Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

www.diw.de

SOEP The German Socio-Economic Panel Study

SOEPpapers

on Multidisciplinary Panel Data Research

172

Nicolas R. Ziebarth

Long-Term Absenteeism and Moral Hazard–Evidence from a Natural Experiment

Berlin, April 2009

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

Long-Term Absenteeism and Moral Hazard–
Evidence from a Natural Experiment[‡]

Nicolas R. Ziebarth
SOEP at DIW Berlin and TU Berlin[*]

April 16, 2009

[‡]I am grateful to Daniela Andrén, Eberhard Feess, Joachim R. Frick, John P. Haisken DeNew, Guido Imbens, Per Johansson, Martin Karlsson, Sonja Kassenböhmer, Henning Lohmann, Hendrik Schmitz, Per Pettersson-Lidbom, Tom Siedler, Gert G. Wagner, the participants in seminars at the 2nd IZA/IFAU Conference on Labour Market Policy Evaluation, the 20th Annual Conference of the European Association of Labour Economists, the 13th Annual Meeting of the Latin American and Caribbean Economic Association, the 4th International Young Scholars SOEP Symposium, the SOEP Brown Bag, and the Berlin Network of Labour Market Researchers (BeNA) for their helpful comments and discussions. Special thanks goes to Deborah Bowen who did the proofreading. I am responsible for all remaining errors and shortcomings of the article.

[*]German Institute for Economic Research (DIW) Berlin, Socio-Economic Panel Study (SOEP), Graduate Center of Economic and Social Research, Mohrenstraße 58, 10117 Berlin, Germany, and Berlin University of Technology (TUB) `e-mail: nziebarth@diw.de`

Abstract

Sick leave payments represent a significant portion of public health expenditures and labor costs. Reductions in replacement levels are a commonly used instrument to tackle moral hazard and to increase the efficiency of the health insurance market. In Germany's Statutory Health Insurance (SHI) system, the replacement level for periods of sickness of up to six weeks was reduced from 100 percent to 80 percent of an employee's gross wage at the end of 1996. At the same time, the replacement level for individuals absent for a long-term period, i.e., from the seventh week onwards, was reduced from 80 to 70 percent. We show theoretically that the net reform effects on long-term absenteeism can be disentangled into a direct and an indirect effect. Using SOEP data, a natural control group, and two different treatment groups, we estimate the net and the direct effect on the incidence and duration of long-term absenteeism by difference-in-differences. Our findings suggest that, on population average, the reforms have not affected long-term absenteeism significantly, which is in accordance with our theoretical predictions, assuming that employees on long-term sick leave are seriously sick. However, we find some heterogeneity in the effects and a small but significant decrease in the duration of long-term absenteeism for the poor and middle-aged full-time employed persons. All in all, moral hazard and presenteeism seem to be less of an issue in the right tail of the sickness spell distribution. Finally, our calculations suggest that from 1997 to 2006, around five billion euros were redistributed from persons on long-term sick leave to the SHI insurance pool.

Keywords: long-term absenteeism, sick pay, moral hazard, natural experiment, SOEP

JEL classification: I18, J22

1 Introduction

The average number of sickness absence days per year and employee varies between 5 and 29 among the OCED countries (OECD, 2006). Average absence days are to a large degree determined by long-term absence spells. In Germany, which lies in the middle field of the ranking with 15 days, absence spells of more than six weeks account for 40 percent of all absence days although they only represent 4 percent of all sickness cases (Badura et al., 2008).

Sick leave payments play a central role in determining public health expenditures and labor costs. Depending on the legislative framework, which differs widely from one country to the next, either the employer or the health insurance provider compensates employees for foregone earnings. What is referred to in Europe as "sickness absence insurance" is called "temporary disability insurance" in the US; in both cases, however, it provides compensation for wages losses due to temporary *non*-work-related illnesses or injuries. Yet the literature reveals a surprising gap of research on such insurance programs and thus a relative paucity of findings – particularly compared to the vast literature on unemployment benefits and unemployment duration. The importance of sickness insurance programs can be seen on the example of the US, where five states have such programs, among them the most populous state of California. There, in 2005, the total sum of net benefits for temporary disability insurance amounted to \$4.2 billion, while the total sum for unemployment insurance amounted to \$4.6 billion (Social Security Administration, 2006, 2008).

A common problem in insurance markets is moral hazard, which drives up insurance costs and leads to an inefficient allocation of resources. With sick leave, moral hazard exists if insured employees call in sick despite being able to work. Consequently, full compensation of foregone earnings is seldom provided either by private or by public health insurance systems.

This study exploits a natural experiment that occurred in Germany at the end of 1996. At that time, compensation payments for long-term absentees with sickness spells of more than six weeks reached the amount of 9.3 billion euros, comprising 7.3 percent of all expenditures by the German Statutory Health Insurance (SHI) system. Employers – who are legally obligated to pay employees for the first six weeks of sick leave – were forced to shoulder a burden of 28.2 billion euros (German Federal Statistical Office, 1998). In reaction, two health reforms were implemented, both of which cut the level of paid sick leave. The main aim of this paper is to analyze how these reforms affected work absence spells of *more* than six weeks, and to what extent moral hazard or presenteeism played a role in that part of the sickness spell distribution. Additionally, we calculate the SHI reform savings and redistributional effects.

At the end of 1996, sick leave compensation for the first six weeks was reduced from 100 percent to 80 percent of foregone gross wages. The second reform came into force at the beginning of 1997 and reduced the compensation level from 80 to 70 percent from the seventh week onwards.[1] Both reforms generate exogenous sources of variation and yield testable implications. We analyze the causal effects of the two health reforms on long-term absenteeism.

To theoretically predict the effects of both reforms on long-term absenteeism, we employ a simple dynamic model of absence behavior. First, if moral hazard plays a role and employees on long-term sick leave react to economic incentives, long-term absenteeism should decrease as the direct costs of long-term absenteeism unambiguously increase. Second, the costs of long-term absences decrease relative to the costs of short-term absences. This indirect effect would theoretically impact long-term absenteeism in a positive way. However, under the assumption that employees on long-term sick leave are indeed severely ill, the incentive structure of the sick pay scheme would break down and individuals would not adapt their labor supply to moderate cuts in sick pay.

Since Germany has two independent health care systems existing side by side, we are able to define subsamples that were affected by none, one, or both of the reforms. Thus, using data from the German Socio-Economic Panel Study (SOEP) and difference-in-differences methods, we can directly estimate the net effect and the direct effect of the two reforms on the incidence and duration of long-term absence spells. Since the legislator also decreased the upper limit of long-term sick pay from 100 percent to 90 percent of monthly net wages, the treatment intensity is likewise exogenously varied. Hence, we are not only able to define treatment and control groups but also to analyze the reform effects by treatment intensity in relation to gross wages.

We are confident for several reasons that our study measures causal reform effects. First, the control and treatment groups are legally clearly defined by political decisions, and the effect of the reforms on the individual was unambiguously exogenous. Second, the legal regulations do not allow selection into or out of the treatment group. Moreover, we control for many socioeconomic characteristics as well as health status, which is by far the most important determinant of long-term absenteeism. Third, due to the panel data format, the composition of the labor force can be considered.

[1] Henceforth, sickness spells that last less than six weeks are defined as short-term absenteeism and sickness spells that last longer than six weeks are defined as long-term absenteeism.

Finally, the reform effects on the incidence and length of long-term absenteeism are taken into account, and differentiation by treatment intensity is possible.

Our results indicate that on average, the cut in replacement levels did not produce an effect on the incidence and duration of long-term sickness spells, either directly or indirectly. This result is in line with our model predictions if we assume that employees on long-term sick leave are indeed seriously ill. However, we find evidence of heterogeneity in the effects. For the poor and middle-aged persons employed full-time, the duration of long-term absenteeism decreased significantly, although this decrease was of small magnitude. In contrast to the previous literature, these findings suggest that work absence behavior of more than thirty days is not very responsive to economic incentives, which implies that moral hazard is of little importance in this context. We calculate that the SHI saved around 5.5 billion euros due to the cut in long-term sick pay from 1997 to 2006. Five billion thereof were redistributed from the long-term sick to the insurance pool in order to achieve lower contribution rates.

The remainder of this paper is organized as follows. The next section provides more background information on the sickness absence literature. Section 3 explains the institutional features of the German health care sector, outlines the two health reforms, and describes the subsamples affected by the health reforms. In Section 4, we derive a theoretical explanation of how both reforms affected long-term absenteeism, expressed as a simple dynamic model of absence behavior. Section 5 describes the data used and how our variables were generated, whereas our estimation and identification strategy is detailed in Section 6. Section 7 presents our estimation results, which are discussed and summarized in Section 8.

2 Literature and Background

There is a large body of literature on absenteeism, but only a few studies explicitly analyze the role of sick leave regulations and the design of insurance contracts. Some studies have modeled the determinants of sick leave behavior theoretically and empirically (Jensen and McIntosh, 1999; Johansson and Palme, 1996), and others have shown how workplace conditions affect sickness absence (Dionne and Dostie, 2007; Ose, 2005). It is well documented that unemployment rates and absenteeism are negatively correlated. This is due partly to changes in the composition of the labor force; but behavioral factors seem to play a major role (Askildsen et al., 2005). It has also been found that workers take sick leave more frequently after the end of their probationary period, when full employment protection is provided (Lindbeck et al., 2006; Ichino and Riphahn, 2005; Engellandt and Riphahn, 2005; Riphahn, 2004). The theoretical paper by Chatterji and Tilley (2002) is one of the few that has explicitly discussed the role of presenteeism as a possible source of behavioral changes due to cuts in sick pay.

Only a handful of studies have empirically analyzed the relationship between absence behavior and compensation levels using only data from Sweden or the US. The US studies solely focus on the workers' compensation insurance which compensates for work-related illnesses or injuries. Curington (1994) used US data on claim records of "minor permanent partial impairments" and estimated the effects of several legislative changes in benefit levels on the length of work absences from 1964 through 1983. The results are mixed; some amendments induced changes in the work absence behavior, others did not. Another study from the US showed that increases in workers' compensation for "temporary total disabilities" due to work-related injuries led to an increase in injury duration in several states in the US in the 80s (Meyer et al., 1995). Johansson and Palme (2002) modeled the impact of a tax reform and a reduction in replacement levels in the Swedish health insurance system in 1991 on the hazard of work absences. They found that the increase in the absence costs reduced the incidence and length of sickness spells. Henrekson and Persson (2004) used long time series data for Sweden and took advantage of several legislative changes in the compensation levels to show that economic incentives strongly affect absence behavior. Pettersson-Lidbom and Skogman Thoursie (2008) showed that an increase in the benefit levels in Sweden in 1987 led to an increase in the incidence of absence spells. The study that comes closest to the one at hand was conducted by Johansson and Palme (2005), who took the health reform in Sweden in 1991 as an exogenous source of variation. They found that even for absence spells of more than 90(!) days, employees adapt their absence behavior to changes in replacement levels. To our knowledge, this was also the only study to date that has (indirectly) analyzed how long-term absenteeism is affected by reductions in replacement levels. However, all published Swedish studies lack a sound control group, which makes it difficult to disentangle the effects of the sick pay cut from overall economic trends like the deep recession in Sweden of 1991.

All in all, the existing literature suggests that sick people react to economic incentives as classical economic theory would predict. These behavioral reactions could be induced by moral hazard, where employees call in sick from work despite being healthy, or presenteeism, where employees go to work despite being sick.

3 The German Health Care System and the Policy Reforms

3.1 The Two Track German Health Care System

The German health care system actually consists of two independent health care systems existing side by side. The more important of the two is the Statutory Health Insurance (SHI) system, which covers about 90 percent of the German population. Employees whose income from salary is below a politically defined income threshold (2007: €3,975 per month) are compulsorily insured under the SHI. High-income earners who exceed that threshold, as well as the self-employed, have the right to choose between the SHI, a private health insurance (PHI) provider, or remaining uninsured. Non-working spouses and dependent children of individuals insured under the SHI are automatically insured by the SHI family insurance at no charge. Special groups such as students or unemployed are subject to special arrangements but are mostly insured under the SHI. In principle, insurance coverage is the same for all those insured under the SHI (German Ministry of Health, 2008).

The SHI is primarily financed by mandatory payroll deductions that are not risk-related. These contributions are paid equally by employer and employee up to a contribution ceiling (2007: €3,562.50 per month). Despite several health care reforms that tried to tackle the problem of rising health care expenditures, the contribution rates rose from 12.6 percent in 1990 to 13.9 percent in 2007, mainly due to demographic changes, medical progress, and system inefficiencies. The SHI is embedded in the German social legislation and is subject to the Social Code Book V (German Federal Statistical Office, 2008).

The second component of the German health care system is Private Health Insurance (PHI). It basically covers private-sector employees who earn above the income threshold, public sector employees, and self-employed persons.[2] Privately insured people pay risk-related insurance premiums based on a health check-up at the beginning of the insurance period. The premiums exceed the expected expenditures in younger age groups as the health insurer makes provisions for rising expenditures with increasing age. Coverage is provided under a variety of different health plans,

[2] We need to distinguish two types of employees in the German public sector. First, there are civil servants with tenure (called *Beamte*), henceforth called "civil servants." They are primarily PHI-insured since the state reimburses around 50 percent of their health expenditures (*Beihilfe*) and almost all of them insure the non-reimbursable expenditures privately. Second, we need to consider employees in the public sector without tenure (called *Angestellte im öffentlichen Dienst*). They have some privileges, too, but are mostly insured with the SHI (under the same conditions as everybody else). We call them "public servants."

and insurance contracts are subject to private law. Consequently, in Germany, public health care reforms affect the SHI rather than the PHI.

It is important to note that once an optionally insured person opts out of the SHI system, a return is virtually impossible. Employees above the income threshold are legally not allowed to switch back, and employees who fall below the income threshold in subsequent years may switch back but lose their provisions, which are not transferable (neither between PHI and SHI, nor between the different private health insurance providers). In reality, a change to a private health insurance provider can be regarded as a lifetime decision, and switching between the SHI and the PHI system as well as between private health insurance providers is very rare.

3.2 The German Sick Pay System

If an employee falls sick, a certificate from a physician is required from the third day of sick leave. The employer is legally obliged to pay sickness compensation up to six weeks per sickness spell regardless of the employee's health insurance. From the seventh week onwards, the physician needs to issue different certificates at reasonable time intervals of usually one week, and sick leave is paid by the SHI or the PHI. The replacement level for persons on long-term sick leave insured under the SHI is codified in the social legislation and is the same for all those with SHI insurance. In 1996, SHI payments for long-term absenteeism made up 7.3 percent of all SHI expenditures, which equaled 9.3 billion euros (German Federal Statistical Office, 1998).

The system for monitoring employees on sick leave is a potentially important determinant of the degree of moral hazard in the insurance market. In Germany, the "Medical Service of the SHI" exists for this purpose. One of the original objectives of the medical service was to monitor absenteeism. It is explicitly stated in the guidelines of this institution that long-term absenteeism in particular should be prevented in order to reduce the risk of patients descending the social ladder (Medizinischer Dienst der Krankenversicherung (MDK), 2008). The German social legislation stipulates that the SHI is obligated to call upon the Medical Service to provide an expert opinion, in order to dispel any doubts about work absences. Such doubts may arise if the insured person is absent unusually often or repeatedly sick for short-term periods on Mondays or Fridays. If physicians certify sickness uncommonly often, the SHI may ask for an expert opinion. The employer also has the right to call upon the Medical Service to provide an expert opinion. Expert opinions are based on available medical documents, information about the workplace, and a compulsory statement from the patient. If necessary, the medical service has the right to examine the patient physically and to cut benefits.[3] In 1997, about 2,000 full-time equivalent employees and independent physicians worked for the medical service and examined 1,719,386 cases of absenteeism (Medizinischer Dienst der Krankenversicherung (MDK), 2008).

3.3 The Policy Reforms

Two health reforms were implemented at the end of 1996. From October 1996 on, the replacement level during the first six weeks of sickness was reduced from 100 percent to 80 percent of foregone gross wages.[4] This reform had, at least theoretically, an indirect influence on sickness spells of more than six weeks and should therefore be considered. A second health reform act became effective on January 1, 1997. The replacement level from the seventh week onwards was cut from 80 percent to 70 percent of foregone earnings for those insured under the SHI.[5] Figure 1 illustrates the reduction in the replacement rates for short and long-term absence spells.

[Insert Figure 1 about here]

Sick leave payments for long-term absence spells are additionally limited by two benefit caps. First, if the wage of an employee insured under the SHI exceeds the legally defined contribution ceiling, then long-term sick pay is limited to 70 (80) percent of this contribution ceiling (2007: €0.7*3,562.50 per month) as contributions are capped over this ceiling as well. Second, before 1997, the replacement level was 80 percent of the gross wage if the total amount did not exceed 100 percent of the net wage. After 1997, the replacement level decreased to 70 percent of the gross wage and the benefit cap to 90 percent of the net wage. These upper limits introduce additional exogenous variation and allow us to generate an index that mirrors the cut in long-term sick pay on a continuous scale from zero percent of gross wages to 10 percent of gross wages.

To deter people from substituting several short-term absence spells for a single long-term absence spell, with the former compensated by a higher amount of sick

[3] The wordings of the law can be found in the Social Code Book V, article 275, para. 1, 1a; article 276, para. 5.

[4] Passed on September 15, 1996 this law is the *Arbeitsrechtliches Gesetz zur Förderung von Wachstum und Beschäftigung (Arbeitsrechtliches Beschäftigungsförderungsgesetz), BGBl. I 1996 p. 1476-1479.* It became effective at October 1, 1996. It should be noted that we are not able to precisely identify those employees who were affected by this law, as employers and unions voluntarily agreed in some collective wage agreements to continue the old sick pay scheme. However, as this reform is not the focus of this paper, this is of minor importance.

[5] Passed on November 1, 1996, this law is the *Gesetz zur Entlastung der Beiträge in der gesetzlichen Krankenversicherung (Beitragsentlastungsgesetz - BeitrEntlG), BGBl. I 1996 p. 1631-1633.*

pay in total, the law on employer-provided sick pay contains a specific passage.[6] The passage stipulates that if employees repeatedly have absence periods due to the same illness, they are no longer entitled to 100 percent employer-provided sick pay. Consequently, there is no incentive to substitute multiple short-term spells for a single long-term spell.

We now define subsamples that have been affected differently by the two health reforms, thereby serving as treatment and control groups in the evaluation of this natural experiment. As the sickness compensation for long-term absence is paid for by the health insurance and not by the employer, the second reform did not affect privately insured people, whose replacement levels are subject to individual insurance contracts.

[Insert Table 1 about here]

We can easily see from Table 1 that private-sector employees who were insured with the SHI (subsample 1) were affected by both reforms. In contrast, SHI-insured public-sector employees (subsample 2) were affected by the reduction in long-term sick pay but not by the cut in short-term sick pay due to political decisions. The same holds for SHI-insured trainees (subsample 3). The last two subsamples, PHI-insured public-sector employees and self-employed persons, were not affected by any of the reforms. As Table 2 visualizes, we accordingly defined two treatment groups and one control group.

[Insert Table 2 about here]

4 A Dynamic Model of Absence Behavior

In the following, we analyze the absence behavior of an individual i within a two-period model. We modify a model by Brown (1994) so as to be able to study the theoretical effects of the German health reforms on long-term absence behavior. The individual's utility function can be specified as:

$$u_t = (1 - \sigma_t)c_t + (\sigma_t)l_t, \qquad t = t, t+1;\ \sigma_t \in [0, 1] \tag{1}$$

[6] *Gesetz über die Zahlung des Arbeitsentgelts an Feiertagen und im Krankheitsfall (Entgeltfortzahlungsgesetz - EntgFG), BGBl. I 1994 p. 1014, 1065.* Para. 3 contains the passage.

where t is the time period, c_t represents consumption in period t, and l_t leisure in period t. The sickness level in t is specified by σ_t, where larger values of σ_t represent a higher degree of sickness. If the sickness index tends towards unity, i.e., a high level of sickness prevails, the individual draws utility only from leisure or recuperation time rather than consumption. On the other hand, if the sickness level is relatively low, the individual attaches more weight to consumption as opposed to leisure. To simplify the analysis, we assume that $f(\sigma_t)$ follows a uniform distribution:

$$f(\sigma_t) = \begin{cases} 1 \quad if \quad 0 \leq \sigma_t \leq 1 \\ 0 \qquad\quad otherwise \end{cases}$$

This means that each sickness level is equally probable. At time t, individuals are aware of their sickness level σ_t but concerning the subsequent period, only the probability distribution $f(\sigma_{t+1})$ is known.

To adequately model the German sick pay scheme, we define the replacement level during long-term sickness spells as r_l with $0 < r_l < 1$ and the replacement level during short term sickness spells as r_h with $0 < r_h < 1$. Moreover, $r_l < r_h < w$, where w represents the gross wage and is normalized to one. Sick pay is always provided when the individual is absent from work. Long-term sickness is when an individual is on sick leave for at least two continuous periods. Hence, in the first absence period after a working period, the sick pay is r_h, which is reduced to r_l in the second period. If a working period follows a long-term sickness period, the replacement level for the next sickness period is again r_h.

A key feature of this simple dynamic model is the concept of the reservation sickness level, σ_t^*, as introduced by Barmby et al. (1994). The reservation sickness level is defined as the value of σ_t such that an individual is indifferent between going to work and staying home. To be more precise, at σ_t^* the utility from working in period 1 plus the expected utility in period 2 equals the utility from being absent in period 1 plus the expected utility in period 2. As we are primarily interested in the reform effects on long-term absenteeism, we assume that our individual was on sick leave in $t-1$ and is eligible for sick pay in t with r_l as the replacement level. In t, the reservation level is hence implicitly defined by:

$$(1-\sigma_t^*)r_l + \sigma_t^* T + \frac{1}{1+\rho}E(U_{t+1}^{absent}) = (1-\sigma_t^*)w + \sigma_t^*(T-h) + \frac{1}{1+\rho}E(U_{t+1}^{work}) \qquad (2)$$

The left hand side of this equation represents the utility in period t if the individual continues to be on sick leave with sick leave compensation r_l and leisure T, where T is the total time available. The expected utility from period $t+1$ is added and discounted with the individual's time preference rate ρ. Analogously, the right hand side adds up the discounted utility in $t+1$ with the utility from working h hours and enjoying $T-h$ hours leisure in t.[7]

The individual decides whether to be absent from work by maximizing utility over both periods. If $\sigma_t > \sigma_t^*$, i.e., the actual sickness level exceeds the reservation sickness level, the individual stays away from work as more weight is placed on leisure rather than consumption. In other words, if employees are seriously sick, they value recuperation time far more than materialistic needs and go on sick leave. On the other hand, if $\sigma_t < \sigma_t^*$, individuals maximize their utility by working h hours.

One has to bear in mind that the decision to be absent from work or not has implications for the sick pay level in the next period. If individuals are absent from work in t, they get r_l in t as well as in $t+1$ if their sickness continues to be so severe that $\sigma_{t+1} > \sigma_{t+1}^{a*}$, where σ_{t+1}^{a*} is the reservation sickness level in $t+1$ conditional on having been absent in t. If they work in t and fall sick in $t+1$, with $\sigma_{t+1} > \sigma_{t+1}^{w*}$, their sick pay is r_h. Hence we can define $E(U_{t+1}^{absent})$ which is the expected utility in $t+1$ conditional on having been absent at time t:

$$
\begin{aligned}
&E(U_{t+1}^{absent}) \\
&= \quad (1-\sigma_{t+1}^{a*})\left[\left(1-E(\sigma_{t+1}\,|\sigma_{t+1}^{a*}<\sigma_{t+1}<1)\right)r_l + E(\sigma_{t+1}\,|\sigma_{t+1}^{a*}<\sigma_{t+1}<1)\,T\right] + \\
&\qquad \sigma_{t+1}^{a*}\left[\left(1-E(\sigma_{t+1}\,|0<\sigma_{t+1}<\sigma_{t+1}^{a*})\right)w + E(\sigma_{t+1}\,|0<\sigma_{t+1}<\sigma_{t+1}^{a*})\,(T-h)\right] \\
&= \quad (1-\sigma_{t+1}^{a*})\left[\left(1-\left(\frac{1+\sigma_{t+1}^{a*}}{2}\right)\right)r_l + \left(\frac{1+\sigma_{t+1}^{a*}}{2}\right)T\right] + \\
&\qquad \sigma_{t+1}^{a*}\left[\left(1-\left(\frac{\sigma_{t+1}^{a*}}{2}\right)\right)w + \left(\frac{\sigma_{t+1}^{a*}}{2}\right)(T-h)\right]
\end{aligned} \tag{3}
$$

As can be seen from (4), the expected utility in $t+1$ is expressed as the weighted average of the expected utility from attending work and being absent from work. The weights represent the probability that σ_{t+1} is less than the reservation sickness level and exceed the reservation sickness level, respectively. The expected values of consumption and leisure are evaluated by using the conditional probability distribution. Conditional on σ_{t+1} being between 0 and σ_{t+1}^{a*}, the expected value of σ_{t+1}, which is $\frac{\sigma_{t+1}^{a*}}{2}$ for the uniform distribution, is taken to evaluate the utility of a working employee. Analogously, the expected value of σ_{t+1}, conditional on being between σ_{t+1}^{a*} and 1, $\frac{1+\sigma_{t+1}^{a*}}{2}$, is substituted into the utility function for an absent employee.

Equivalently defined is $E(U_{t+1}^{work})$ which is the expected utility in $t+1$ conditional

[7] We assume a rigid employment contract without the possibility of working overtime or less than the contracted hours h.

on having worked in t:

$$E(U_{t+1}^{work}) = \sigma_{t+1}^{w*}\left[\left(1-\left(\frac{\sigma_{t+1}^{w*}}{2}\right)\right)w+\left(\frac{\sigma_{t+1}^{w*}}{2}\right)(T-h)\right]+ \\ (1-\sigma_{t+1}^{w*})\left[\left(1-\left(\frac{1+\sigma_{t+1}^{w*}}{2}\right)\right)r_h+\left(\frac{1+\sigma_{t+1}^{w*}}{2}\right)T\right] \tag{4}$$

Finally, we derive σ_{t+1}^{a*} and σ_{t+1}^{w*} as:

$$\sigma_{t+1}^{a*} = \frac{w-r_l}{w-r_l+h} \tag{5}$$

$$\sigma_{t+1}^{w*} = \frac{w-r_h}{w-r_h+h} \tag{6}$$

We find that $\frac{\partial\sigma_{t+1}^{a*}}{\partial r_l} < 0$ and $\frac{\partial\sigma_{t+1}^{w*}}{\partial r_h} < 0$, which means that a decrease in sick pay levels has a positive impact on the reservation sickness levels, resulting, ceteris paribus, in a lower probability to be absent from work. This is what we would expect intuitively when the costs of sickness rise. Moreover, static labor supply models also predict a decrease in absenteeism with decreasing sick pay rates (Brown and Sessions, 1996). Henceforth, we call this the direct effect of a reduction in sick pay.

As $r_l < r_h < w$, we get $\sigma_{t+1}^{a*} > \sigma_{t+1}^{w*}$ meaning that the probability to work in $t+1$ is higher for an employee who stayed home in t as opposed to an employee who worked in t. The reason is that the gap between wages and sick pay, i.e., the cost of absence, is bigger for long-term absenteeism as compared to a short-term absenteeism. This is a reasonable approximation of the statutory sick leave regulations in Germany.

Plugging equations (3) to (6) into (2) and solving for the reservation sickness level σ_t^* yields:

$$\sigma_t^* = \sigma_{t+1}^{a*} + \frac{\varpi}{(1+\rho)(w-r_l+h)} \tag{7}$$

$$\varpi = \frac{(r_h-r_l)h^2}{2(w-r_l+h)(w-r_h+h)} > 0 \tag{8}$$

We see that σ_t^* equals σ_{t+1}^{a*} plus a discounted positive term which we interpret as the impact of future absence costs on the today's decision to be absent from work or not. It illustrates how the German sick pay scheme, which penalizes long absence spells more severely than short absence spells, impacts the probability to stay at home in the current period. In the case of a flat sick pay level, which would not depend on the length of absence, the second term would vanish and the probability of being absent from work today would equal the probability of being absent from work tomorrow. Remember that this holds under the assumption that every health status is equally probable and outside the individual's influence. Utility-maximizing individuals need to take the impact of today's absence behavior on future sick pay entitlements into account.

We now predict how long-term absenteeism is affected if the sick pay levels for short and long absence spells decrease and the employee is entitled to r_l in case of being absent. Consider first the effects of a reduction in r_l.

$$\frac{\partial \sigma_t^*}{\partial r_l} = \underbrace{\frac{\partial \sigma_{t+1}^{a*}}{\partial r_l}}_{<0} + \underbrace{\frac{\frac{\partial \varpi}{\partial r_l}(w - r_l + h) + \varpi}{(1+\rho)(w - r_l + h)}}_{<0} \tag{9}$$

We see from equation (9) that the total effect of a decrease in r_l is the sum of the direct effect $\frac{\partial \sigma_{t+1}^{a*}}{\partial r_l}$ and an additional factor. Hence, it is crucial to consider the impact of the discounted future term when evaluating the impact of a reduction in r_l. The second term represents the indirect effect that arises from the gap in the replacement levels between long and short-term absence spells, $r_h - r_l$. In case of a flat compensation scheme the gap closes and the indirect effect disappears. Ceteris paribus, a reduction in r_l widens the compensation gap, increases future absence costs, and thus affects long-term absenteeism negatively, thereby strengthening the direct effect.

Now we consider a reduction in r_h. Note that there is no direct effect of a decrease in r_h for people in an ongoing long-term sickness spell. These people continue to get r_l if they remain absent, and get their full wage if they go back to work. However, a reduction in r_h would, ceteris paribus, diminish the compensation gap between short and long-term absences and thus exert a positive effect on long-term absenteeism.

$$\frac{\partial \sigma_t^*}{\partial r_h} = \underbrace{\frac{\partial \sigma_{t+1}^{a*}}{\partial r_h}}_{=0} + \underbrace{\frac{\frac{\partial \varpi}{\partial r_h}}{(1+\rho)(w - r_l + h)}}_{>0} \tag{10}$$

We now want to relax the rather restrictive assumption that the sickness level σ_t is independent of the sickness level in the previous period and that every sickness level is equally probable in every period. Suppose that the sickness levels are serially correlated and that r_h is paid for sickness spells up to six periods. If the employee continues to be on sick leave in the seventh period, r_l is paid. For a sickness spell to last more than six periods, the illness must to be so severe that $\sigma_t > \sigma_t^*$ in every period. If that is the case, the incentive structure of our sick leave scheme breaks down and the employee is absent from work in every period. Hence, if employees are seriously sick, which means that their degree of sickness tends towards unity, and the replacement levels change only moderately without taking on extreme values, then these employees do not react to economic incentives.

In Section 7, we empirically estimate the net effect as well as the direct effect of the German health reforms on long-term absenteeism.

5 Data and Variable Definitions

The dataset that we use is the German Socio-Economic Panel Study (SOEP). The SOEP is an annual representative household survey that was started in 1984 and sampled more than 20,000 persons in 2006. Further details can be found elsewhere (Wagner et al., 2007).

Depending on our empirical estimation strategy, we use data of the years 1994 to 1999. As our goal is to evaluate a reduction in wage compensation levels, we drop non-working respondents and those who are not eligible for long-term sickness compensation (i.e., people who earn less than 400 euros per month and working students). Furthermore, we drop observations with missings and restrict our sample to respondents aged 18 to 65.

5.1 Endogenous and Exogenous Variables

The SOEP contains various questions about the usage of health services and health insurance. We generate our first dependent dummy variable, which measures the incidence of long-term absenteeism, from the following question that was asked continuously from 1994 on: *"Were you sick from work for more than six weeks at one time last year?"* Since sick pay is reduced after six weeks because it is no longer disbursed by the employer but by the health insurance, and since a different certificate needs to be issued by the physician, measurement errors should play a minor role here.

To measure how many days long-term sick pay was received, we use the following SOEP question: *"How many days were you not able to work in 199X because of illness?"* We generate our second dependent variable by subtracting, for those who had a long-term absence spell, the number of employer-paid sick days – namely 30 –

from the total number of days absent.[8] [9] Clearly, this duration variable is subject to measurement errors as we assume that the respondents had no other absence spells. Moreover, comparing the average duration of long-term sick pay with official data, it becomes clear that we face a systematic underreporting in the survey data, as persons with long-term sickness spells are less likely to participate in the survey. However, if the cut in long-term sick pay did not affect the probability to participate in the survey and did not affect the sickness spell distribution, this duration measure is sufficient to evaluate the reform effects. While the former assumption clearly holds, one might argue that the latter is more problematic. Those who were only affected by the cut in long-term sick pay have an incentive to interrupt their long-term sickness spell and to start a new one. Luckily, we do not need to fear such behavioral effects since, according to German law, the claim for employer-provided sick pay expires in case of such sickness spell substitutions (see Section 3.3 for more details). Once more, the importance of having various treatment groups is emphasized here. By comparing Treatment Group 1 with our controls, we cannot identify potential reform effects, since a negative impact on the duration measure might be caused by the decrease in short-term sick pay. Contrasting Treatment Group 2, which was affected only by the cut in long-term sick pay, with the Control Group, and bearing in mind that sickness spell substitutions are not of relevance here, we can reliably estimate the impact of the reform on the length of long-term sickness spells.

As both questions on absenteeism refer to the last year, we take the information of time variant covariates from the previous year if the respondent was interviewed the year before. For respondents who were not interviewed in the previous year, we take the current information and assume that it did not change meanwhile.

The whole set of explanatory variables can be found in Appendix A and is categorized as follows. A first group incorporates variables on personal characteristics, like the dummies on *gender, immigrant, East Germany, partner, married, children, disabled, good health, bad health, no sports*, and *age* (age^2). The second group consists of educational controls such as the degree obtained, the number of years

[8] As noted above, public servants enjoy special privileges. The period in which their employer provides a 100 percent sickness compensation varies from 6 weeks to 26 weeks depending on seniority. Since we have detailed information about the seniority levels, we are able to identify privileged public servants and redefine for them long-term absence spells which eventually coincide with the period of the lower SHI sick pay. Hence, for public servants, we subtract the benefit days that are provided by the employer and that vary between 6 and 26 weeks.

[9] For those respondents who indicated having been absent for more than six weeks but who reported a total number of sick days of less than 30, we replaced the values with a one. By estimating a Zero-inflated Negbin-2 model (see Section 6.2) and predicting the benefit days, we imputed the values for respondents with a missing on the benefit day variable. We imputed the values only for respondents who indicated that they were on long-term sick leave and who had no missings on the other relevant variables.

with the company, and whether the person was trained for the job. The last group contains explanatory variables on job characteristics. Among them are *blue-collar worker, white-collar worker, the size of the company*, and the *monthly gross wage*.

5.2 Control Group, Treatment Groups, and Treatment Intensity Indices

As described in Section 3.3 and visualized in Table 2, we generate one control group and two treatment groups. For each of the treatment groups we compute a treatment index that represents the treatment intensity. By these means, we estimate the net and the direct reform effects.

The SOEP is very detailed about the insurance status and the workplace of the respondents, which allows us to precisely assign them to the control and treatment groups. However, self-employed persons insured under the SHI have the option to opt out of long-term sick pay in order to obtain lower contribution rates. As we are unable to identify respondents with such contracts, we drop them.

Another advantage of the SOEP is the extensive data about gross wages, net wages, and variable income components such as Christmas or vacation bonuses. The SOEP group deals precisely with the problem of missing income data, and imputes values thoroughly (Frick and Grabka, 2005). Thanks to this information and the legally defined upper limits for long-term sick pay, we are able to accurately generate treatment indices that display the decrease in replacement levels continuously from 0 to 10 percent of individual gross wages.

We firstly specify three treatment dummy variables. *Treatment Group 1* is a dummy variable that equals 1 if the respondent belongs to Treatment Group 1 and 0 if the respondent is in the Control Group. *Treatment Group 2* is a dummy variable that takes on the value 1 for respondents in Treatment Group 2 and 0 for respondents in the Control Group. Finally, *Treatment Group 3* has a 1 for people belonging to Treatment Group 2 and a 0 for people belonging to Treatment Group 1. In our basic specification, Treatment Group 1 contains 16,006 observations, Treatment Group 2 has 6,500 observations, and the Control Group contains 2,693 observations.

Beside the universal rule that long-term sick pay is 70 (80) percent of the gross wage up to the contribution ceiling, legally defined upper limits induce an additional, continuous, and more precise source of exogenous variation. The maximum amount of long-term sick pay was restricted to 100 percent of the net wage before the reform and to 90 percent of the net wage after the reform. Depending on the individual gross and net wages for those being treated, we can calculate the individual decrease in long-term sick pay in percent of the gross wage. Hence, the treatment intensity varies from 0 percent of the gross wage for those unaffected by the reform to a maximum decrease of 10 percent of the gross wage. We generate a continuous variable called *Treatment Index 1* that has the value 0 for those in the Control Group and values from 0.57 (percent) up to 10.00 (percent) for those in the *Treatment Group 1*. Equivalently built is *Treatment Index 2*, which includes people in the Control Group and *Treatment Group 2*. The density of both variables *Treatment Index 1* and *Treatment Index 2* peaks around 6 (percent) and 10 (percent). About 80 percent of the treated faced a cut in long-term sick pay between 4 and 8 percent, and about 12 percent experienced a cut of 10 percent.

6 Estimation Strategy and Identification

6.1 Probit Specification

To estimate the causal reform effects on the incidence of long-term absence spells, we fit a difference-in-differences (DiD) probit model of the following type:

$$Pr(y_{it} = 1) = \Phi(\alpha + \beta p97_t + \gamma D_{it} + \delta DiD_{it} + x'_{it}\zeta) \tag{11}$$

with i representing individuals and t representing time. The dummy $p97_t$ has a one for post-treatment years and a zero for pre-treatment years. The treatment dummy variable D_{it} indicates whether respondents belong to the treatment or the control group. Depending on the specifications, we use the variables *Treatment Group 1*, *Treatment Group 2*, *Treatment Group 3*, *Treatment Index 1*, or *Treatment Index 2*, respectively, which were explained in the previous subsection. The variable DiD_{it} can be interpreted as the interaction term between D_{it} and $p97_t$. Consequently, it always has the value zero for individuals in the control group as well as for the treated in pre-treatment periods. The vector x'_{it} includes all personal, educational, and job-related controls as well as year dummies, state dummies and a variable for the state unemployment rate in a given year. $\Phi(.)$ is the cumulative distribution function for the standard normal distribution.

The marginal effect of the interaction term DiD_{it} delivers us the causal reform effect and is henceforth always displayed when output tables are presented.[10]

[10] Puhani (2008) has shown that the advice of Ai and Norton (2004) to compute the discrete double difference $\frac{\Delta^2\Phi(.)}{\Delta \mathrm{p97}\Delta \mathrm{D}}$ is not of relevance in nonlinear models when the interest lies in the estimation of a treatment effect. Using treatment dummy variables, the average treatment effect

6.2 Count Data Specification

The second empirical specification intends to estimate how the policy reform affected the length of long-term absence spells in post-treatment periods. As the number of benefits days is a count with excess zero observations and overdispersion, i.e., the conditional variance exceeding the conditional mean, we fit count data models. Based on the Akaike (AIC) and Bayesian (BIC) information criteria as well as on Vuong tests, we found two model specifications to be appropriate.

The first is a *Hurdle-at-Zero Negative Binomial Model*, also simply referred to as a two-part model, which models two distinct statistical processes for the incidence and the duration of long-term absenteeism. The first part represents the probability of crossing the hurdle, e.g., of being absent long-term, and can be estimated by a logit or probit model equivalent to that in equation (11). The second part models the duration of long-term absenteeism by fitting a truncated at zero Negative Binomial-2 (NegBin-2) model (Deb and Trivedi, 1997).

The second count data model to be employed is the so-called *Zero-Inflated Negative Binominal Model* that equally allows diverging statistical processes for the incidence and duration of long-term absenteeism. The underlying statistical mechanism differentiates between employees on long-term sick leave and those not on long-term sick leave, and assigns different probabilities that are parameterized as functions of the covariates to each group. The binary process is again specified in form of a logit or a probit model, and the count process is now modeled as an untruncated NegBin-2 model for the binary process to take on value one. Hence, zero counts may be generated in two ways: as realizations of the binary process and as realizations of the count process when the binary process is one (Winkelmann, 2008).

Both count data models incorporate the negative binomial distribution. The reason is that, in contrast to the more restrictive Poisson distribution, it does not only take excess zeros into account but also allows for overdispersion and unobserved heterogeneity. The NegBin model can be seen as a special case of a continuous mixture model. In the notation of Cameron and Trivedi (2005), the NegBin distribution can be described as a density mixture of the following form:

on the treated at the time of the treatment is given by $\frac{\Delta\Phi(.)}{\Delta(\text{p97*D})} = \Phi(\alpha + \beta p97_t + \gamma D_{it} + \delta DiD_{it} + x'_{it}\zeta) - \Phi(\alpha + \beta p97_t + \gamma D_{it} + x'_{it}\zeta)$ which is exactly what we calculate and present throughout the paper.

$$
\begin{aligned}
\varphi(y|\mu,\alpha) &= \int f\,(y|\mu,\nu) \times \gamma(\nu|\alpha)\, d\nu \\
&= \int_0^\infty \left(\frac{e^{-exp(\boldsymbol{X\beta})\nu}\{exp(\boldsymbol{X\beta})\nu\}^y}{y!} \right) \left(\frac{\nu^{\delta-1}e^{-\nu\delta}\delta^\delta}{\Gamma(\delta)} \right) d\nu \\
&= \frac{\Gamma(\alpha^{-1}+y)}{\Gamma(\alpha^{-1})\Gamma(y+1)} \left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu+\alpha^{-1}} \right)^y
\end{aligned}
\tag{12}
$$

where $f\,(y|\mu,\nu)$ is the conditional Poisson distribution and $\gamma(\nu|\alpha)$ is assumed to be gamma-distributed with ν as an unobserved parameter with variance α. $\Gamma(.)$ denotes the gamma integral and $\mu = exp(\boldsymbol{X\beta})$ where the matrix $\boldsymbol{X}$ incorporates the same variables as the probit model in equation (11). The NegBin can be derived in different ways; it has different variants and different interpretations. Note that in the special case of $\alpha = 0$ the NegBin collapses to a simple Poisson model.

6.3 Identification

The core identification assumption in every DiD model is the common time assumption. It assumes that the estimated effect is due entirely to the policy intervention and that in the absence of the intervention and conditional on the covariates, the outcome variable of the treated group would have developed in the same way as the outcome variable of the controls. Depending on the context, this may be a more or less strong assumption. Our identification strategy is based on various pillars, making us confident that we have reliably identified causal reform effects.

First, we should point out that we use a distinct control group that was not affected by the reforms. Additionally, the identification of two different treatment groups that were affected by a single and both reforms, respectively, makes it possible to distinguish between direct and net reform effects. Since the insurance status of the respondents as well as their job characteristics and earnings are collected accurately, we can assign people very precisely to the control and treatment groups.

Second, we exploit an additional source of exogenous variation which allows us to distinguish effects by treatment intensity (see Section 3.3 for more details). By using income information that differentiates between gross wages, net wages, and fringe benefits, we are able to generate treatment intensity variables remarkably exactly. In the period under observation, the implementation of the reform and the variation in the treatment intensity were clearly exogenous to the individuals and politically determined. We have not found evidence that the policy change was endogenous in the sense that the reform was a reaction to increasing absence rates (Besley and

Case, 2000; German Federal Statistical Office, 2008). Rather, it was a fairly random means of cutting health expenditures and was used mainly as an instrument of the unpopular Kohl administration, which took office in 1982, to demonstrate strength and capacity to act.

Third, as in almost every study that builds upon natural experiments, the control group and the two treatment groups differ significantly with respect to most of the observed characteristics (see Table 3). For example, in comparison to the Control Group, Treatment Group 1 includes fewer females but more immigrants, and the employees are less educated. Treatment Group 2 is younger than the other subsamples, less often married, and includes more white-collar workers without tenure. The heterogeneity in most of the observable characteristics is due to the federal regulations of the German health insurance and hence unavoidable. However, we argue that it is very unlikely that the common time trend assumption is violated as a) the differences in characteristics are not the result of treatment-related self-selection but politically determined, b) we have a very rich dataset and are able include a variety of controls, c) the key determinant of long-term absenteeism is the health status, which we are able to control for. Recall that it poses no problem if the subsamples have different probabilities of being affected by long-term sickness; the identifying assumption would only be violated if unobservables existed that would impact the *change* of these probabilities differently. In case of long-term absenteeism it is difficult to think of unobservables that have a diverging effect on the dynamic of the outcome – all the more after having controlled for a rich set of health-related, personal, educational, and job-related covariates as well as the annual regional unemployment rate, regional time-invariant effects, and annual time trends.

[Insert Table 3 about here]

We can see from Table 4 that relatively few covariates affect long-term absenteeism significantly. More educated employees are less often absent for long-term periods, and firm size is positively correlated with long absence spells. As expected, the most important driver of long-term absenteeism is health status. This is not surprising since the main reasons for long-term absences are persistently low health stocks and health shocks like unexpected illnesses and accidents (Müller et al., 1998).

[Insert Table 4 about here]

Figure 2 reinforces our presumption of the validity of the common time trend assumption. As for long-term sick leave duration and the pre-reform periods, the two treatment groups as well as the control group show an almost parallel time trend.

[Insert Figure 2 about here]

Fourth, we not only estimate the reform effects on the incidence of long-term absenteeism but also the effects on the length of long-term absence spells. Although we work with survey data, which makes it possible to take a rich set of background variables into account (at the cost of having no detailed spell data), we have good arguments why the available sick absence information is sufficient (see Section 5.1).

Fifth, to prove the consistency of our results, we perform various robustness checks. Thanks to the panel structure, we are able to control for the labor force composition by using balanced panels. Moreover, we experiment with different pre- and post-reform years and pool the data over only two years. Additionally, we restrict the sample size to singles, persons aged 25 to 55 employed full-time, and split the sample at the median wage.

In recent years, there has been an extensive debate about the drawbacks and limitations of DiD estimation. A particular concern is the underestimation of OLS standard errors due to serial correlation in case of long time horizons and unobserved (treatment and control) group effects. To deal with the serial correlation issue, we focus on short time horizons. As Bertrand et al. (2004) have shown, the main source for understating the standard errors stems from serial correlation of the outcome and the intervention variable and is basically eliminated when focusing on less than five periods. While there is consensus about the serial correlation problem, the issue with unobserved common group effects is still a matter of considerable debate. If one takes the objection of Donald and Lang (2007) seriously, then it would not be possible to draw inferences from DiD analyses in the case of few groups, meaning that no empirical assessment could be performed. We subscribe to the view of Wooldridge (2006), who says of the study by Donald and Lang (p. 18):

"DL criticize Card and Krueger (1994) for comparing mean wage changes of fast-food workers across two states because Card and Krueger fail to account for the state effect (New Jersery or Pennsylvania) [...]. But the DL criticism in the G = 2 case is no different from a common question raised for any difference-in-differences analyses: How can we be sure that any observed difference in means is due entirely to the policy change? To characterize the problem as failing to account for an unobserved group effect is not necessarily helpful."[11]

[11] In this very readable extended version of an older published AER paper (Wooldridge, 2003), Wooldridge (2006) discusses several other shortcomings and assumptions of the estimation approach

Besides our focus on short time spans to resolve serial correlation concerns, we use robust standard errors and correct for clustering at the individual level throughout the analysis.

Finally, an important feature of this study is that there is no selection into or out of the treatment group, which is a central issue in other settings, e.g., when labor market programs are evaluated. Switching between the two diverse health care systems that were affected differently by the reforms is not allowed for the great majority. We are able to identify the only subsample that has this right and exclude it in our robustness checks.[12]

Our basic empirical strategy is thus to pool the data for the years 1995 to 1998 and to estimate DiD probit as well as count data models where we employ the variables Treatment Group 1 to 3 as well as Treatment Index 1 and 2, respectively.

7 Results

Table 5 provides the unconditional DiD estimate of the reform's net effect and direct effect on the incidence of long-term absenteeism. The unconditional long-term absence rate fell for Treatment Group 1 from 6.16 percent in the pre-treatment years 1995 and 1996 to 5.92 percent in 1997 and 1998. The rate for Treatment Group 2 fell from 3.77 to 3.56 percent. Without the availability of a control group and by means of before-after estimators one could erroneously attribute the total decrease to the reform. However, the absence rate for the Control Group also fell from 3.49 to 3.11 percent, resulting in overall difference-in-differences estimates of +0.13 and +0.17 percent, respectively. Table 6 shows the same estimates for the duration of long-term absence spells. The average number of benefit days per insured person

proposed by Donald and Lang (2007). In another place, Wooldridge (2007) asks rhetorically whether introducing more than sampling error into DiD analyses was either necessary or desirable. *"Should we conclude nothing can be learned in such settings?"*, he questions (p. 3). Moreover, he uses the well known Meyer et al. (1995) study as another example:
"It seems that, in this example, there is plenty of uncertainty in estimation, and one cannot obtain a tight estimate without a fairly large sample size. It is unclear what we gain by concluding that, because we are just identifying the parameters, we cannot perform inference in such cases. In this example, it is hard to argue that the uncertainty associated with choosing low earners within the same state and time period as the control group somehow swamps the sampling error in the sample means." (p.3 to 4).

[12] The only group that has the right to opt out of the SHI is that of optionally insured employees (self-employed and high-income earners above the income threshold). However, it is very unlikely that employees opted out of the SHI as a reaction to the cut in long-term sick pay. Opting out is a lifetime decision that is practically not feasible for the elderly due to extremely high premiums and that makes no sense for the young since they are very likely to be unaffected by long-term absenteeism anyway. We consider the possibility that selection out of the treatment played a role in Section 7.

fell from 3.62 to 3.17 days for Treatment Group 1 and from 2.58 to 1.95 days for Treatment Group 2. It also decreased slightly from 1.98 to 1.95 days for the Control Group leading to unconditional DiD estimates of -0.42 and -0.61 days.

[Insert Table 5 and 6 about here]

The DiD estimator is now incorporated into a regression framework. Table 7 reports the results from six model specifications that differ with respect to the inclusion of additional controls and measure the impact on the incidence of long-term absenteeism. Each specification represents a probit model equivalent to equation (11) with a dependent variable that is 1 if the respondent had a long-term sickness spell in the previous year and zero otherwise. The variable of interest is displayed as DiD1 and consists of an interaction between the dummy *Treatment Group 1* and the year dummy *p1997*. In every specification, marginal effects are calculated and displayed. In none of the model specifications is the DiD estimate statistically different from zero. The estimated coefficients are very close to zero, 0.0063 in the preferred specification, and positive. Note that there was no time trend in 1997 that significantly affected the absence rates, and that the DiD coefficients are robust to the inclusion of covariates and close to the unconditional DiD estimates. This can be interpreted as an additional evidence for a common time trend.

[Insert Table 7 about here]

In the next step, we disentangle the net effect of the reform into a direct effect and an indirect effect, and estimate their impact on the incidence of long-term absenteeism separately. As has been shown theoretically in Section 4, this is crucial since it may be that the indirect reform effect compensated the direct effect, rendering the net reform effect insignificant and highlights the importance of the separate analysis displayed in Table 8. Column 1 once again shows the net effect; the regression model equals Model 6 in Table 7. Column 2 displays the direct effect of the reduction in long-term sick pay on the absence rate. Again, we used equation (11) but in contrast to column 1, *Treatment Group 2*, i.e., those *only* affected by the cut in long-term sick pay, has been interacted with the post-reform year dummy to get the DiD2 estimate. It is easy to see that the DiD2 coefficient is statistically not different from but close to zero, which is also the case for DiD3 in column 3 where we used *Treatment Group 3* which contrasts those solely affected by the cut in long-term sick pay with those affected by both reforms.

[Insert Table 8 about here]

Treatment Index 1 and *2* represent the treatment intensity of the reform, namely the cut in long-term sick pay as a percentage of the individual's gross wage. As before, we use these variables to estimate the net effect as well as the direct effect of the reforms on the incidence of long-term absenteeism. And as before, we are unable to reject the hypothesis that the difference-in-differences estimate is statistically different from zero. Note that all DiD point estimates are practically zero (see Table 9).

[Insert Table 9 about here]

Table 10 gives us the DiD estimates when we use the number of days that long-term benefits were received as dependent variable and estimate count data models using Treatment Index 2. The non-significant point estimate for the whole sample is -0.041, and conditional on those who had a long-term absence spell, it is -0.904 (days). To sum up, we do not find evidence that the reforms had an overall significant impact on absenteeism – either on the incidence or on the length of long-term absence spells.

[Insert Table 10 about here]

One piece of "eyeball evidence" supporting this conclusion is descriptive statistics from the German Federal Statistical Office. These statistics show a slight decrease from 93.87 benefit days per case and SHI insured in 1996 to 88.93 benefit days in 1997 which lies within the usual fluctuation range (e.g. 1995: 86.47) and is in line with our results (German Federal Statistical Office, 2008).

7.1 Robustness Checks and Heterogeneity in Effects

Until now our estimation strategy was to pool the data over four years, which means that we allowed the sample composition to change over the years. As people with long-term absence spells have a higher probability to leave the labor force as a result of their (probably severe) illness, we should check whether this selection out of the labor market distorted our results. From those who had a long-term absence spell in 1996, 7.1 percent did not answer the questionnaire one year later for unknown reasons (one respondent died and one moved abroad). We do not find evidence that long-term illness led to a higher probability of dropping out of the sample in the subsequent year, since 7.7 percent of the respondents without long-term absence spells did not participate in the following year. One the other hand, 74.6 percent of those who were absent for a long-term period in 1996 were employed full-time at that time, whereas one year later, this number decreased to 62.3 percent for those who remained in the sample.[13] Especially if we had found a significant reform effect, one could have argued that the estimate was biased and caused by selection out of the labor market. There are several reasons why this selection effect is only of minor importance in our setting.

First, in light of the selection, it is even more remarkable that we do not find significant reform effects. Second, in 1998 (with information about 1997) the SOEP group drew a random refreshment sample that covered all existing subsamples and a total of 1,067 observations (Wagner et al., 2007). Thanks to this refreshment sample, the employment status distribution over those who had long-term sickness spells in 1996 and 1997 remained very stable. Under the consideration of the new observations, in total 73.1 percent of those who suffered long-term absence spells in 1997 were employed full-time (as compared to 62.3 percent without considering the refreshment sample).[14]

Third, through the availability of a control group that we observe over time, we are able to control for treatment-independent selection out of the labor market.[15] In the absence of a control group one could easily confuse the illness-related selection out of the labor market with a causal reform effect, since it is natural that sickness absence rates decrease over time as the sample ages.

Finally, as we use panel data, we can take account of labor force composition by using a balanced sample. In the following, we perform additional tests to prove the robustness of our results and to check whether heterogeneity in the reform effects is of importance. Table 11 reports results for the direct effect specification on the incidence of long-term absenteeism using *Treatment Index 2*.

[Insert Table 11 about here]

As a first test, we center the data two years around the reform (column 1). Af-

[13] The ratio of full time employed who were not absent for long-term periods was 71.9 percent in 1996 and 72.6 percent in 1997.

[14] For the other employment status groups, the deviation was less than 1.6 percent.

[15] We cannot, however, entirely exclude the possibility that the reform had an effect on the decision to leave the labor market voluntarily. We are unable to observe how large the share of voluntary labor market quitters was. However, as the cut in long-term sick pay was moderate and financial penalties are substantially higher for unemployed or retirees, we believe that reform-induced selection out of the labor market plays a negligible role.

terwards, we restrict our sample to the years 1996 and 1997, balance it, and consider only employees who were eligible for long-term sick pay in both years and answered the SOEP questionnaire in both years (column 2). An alternative robustness check would be to take 1995 as reference year and contrast it to 1997 and 1998. It might be that pull-forward effects played a role and that people adapted their behavior in 1996, when the reform plans were made public (column 3). However, this is not very probable as many catalysts of long-term absences happen unexpectedly. Since people who started their long-term absence spell in 1996 and carried it over to 1997 took advantage of a transitory arrangement and were not exposed to reduced sick pay, we contrasted the years 1995/1996 and 1998 in column 4. Another check would be to restrict the sample to full-time employed people aged 25 to 55 (column 5) and to singles (column 6) as the income of other household members may have an impact on the exposure to treatment. On the household level, the relevant parameter might be the decrease in total household income rather than individual wages. Since optionally SHI insured could have switched to the PHI system as a reaction to the reform, we exclude all optionally insured people in column 7. We also split the sample at the median gross wage (columns 8 and 9). As can be seen easily in Table 11, none of the difference-in-differences estimates is statistically significant and all point estimates are very close to zero in magnitude.

[Insert Table 12 about here]

We employ the same specifications with the number of benefit days as dependent variable and estimate count data models using Treatment Index 2. As can be seen in Table 12, we find significant and negative reform effects on the length of long-term absence spells for middle-aged full-time employed and the poor, which suggests heterogeneity in the reform effects. Middle-aged full-time employed people most likely need to support a family and might be the main earners in their household. The poor are also likely to be more crucially dependent on their full salary, which would imply that the reform induced a higher degree of presenteeism in these subsamples. On the other hand, the poor are more likely to work in less satisfying jobs and, thus, the reform might have reduced the degree of moral hazard as well.

According to the estimates, a one unit increase in Treatment Index 1 which equals an increase in the absence costs of about 5 percent, led to a decrease in the average number of benefit days per case of around 0.04 and 0.11, respectively.

[Insert Table 13 about here]

Besides displaying graphs on the outcome variable trends by treatment status, another standard method for checking the robustness of DiD estimates is to perform placebo regressions and to estimate the reform effects for years without a reform. For the assumption of common time trends of control and treatment group to hold, none of the placebo reform effects should be significant. Table 13 displays placebo regression results on the incidence and duration of long-term absenteeism for the years 1994 to 1996. All placebo estimates turn out to be insignificant.

[Insert Table 13 about here]

7.2 Calculation of SHI Reform Savings

In this subsection we calculate the total amount that the SHI has saved from 1997 to 2006 through the cut in long-term sick pay. The sum reflects the redistributional effect of the reform; reducing the replacement level for the long-term sick benefits the rest of the statutory health insurance pool through lower contribution rates.

For every eligible individual and the years 1997 to 2006, we calculate the sick pay according to the old and the new regulations, take the difference, and sum over the frequency-weighted number of long-term absences for the whole period. The long-term sick pay amounted to 80 percent of the monthly gross wage before the reform and to 70 percent after the reform up to the contribution ceiling. The benefit cap decreased from 100 percent of the monthly net wage before the reform to 90 percent after the reform.

Already in 1995, the German Federal Constitutional Court (*Bundesverfassungsgericht*) pronounced the common practice to calculate long-term sick pay to be unconstitutional.[16] The Court criticized that those insured under the SHI would be forced to pay contribution rates on lump-sum payments like Christmas or vacation bonuses (up to the contribution ceiling) but that these lump sum payments would not be considered in the calculation of the sick pay. However, the legislator ignored these objections when passing the reform bill at the end of 1996. From 1997 to 2000, sick pay was calculated without considering lump-sum payments, but several Federal Social Court (*Bundessozialgericht*) actions were filed. In 2003, the Federal Social Court judged in favor of the plaintiff.[17] The claimants whose sick pay was miscalculated between January 1, 1997 and June 22, 2000 were set a time limit of about five months to make an application for reimbursement of their miscalculated

[16] The judgment was pronounced at January 11, 1995 and is categorized under BVerfGE 92, 53.
[17] The judgement was pronounced at March 25, 2003 and is categorized under B 1 KR 36/01 R.

sick pay. From June 22, 2000, on, lump sum payments were considered (up to the contribution ceiling) in the calculation of long-term sick pay.

As it is unknown how many percent of the claimants filed an application within this rather restrictive time frame, our calculation specifications assume both full and zero reimbursement. Another question is whether the cut in long-term sick pay sensitized the population and caused the lawsuits. To deal with these unknowns, we formulate three scenarios. Specification I assumes that zero reimbursement of the miscalculated sick pay was provided – if no reform had taken place as well as in reality. Specification II assumes full reimbursement of the miscalculated sick pay which equals the assumption that lump sum payments have been considered from 1997 onwards. Both specifications assume independence of the cut in long-term sick pay and the lawsuits. Specification III assumes that there had not been a change in the basis of calculation without the reform and that in reality, the change became not effective until 2000.

We take advantage of the rich SOEP dataset that not only provides generated gross and net income measures but also provides the sum of yearly bonuses per employee. In a first step, we calculate the amount of long-term sick pay that every eligible individual would receive per day according to the pre- and the post-reform regulations and our three specifications. Observations with nonsense income data were dropped.[18]

In a second step, we use administrative data from the German Ministry of Health on the total number of SHI long-term sick pay cases and the average number of benefit days for SHI-insured people. Every statutory health insurance provider (2006: 253) is legally obligated to file information about the insured person and the benefits provided. The data are collected, aggregated, and published by the German Ministry of Health. Unfortunately, only the total number of long-term sickness cases and the average length of sick pay received is available. No personal data or income information is collected. Hence, we combine administrative data with the SOEP dataset, which contains very detailed income information.

Comparing the frequency-weighted number of SHI long-term sickness cases in the SOEP with the administrative data reveals that the SOEP slightly underestimates the number of cases as well as the average benefit days per case. This is not surprising since especially long-term sick people with very long sickness spells have a higher probability of not participating in the survey.

[18] We dropped respondents who claimed to be employed full-time and to earn less than €400 per month. Additionally, we dropped part-time employees who claimed to earn less than €200 per month.

Now consider Table 14. All values are expressed in euros and inflation-adjusted, with 2005 as the reference year. Columns (1), (2), and (3) show the estimates according to our three model specifications. The first row displays the difference between the average sick pay per case when the pre- and the post-reform regulations are compared. The sick pay per day and individual affected is calculated with SOEP data and is then multiplied by the average number of benefit days for those who had a long-term absence spell according to the Ministry of Health (2006: 76.07 days per case). Through the reform, the long-term sick pay has been cut on average by approximately €300 per case and year. Since (reduced) social contributions are charged on long-term sick pay, the net cut per case was about €250. Given that the average number of benefit days equals about 2.5 months, this translates into a monthly net cut of about €100, which represents about seven percent of the average monthly net wage.

[Insert Table 14 about here]

The second row presents the estimates when we consider the frequency-weighted long-term absence cases in SOEP. All eligible SHI-insured people are included; since we slightly underestimate the total number of cases, we take these estimates as the lower bound. According to these estimates, the SHI expenditures decreased between 4.3 and 5 billion euros as a result of the reform. The third row displays the total amount saved when we only consider compulsorily SHI-insured who are employed[19] and use administrative data on the number of cases instead of SOEP data. These values can also be seen as lower bound estimates. Row four, by contrast, shows the estimates when we consider all SHI-insured who are eligible for long-term sick pay according to official statistics.[20] All in all, we estimate that the total reform-induced SHI health expenditure savings from 1997 to 2006 was between 3.8 and 5.7 billion euros depending on the assumptions. When considering all eligible SHI-insured people and assuming that the change in the calculation basis was independent of the reform and full reimbursement of miscalculated sick pay was provided, our

[19] Students and (early) retirees are not considered although they might be eligible under special conditions.

[20] These values sighlty overestimate the true effect since short-term unemployed, who are eligible for long-term sick pay, are included. However, in theory, we would need to differentiate between two types of unemployed with long-term sick pay. First, if people become unemployed during their sickness, they receive long-term sick pay according to their last wage and are affected by cuts in sick pay. These people are of interest for us and they are included in row four as well as in the SOEP if they became unemployed during the calendar year prior to the interview. However, the second type of unemployed are those who become long-term sick during their unemployment period. In that case, they receive long-term sick pay which equals the unemployment benefits. Those unemployed were not affected by the cut in long-term sick pay but are included in row four.

estimate yields a total SHI saving sum of 5.62 billion euros. Under this specification, deducting social contributions yields a net loss for the long-term sick of about five billion euros.

8 Discussion and Conclusion

Economists often assume that moral hazard is responsible for a significant fraction of workplace absences, thereby contributing to rising health expenditures and labor costs. If this assumption holds true, it justifies reductions in sick pay replacement levels, which would eventually lower absence rates and durations, increase efficiency in the insurance market, and decrease health expenditures and labor costs. Several countries with public health insurance systems have indeed reduced the replacement levels for sick pay in recent years. Concurrently, various studies have found that people adapt their short-term absence behavior to economic incentives, providing evidence of the existence of a considerable degree of moral hazard in the decision to go on sick leave.

The aim of this study has been to analyze the causal effects of cuts in sick pay on long-term absenteeism. In Germany, two health reforms came into force at the end of 1996. The first reduced the compensation level for the first six weeks of receiving sick pay from 100 percent to 80 percent of foregone gross wages. The second reduced the compensation level from the seventh week onwards from 80 to 70 percent.

We show that within a simple dynamic model of absence behavior, the net effect of the two reforms on long-term absenteeism is a priori unclear, as it is composed of two diverging effects. The direct effect increases the costs of being absent for long-term periods and leads to a decrease of long-term absenteeism. The indirect effect has a positive impact on long-term absenteeism since through the two reforms, the costs of being absent for a long term decreased relative to the costs of being absent for a short term. The reform effects are derived under the assumption that the individuals' sickness levels are independent of previous periods and that every sickness level is equally probable. If we relax this assumption and assume that employees who are sick for long-term periods are seriously ill, the sick pay incentive structure breaks down and employees who are sick long-term do not change their absence behavior as a reaction to moderate cuts in replacement levels.

The identification and estimation of the direct as well as the net effect is feasible by difference-in-differences. SOEP data and the two-tiered health insurance system in Germany allow us to identify subsamples that were affected by both reforms, only by the reduction in long-term sick pay, and by neither reform. Moreover, the legislator defined an upper limit for long-term sick pay that decreased from 100 percent of net wages to 90 percent of net wages as a consequence of the reform. Hence, an additional source of exogenous variation is provided that does not only allow us to assign employees to treatment and control groups but also makes it possible to differentiate by treatment intensity in percent of the gross wage. Every part of the reform was distinctly exogenous to the individual and politically determined. Moreover, selection into or out of the treatment is not an issue here, as switching between the SHI and the PHI is not allowed due to rigid legal restrictions.

Our empirical findings suggest that the health reforms have, for the population average, not led to a significant change in the incidence and length of long-term absence spells. These results are robust to various specifications. Although we do not find general reform effects, we find evidence of heterogeneity in the effects. According to our estimates, the reform induced significant decreases in the length of long-term sickness spells for the poor and middle-aged employees employed full-time.

The finding that the long-term sick have not adapted their sickness behavior to the monetary reform incentives in a significant manner is in line with our model predictions if long-term sick people are assumed to be seriously ill. This is plausible since, in Germany, the most common causes for sickness spells of more than six weeks are chronic diseases of the spine, arthritis, accidents, cancer, and mental diseases. Moreover, 43 percent of the persons concerned have strong or very strong fears of being laid off and becoming unemployed (Müller et al., 1998). Interestingly, our results are in contrast to a study from Sweden that found absence behavior to be affected considerably by economic incentives even when absence spells of more than 90 days are assessed. The differences in the findings might be due to a) cultural peculiarities, e.g. Germans are said to have a particularly strong work ethic, b) different monitoring systems for sick leave, c) different reform settings, e.g. in this study, on average, the treated faced a monthly long-term sick pay cut of €100 or seven percent of their net wage, d) the application of different econometric techniques, e.g., in contrast to the Swedish study, we do not rely on before-after estimates but use a control group.

By combining SOEP income data with administrative data, we estimate the total SHI reform savings from 1997 to 2006 to lie between 3.8 and 5.7 billion euros in real terms as of 2005. The most realistic scenario yields a sum of about five billion euros that was redistributed from the long-term sick to the insurance pool in order to achieve lower contribution rates.

Various pieces of evidence throughout this study allow us to infer that moral hazard is of minor importance when sickness spells of more than 30 days are considered. Consequently, health reforms like the German one do not lead to more efficient sickness insurance markets by decreasing moral hazard but are merely an instrument to cut health expenditures or labor costs. On the other hand, if introduced together with with moderate cuts in replacement levels, this cost containment instrument seems to be economically efficient in the sense that it induces no major behavioral changes that might lead to undesirable equilibria. Policy makers should be aware of the reform effects and the redistributional consequences. It is simply a normative question whether this instrument to cut health expenditures should be applied.

Further research on how sickness absence, moral hazard, and presenteeism are related to the design of insurance contracts is essential as it has short and long-term consequences for health expenditures, health outcomes, labor costs, and productivity.

References

Ai, C. and E. C. Norton (2004). Interaction terms in logit and probit models. *Economics Letters 80*, 123–129.

Askildsen, J. E., E. Bratberg, and Ø. A. Nilsen (2005). Unemployment, labor force composition and sickness absence: A panel study. *Health Economics 14*, 1087–1101.

Badura, B., H. Schröder, and C. Vetter (2008). *Fehlzeiten-Report 2007: Arbeit, Geschlecht und Gesundheit* (1 ed.). Springer Medizin Verlag.

Barmby, T., J. Sessions, and J. Treble (1994). Absenteeism, efficiency wages and shirking. *Scandinavian Journal of Economics 96*(4), 561–566.

Bertrand, M., E. Duflo, and M. Sendhil (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Besley, T. and A. Case (2000). Unnatural experiments? Estimating the incidence of endogenous policies. *Economic Journal 110*(467), 672–694.

Brown, S. (1994). Dynamic implications of absence behaviour. *Applied Economics 26*, 1163–1175.

Brown, S. and J. G. Sessions (1996). The economics of absence: Theory and evidence. *Journal of Economic Surveys 10*(1), 23–53.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications* (1 ed.). Cambridge University Press.

Card, D. and A. B. Krueger (1994). Wages and employment: A case study of the fast-food industry in new jersey and pennsylvania. *American Economic Review 84*(4), 772–793.

Chatterji, M. and C. J. Tilley (2002). Sickness, absenteeism, presenteeism, and sick pay. *Oxford Economic Papers 54*, 669–687.

Curington, W. P. (1994). Compensation for permanent impairment and the duration of work absence: Evidence from four natural experiments. *The Journal of Human Resources 29*(3), 888–910.

Deb, P. and P. K. Trivedi (1997). Demand for medical care by the elderly: A finite mixture approach. *The Journal of Applied Econometrics 12*(3), 313–336.

Dionne, G. and B. Dostie (2007). New evidence on the determinants of absenteeism using linked employer-employee data. *Industrial & Labor Relations Review 61*(1), 108–120.

Donald, S. G. and K. Lang (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics 82*(2), 221–233.

Engellandt, A. and R. T. Riphahn (2005). Temporary contracts and employee effort. *Labor Economics 12*, 281–299.

Frick, J. R. and M. M. Grabka (2005). Item-non-response on income questions in panel surveys: Incidence, imputation and the impact on inequality and mobility. *Allgemeines Statistisches Archiv 89*(1), 49–60.

German Federal Statistical Office (1998). *Statistical Yearbook 1998 For the Federal Republic of Germany*. Metzler-Poeschel.

German Federal Statistical Office (2008). *Federal Health Monitoring*. www.gbe-bund.de, last accessed at June 25, 2008.

German Ministry of Health (2008). www.bmg.bund.de, last accessed at February 22, 2008.

Henrekson, M. and M. Persson (2004). The effects on sick leave of changes in the sickness insurance system. *Journal of Labor Economics 22*(1), 87–113.

Ichino, A. and R. T. Riphahn (2005). The effect of employment protection on worker effort. A comparison of absenteeism during and after probation. *Journal of the European Economic Association 3*(1), 120–143.

Jensen, S. and J. McIntosh (1999). Absenteeism in the workplace: Results from danish sample survey data. *Applied Economics Letters 6*, 337–341.

Johansson, P. and M. Palme (1996). Do economic incentives affect work absence? Empirical evidence using swedish micro data. *Journal of Public Economics 59*(1), 195–218.

Johansson, P. and M. Palme (2002). Assessing the effect of public policy on worker absenteeism. *Journal of Human Resources 37*(2), 381–409.

Johansson, P. and M. Palme (2005). Moral hazard and sickness insurance. *Journal of Public Economics 89*, 1879–1890.

Lindbeck, A., M. Palme, and M. Persson (2006). Job security and work absence: Evidence from a natural experiment. Working Paper Series 660, Research Institute of Industrial Economics. http://ideas.repec.org/p/hhs/iuiwop/0660.html, last accessed at October 23, 2008.

Medizinischer Dienst der Krankenversicherung (MDK) (2008). www.mdk.de, last accessed at 23.10.2008.

Meyer, B. D., W. K. Viscusi, and D. L. Durbin (1995). Workers' compensation and injury duration: Evidence from a natural experiment. *American Economic Review 85*(3), 322–340.

Müller, R., D. Hebel, B. Braun, R. Beck, U. Helmert, G. Marstedt, and H. Müller (1998). *Auswirkungen von Krankengeld-Kürzungen: Materielle Bestrafung und soziale Diskriminierung chronisch erkrankter Erwerbstätiger. Ergebnisse einer Befragung von GKV-Mitgliedern* (2 ed.). Schriftenreihe zur Gesundheitsanalyse, Volume 1. GEK Edition.

OECD (2006). *OCED Health Data 2006.*

Ose, S. O. (2005). Working conditions, compensation and absenteeism. *Journal of Health Economics 24*, 161–188.

Pettersson-Lidbom, P. and P. Skogman Thoursie (2008). Temporary disability insurance and labor supply: evidence from a natural experiment. Working paper, Stockholm University, Department of Economics. http://people.su.se/ pepet/tdi.pdf, last accessed at March 19, 2008.

Puhani, P. A. (2008). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. IZA Discussion Paper Series 3478, IZA. http://www.iza.org, last accessed at February 22, 2008.

Riphahn, R. T. (2004). Employment protection and effort among German employees. *Economics Letters 85*, 353–357.

Social Security Administration (2006). *Annual Statistical Supplement 2006, Table 9.A2.* http://www.ssa.gov/policy/docs/statcomps/supplement/2006/9a.html, last accessed at March 19, 2009.

Social Security Administration (2008). *Annual Statistical Supplement 2006, Table 9.C1.* http://www.ssa.gov/policy/docs/statcomps/supplement/2008/9c.html, last accessed at March 19, 2009.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel study (SOEP) - evolution, scope and enhancements. *Journal of Applied Social Science (Schmollers Jahrbuch) 127*(1), 139–169.

Winkelmann, R. (2008). *Econometric Analysis of Count Data* (5 ed.). Springer.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review 93*(2), 133–138.

Wooldridge, J. M. (2006). Cluster-sample methods in applied econometrics: an extended analysis. Working paper, Michigan State University, Department of Economics. https://www.msu.edu/ ec/faculty/wooldridge/current research/clus1aea.pdf, last accessed at March 19, 2009.

Wooldridge, J. M. (2007). What's new in econometrics? Imbens/Wooldridge lecture notes; summer institute 2007, lecture 10: Difference-in-differences estimation, NBER. http://www.nber.org/minicourse3.html, last accessed at March 19, 2009.

Figure 1: Replacement Levels for Short and Long-Term Absence Spells

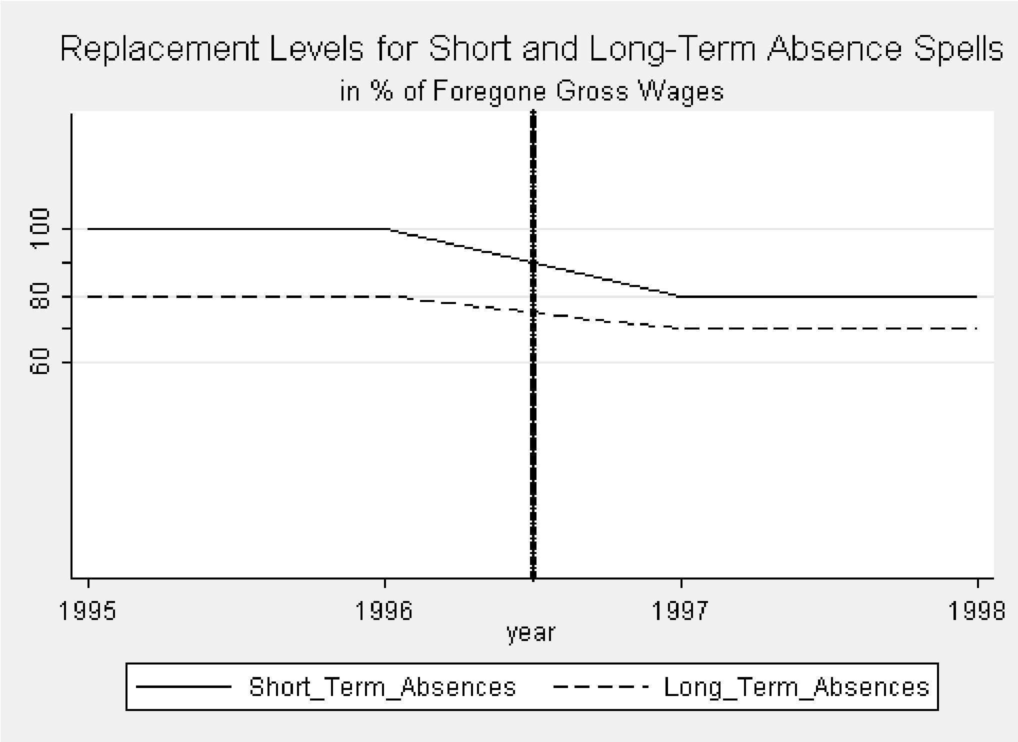

Figure 2: Logarithm of Long-Term Absent Benefit Days

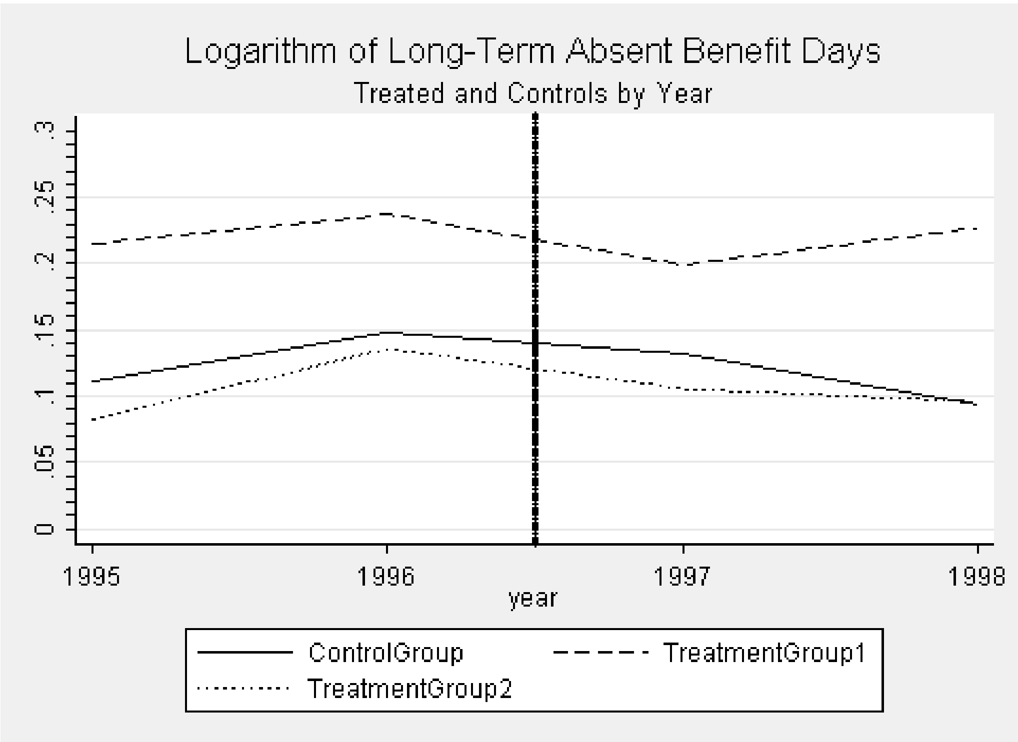

Table 1: Definition of Subsamples

	Reduction Sickness Compensation < 30 days (paid by employer)	**Reduction Sickness Compensation > 30 days (paid by SHI)**
Private-sector employees with SHI (1)	yes	yes
Public-sector employees with SHI (2)	no	yes
Trainees with SHI (3)	no	yes
Public-sector employees with PHI (4)	no	no
Self-employed with PHI (5)	no	no

Table 2: Overview Treatment and Control Groups

Effect to be estimated	**Treatment groups**	**Control group**
Net effect	subsample (1) **Treatment Group 1**	subsamples (4) + (5)
Direct effect	subsamples (2) + (3) **Treatment Group 2**	subsamples (4) + (5)

Table 3: Variable Means by Treatment and Control Groups

Variable	Control Group	Treatment Group 1	Treatment Group 2
Long-term absent	0.033	0.060	0.026
Long-term absent benefit days	1.965	3.392	2.249
Personal characteristics			
Female	0.410	0.366	0.587
Age	40.57	39.86	37.48
Age square/100	17.58	17.01	15.60
Immigrant	0.097	0.215	0.112
East Germany	0.166	0.258	0.378
Partner	0.762	0.803	0.650
Married	0.673	0.696	0.569
Children	0.483	0.470	0.435
Disabled	0.033	0.052	0.053
Good health	0.648	0.607	0.604
Bad health	0.080	0.099	0.104
No sports	0.287	0.409	0.331
Educational characteristics			
Dropout	0.021	0.050	0.044
Degree after 8 years of schooling	0.230	0.357	0.271
Degree after 10 years of schooling	0.290	0.330	0.438
Degree after 12 years of schooling	0.051	0.035	0.035
Degree after 13 years of schooling	0.363	0.115	0.162
Other degree	0.046	0.112	0.051
Work in job trained for	0.608	0.545	0.511
New job	0.204	0.179	0.179
No. of years in company	10.29	9.04	8.79
Job characteristics			
No tenure	0.106	0.051	0.273
One man firm	0.099	0.000	0.000
Small company	0.327	0.274	0.169
Medium company	0.179	0.312	0.281
Big company	0.126	0.221	0.290
Huge company	0.268	0.193	0.260
Self employed	0.308	0.000	0.000
Blue collar worker	0.112	0.528	0.190
White collar worker	0.150	0.472	0.579
Public sector	0.493	0.000	0.829
Civil servant	0.395	0.000	0.031
Self employed	0.307	0.000	0.000
High job autonomy	0.506	0.160	0.152
Gross income per month	2,383.16	2,012.98	1,674.95
Regional unemployment rate	11.49	12.04	13.07
N	2,693	16,006	6,500

Table 4: Probit Model: Determinants of Long-Term Absenteeism

Variable	Coefficient	Standard Error
Personal characteristics		
Female (d)	-0.001	0.003
Age	0.000	0.003
Age squared/100	0.000	0.001
Immigrant (d)	0.004	0.005
East Germany (d)	-0.012	0.011
Partner (d)	0.006	0.004
Married(d)	-0.008*	0.005
Children (d)	-0.006**	0.003
Disabled (d)	0.034***	0.007
Good health (d)	-0.026***	0.003
Bad health (d)	0.076***	0.007
No sports (d)	0.007**	0.003
Educational characteristics		
Degree after 8 years' of schooling (d)	-0.006	0.006
Degree after 10 years' of schooling (d)	-0.008	0.007
Degree after 12 years' of schooling (d)	-0.018***	0.007
Degree after 13 years' of schooling (d)	-0.013**	0.006
Other degree (d)	-0.003	0.007
Work in job trained for (d)	-0.001	0.003
New job (d)	0.006	0.004
No. of years in company	-0.000	0.000
Job characteristics		
No tenure last year (d)	-0.009**	0.004
Medium size company (d)	0.0012***	0.004
Big company (d)	0.015***	0.004
Huge company (d)	0.014**	0.005
White collar worker (d)	-0.013***	0.003
High job autonomy (d)	-0.008*	0.004
Gross wage per month/1000	-0.005**	0.002
Regional unemployment rate	0.003	0.002
Year 1996 (d)	0.004	0.004
Year 1997 (d)	-0.004	0.006
Year 1998 (d)	-0.000	0.005
R-squared	0.106	
χ^2	916.944	
N	25199	

(d) for discrete change of dummy variable from 0 to 1
marginal effects, which are calculated at the means of the covariates, are displayed
* p<0.10, ** p<0.05, *** p<0.01
Dependent variable: dummy that is 1 if respondent had long-term absence spell; probit model is estimated
Standard errors in parentheses are adjusted for clustering on person id
Regression includes state dummies
Left out reference categories are dropout, blue collar worker, and small company

Table 5: Unconditional Difference-in-Differences Estimate on the Incidence of Long-Term Absenteeism

	1995/1996	1997/1998	Difference	Diff-in-Diff
Treatment Group 1	0.0616	0.0592	-0.0024	0.0013
	(0.0027)	(0.0026)	(0.0038)	(0.0078)
Treatment Group 2	0.0377	0.0356	-0.0020	0.0017
	(0.0034)	(0.0032)	(0.0047)	(0.0082)
Control Group	0.0349	0.0311	-0.0038	
	(0.0049)	(0.0048)	(0.0069)	

Average incidence rate of long-term absenteeism is displayed

Standard errors in parentheses

Table 6: Unconditional Difference-in-Differences Estimate on the Average Number of Long-Term Absent Benefit Days

	1995/1996	1997/1998	Difference	Diff-in-Diff
Treatment Group 1	3.6212	3.1747	-0.4464	-0.4219
	(0.2455)	(0.2277)	(0.3344)	(0.7358)
Treatment Group 2	2.5800	1.9461	-0.6339	-0.6094
	(0.3407)	(0.2689)	(0.4304)	(0.7836)
Control Group	1.9767	1.9522	-0.0245	
	(0.4194)	(0.4546)	(0.6177)	

Average number of long-term absent benefit days is displayed

Standard errors in parentheses

Table 7: Difference-in-Differences Estimation on the Incidence of Long-Term Absenteeism

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
DiD1 (d)	0.0035	0.0024	0.0053	0.0032	0.0061	0.0063
	(0.0119)	(0.0108)	(0.0101)	(0.0104)	(0.0088)	(0.0086)
Post reform dummy (d)	-0.0012	-0.0123	-0.0133	-0.0102	-0.0117	-0.0102
	(0.0124)	(0.0140)	(0.0140)	(0.0135)	(0.0127)	(0.0123)
Year 1996 (d)	0.0064	-0.0002	0.0003	0.0003	-0.0007	0.0001
	(0.0048)	(0.0053)	(0.0052)	(0.0052)	(0.0047)	(0.0047)
Year 1997 (d)	-0.0032	-0.0051	-0.0042	-0.0057	-0.0049	-0.0047
	(0.0050)	(0.0046)	(0.0045)	(0.0045)	(0.0042)	(0.0041)
Treatment Group 1 (d)	0.0276***	0.0244***	0.0151**	0.0219***	0.0145***	0.0124**
	(0.0062)	(0.0057)	(0.0063)	(0.0059)	(0.0053)	(0.0059)
Educational characteristics	no	no	yes	no	no	yes
Job characteristics	no	no	no	yes	no	yes
Personal characteristics	no	no	no	no	yes	yes
Regional unemployment rate	no	yes	yes	yes	yes	yes
State dummies	no	yes	yes	yes	yes	yes
R-squared	0.0049	0.0091	0.0308	0.0258	0.1046	0.1153
χ^2	30.368	51.609	187.191	153.235	704.315	780.916
N	18699	18699	18699	18699	18699	18699

(d) for discrete change of dummy variable from 0 to 1

Marginal effects are calculated at the means of the covariates except for *Treatment Group 1* (=1), *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)

* p<0.1, ** p<0.05, *** p<0.01

Dependent variable: dummy that is 1 if respondent had long-term absence spell; every column represents one probit model

Standard errors in parentheses are adjusted for clustering on person id

Table 8: DiD Estimation on Incidence: Direct vs. Indirect Effect

Variable	**Net effect**	**Direct effect**	**Direct vs. indirect effect**
DiD1 (d)	0.006		
	(0.009)		
Treatment Group 1 (d)	0.012**		
	(0.006)		
DiD2 (d)		0.010	
		(0.010)	
Treatment Group 2 (d)		-0.015	
		(0.012)	
DiD3 (d)			-0.000
			(0.004)
Treatment Group 3 (d)			-0.021***
			(0.006)
Post reform dummy(d)	-0.010	0.007	-0.000
	(0.012)	(0.012)	(0.004)
Year 1996 (d)	0.000	0.016*	0.002
	(0.005)	(0.009)	(0.003)
Year 1997 (d)	-0.005	0.009	-0.003
	(0.004)	(0.007)	(0.003)
Educational characteristics	yes	yes	yes
Job characteristics	yes	yes	yes
Personal characteristics	yes	yes	yes
Regional unemployment rate	yes	yes	yes
State dummies	yes	yes	yes
R-squared	0.115	0.106	0.114
χ^2	780.916	298.763	1074.389
N	18699	9193	22506

(d) for discrete change of dummy variable from 0 to 1

Marginal effects are calculated at the means of the covariates except for *Treatment Group 1 (2, 3)* (=1), *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)

* p<0.1, ** p<0.05, *** p<0.01

Dependent variable: dummy that is 1 if respondent had long-term absence spell; every column represents one probit model

Standard errors in parentheses are adjusted for clustering on person id

Table 9: DiD Estimation on Incidence with Varying Treatment Intensity

Variable	**Net effect**	**Direct effect**
DiD1	0.000	
	(0.001)	
Treatment Index 1	0.003***	
	(0.001)	
DiD2		0.000
		(0.001)
Treatment Index 2		0.000
		(0.002)
Post reform dummy(d)	-0.005	0.011
	(0.010)	(0.012)
Year 1996 (d)	0.000	0.016*
	(0.005)	(0.009)
Year 1997 (d)	-0.005	0.009
	(0.004)	(0.007)
Educational characteristics	yes	yes
Job characteristics	yes	yes
Personal characteristics	yes	yes
Regional unemployment rate	yes	yes
State dummies	yes	yes
R-squared	0.116	0.104
χ^2	785.887	291.684
N	18699	9193

(d) for discrete change of dummy variable from 0 to 1

Marginal effects are calculated at the means of the covariates except for *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)

* p<0.1, ** p<0.05, *** p<0.01

Dependent variable: dummy that is 1 if respondent had long-term absence spell

Every column represents one probit model

Standard errors in parentheses are adjusted for clustering on person id

Table 10: DiD Estimation on the Duration of Long-Term Absenteeism

	Zero-Inflated Model	***Hurdle-at-Zero Model***
Variable	**Direct effect: Varying Intensity**	**Direct effect: Varying Intensity**
DiD2	-0.041	-0.904
	(0.058)	(1.915)
Treatment Index 2	0.043	1.188
	(0.044)	(1.006)
Post reform dummy(d)	-0.402	-16.524
	(0.642)	(24.307)
Year 1996 (d)	-0.064	1.509
	(0.275)	(10.047)
Year 1997 (d)	0.242	0.071
	(0.326)	(14.345)
Educational characteristics	yes	yes
Job characteristics	yes	yes
Personal characteristics	yes	yes
Regional unemployment rate	yes	yes
State dummies	yes	yes
χ^2	149.552	108.45
N	9193	327

(d) for discrete change of dummy variable from 0 to 1

Marginal effects are calculated at the means of the covariates except for *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)

* p<0.1, ** p<0.05, *** p<0.01

Dependent variable: Number of long-term benefit days; every column represents one count data model

Standard errors in parentheses are adjusted for clustering on person id

Table 11: Robustness and Heterogeneity of Effects: Direct Effect on Incidence Using Treatment Index 2

Variable	**1996-1997**	**1996-1997; balanced**	**1995 vs. 1997/1998**	**1995/1996 vs. 1998**	**Full-time: age 25 - 55**	**Singles**	**No optionally insured**	**Less than median income**	**More than median income**
DiD2	0.000	0.002	0.001	0.002	0.001	0.002	0.001	0.001	0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Educational characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Job characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Personal characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Regional unemployment rate	yes	yes	yes	yes	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.096	0.123	0.084	0.089	0.095	0.110	0.079	0.118	0.101
χ^2	145.022	126.841	167.372	217.029	144.648	113.32	207.033	212.115	166.736
N	4595	3239	6786	6827	5204	2747	8435	4833	4289

(d) for discrete change of dummy variable from 0 to 1
Marginal effects are calculated at the means of the covariates except for *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)
* p<0.1, ** p<0.05, *** p<0.01
Dependent variable: dummy that is 1 if respondent had long-term absence spell; every column represents one probit model
Standard errors in parentheses are adjusted for clustering on person id

Table 12: Robustness and Heterogeneity of Effects: Direct Effect on Duration Using Treatment Index 2

Variable	**1996-1997**	**1996-1997; balanced**	**1995 vs. 1997/1998**	**1995/1996 vs. 1998**	**Full-time: age 25 - 55**	**Singles**	**No optionally insured**	**Less than median income**	**More than median income**
DiD2	-0.021	0.130	-0.035	-0.025	-0.041***	0.063	-0.093	-0.114**	-0.048
	(0.053)	(0.123)	(0.039)	(0.024)	(0.020)	(0.072)	(0.071)	(0.023)	(0.049)
Educational characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Job characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Personal characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Regional unemployment rate	yes	yes	yes	yes	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes	yes	yes	
χ^2	4608.620	1933.945	5256.873	2111.791	2478.681	222.277	235.314	2332.530	6751.009
N	4571	3334	6786	6812	5186	2798	8435	4833	4289

(d) for discrete change of dummy variable from 0 to 1
Marginal effects are calculated at the means of the covariates except for *Post reform dummy* (=1), *Year 1996* (=0), and *Year 1997* (=1)
* $p<0.1$, ** $p<0.05$, *** $p<0.01$
Dependent variable: number of long-term benefit days; every column represents one Zero-Inflated NegBin-2 Model
Standard errors in parentheses are adjusted for clustering on person id

Table 13: Placebo Estimates Using Treatment Index 2

Variable	**Direct effect (Incidence)**	**Direct effect (Duration)**
DiD96 (d)	0.001	-0.042
	(0.003)	(0.159)
DiD95 (d)	-0.003	-0.171
	(0.005)	(0.277)
DiD94 (d)	-0.010	-0.503
	(0.010)	(0.945)
Educational characteristics	yes	yes
Job characteristics	yes	yes
Personal characteristics	yes	yes
Regional unemployment rate	yes	yes
State dummies	yes	yes
χ^2	339.092	264.462
N	11457	11457

(d) for discrete change of dummy variable from 0 to 1

Marginal effects are calculated at the means of the covariates except for corresponding post reform dummies (=1), pre-treatment(=0), and post-treatment years (=1))

* p<0.1, ** p<0.05, *** p<0.01

Dependent variable: number of long-term benefit days; every column represents one Zero-Inflated NegBin-2 model

Standard errors in parentheses are adjusted for clustering on person id

Table 14: Total Amount Saved by SHI Due to Reform: 1997-2006

Average: 1997-2006	**Specification I (1)**	**Specification II (2)**	**Specification III (3)**
SHI reform savings per case	267	309	302
Total amount redistributed: Frequency weighted SOEP cases	4.266.472.300	4.967.670.277	4.874.958.639
Total amount redistributed: Compulsorily insured (Federal Statistical Office)	3.832.975.534	4.473.828.845	4.391.214.903
Total amount redistributed: All eligible SHI insured (Federal Statistical Office)	4.892.101.168	5.632.182.856	5.735.520.006

Source: SOEP, German Ministry of Health, own calculations

All values are in Euro, inflation-adjusted (2005=100), and weighted

Specification I assumes zero reimbursement of the miscalculated sick pay – if no reform had taken place as well as in reality

Specification II assumes full reimbursement of the miscalculated sick pay – if no reform had taken place as well as in reality

Specification III assumes that there wouldn't have been a change in the basis of calculation at all, if the reform had not been implemented; in reality, zero reimbursement of the miscalculated sick pay is assumed (1997 - 2000).

Appendix A

Table 15: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Long-term absence	0.051	0.221	0	1	25199
Long-term absent benefit days	2.944	19.732	0	335	25199
Treatment Group 1	0.856	0.351	0	1	18699
Treatment Group 2	0.707	0.455	0	1	9193
Treatment Group 3	0.289	0.453	0	1	22506
Treatment Index 1	5.699	2.755	0	10	18699
Treatment Index 2	4.652	3.32	0	10	9193
Personal characteristics					
Female	0.427	0.495	0	1	25199
Age	39.322	11.154	18	65	25199
Age squared/100	16.707	9.067	3.24	42.25	25199
Immigrant	0.176	0.381	0	1	25199
East Germany	0.28	0.449	0	1	25199
Partner	0.759	0.428	0	1	25199
Married	0.661	0.473	0	1	25199
Children	0.463	0.499	0	1	25199
Disabled	0.05	0.218	0	1	25199
Good health	0.611	0.488	0	1	25199
Bad health	0.098	0.298	0	1	25199
No sports	0.376	0.484	0	1	25199
Educational characteristics					
Drop out	0.045	0.208	0	1	25199
Degree after 8 years' of schooling	0.321	0.467	0	1	25199
Degree after 10 years' of schooling	0.354	0.478	0	1	25199
Degree after 12 years' of schooling	0.037	0.188	0	1	25199
Degree after 13 years' of schooling	0.154	0.361	0	1	25199
Other degree	0.089	0.285	0	1	25199
Work in job trained for	0.543	0.498	0	1	25199
New job	0.182	0.386	0	1	25199

Continued on next page...

... Table 15 continued

Variable	Mean	Std. Dev.	Min.	Max.	N
No. years in company	9.106	9.217	0	47.9	25199
Job characteristics					
No tenure	0.114	0.318	0	1	25199
One man company	0.011	0.104	0	1	25199
Small size company	0.253	0.435	0	1	25199
Medium size company	0.289	0.454	0	1	25199
Big company	0.229	0.42	0	1	25199
Huge company	0.218	0.413	0	1	25199
Blue collar worker	0.396	0.489	0	1	25199
White collar worker	0.465	0.499	0	1	25199
Public sector	0.267	0.442	0	1	25156
Civil servant	0.05	0.218	0	1	25199
Self-employed	0.033	0.178	0	1	25199
High job autonomy	0.195	0.396	0	1	25199
Gross wage per month	1965.35	1106.54	204.00	40903.35	25199
Regional unemployment rate	12.25	3.97	7	21.7	25199