44 (2.27)	90.66 (1.92)	89.68 (1.89)	90.68 (1.71)	88.85 (1.81)
Europe	2.31 (1.64)	3.26 (1.42)	1.62 (0.90)	1.70 (0.85)	2.51 (0.97)	1.79 (0.78)	2.02 (0.81)
World	19.24 (4.30)	9.84 (2.38)	9.94 (2.13)	7.64 (1.75)	7.81 (1.66)	7.53 (1.55)	9.13 (1.66)
Edu	10.89 (2.06)	8.09 (11.17)	10.87(2.01)	11.13 (1.84)	11.06 (1.92)	11.03 (2.06)	11.26 (1.69)
Fsize	3.11 (1.20)	2.63 (1.21)	2.59 (1.19)	2.48 (1.25)	2.44 (1.23)	2.34 (1.09)	2.38 (1.17)
Kids	0.69 (0.05)	0.43 (0.04)	0.38 (0.03)	0.36 (0.03)	0.32 (0.03)	0.26 (0.03)	0.23 (0.02)
Age	45.81 (9.34)	47.73 (9.54)	49.88 (9.22)	50.63 (9.12)	50.21 (9.78)	50.84 (9.05)	50.28 (8.92)
Hours	23.17 (19.97)	27.11 (20.44)	23.22 (21.27)	21.90 (22.49)	18.25 (20.85)	20.42 (21.55)	18.84 (20.59)
Gender	66.41 (5.15)	51.41 (4.00)	59.52 (3.49)	59.77 (3.23)	58.54 (3.06)	60.69 (2.88)	61.17 (2.80)
Partner	58.08 (5.38)	64.53 (3.83)	72.62 (3.17)	69.45 (3.04)	65.77 (2.94)	64.81 (2.81)	63.82 (2.76)
Dadlow	17.50 (4.15)	13.27 (2.72)	10.73 (2.20)	9.67 (1.95)	15.04 (2.22)	13.81 (2.03)	14.46 (2.02)
Dadmed	74.49 (4.76)	78.27 (3.30)	77.86 (2.95)	77.91 (2.74)	71.98 (2.79)	71.85 (2.65)	71.38 (2.60)
Dadhigh	8.01 (2.96)	8.46 (2.23)	11.41 (2.26)	12.42 (2.17)	12.99 (2.08)	14.34 (2.06)	14.16 (2.00)
Mothlow	19.83 (4.35)	13.89 (2.77)	12.28 (2.33)	13.53 (2.26)	14.05 (2.16)	15.66 (2.14)	16.20 (2.12)
Mothmed	76.35 (4.64)	85.66 (2.81)	86.70 (2.41)	85.57 (2.32)	83.92 (2.28)	83.06 (2.21)	83.27 (2.14)
mothhigh	3.82 (2.09)	0.46 (0.54)	1.02 (0.71)	0.90 (0.62)	2.03 (0.87)	1.28 (0.66)	0.54 (0.42)
Own	37.43 (5.28)	26.80 (3.55)	28.05 (3.19)	35.63 (3.16)	37.17 (3.00)	64.09 (2.83)	61.42 (2.80)

Appendix F: Description of variables used in the construction of parental level of education.

The GSOEP survey elicits information on parental education and professional level for both the mother and father of surveyed individuals. There are two basic questions identical for each mother and father. The first question identifies vocational training and the second the educational achievement. Because of incomplete information in both questions, only a mixture of information in both allows us to build a categorical variable for the level of education of each parent.

Based on Tables F1 and F2, we build a variable with 4 categories, with individual with the lowest level of education or training are assigned a value 1, and individuals with the highest level of education/ training are assigned a value 4. This employs the following procedure:

1. *If the parent's (mother/father, independently) value in F1 or F2 do not match any of the three following possibilities:*

2. *If the parent's (mother/father, independently) value in F1 is [1,2] but F2 shows [2,3]*

3. *If the parent's (mother/father, independently) value in F1 is [2,3], [5,6] or [13], or if F1 is [1,2], but F2 shows [4,6].*

4. *If the parent (mother/father, independently) value in F1 is in [7,13] or [4]*

Table F1: Parental classification of training and technical experience.

Vocational Training Classification	Assigned Category
No knowledge of any vocational classification achieved	1
No vocational degree achieved	2
Trained in foreign company	3
Trained extensively in foreign company	4
Trained in foreign vocational school	5
Apprentice on trade with farming industry	6
Apprentice with business	7
Health care school training	8
Specialised technical college	9
Civil servant training	10
Technical engineering school	11
Foreign college	12
College or university	13
Other training	14

Table F2: Parental classification of educational achievement.

Education Classification	Assigned Category
No knowledge of what achievement	1
Secondary education degree	2
Intermediate school	3
Technical school	4
Upper secondary school	5
Other degree	6
No school degree achieved	7
Never attended school	8