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

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

December 14, 2010

Abstract

I theoretically and empirically disentangle the effects of cuts in the statutory sick pay levels on long-term absenteeism in Germany. The reforms have not induced significant changes in the average incidence and duration of sick leave periods longer than six weeks. The finding is theoretically confirmed assuming that the long-term sick are seriously sick. Thus, moral hazard seems to be less of an issue in the upper end of the sickness spell distribution. However, I find heterogeneity in the effects and significant duration decreases for certain subsamples. Finally, I calculate that within ten years, the cut in statutory long-term sick pay redistributed five billion Euros from the long-term sick to the insurance pool.

Keywords: long-term absenteeism, statutory long-term sick pay, moral hazard, natural experiment, Socio-Economic Panel Study (SOEP)

JEL classification: H51, I18, J22, J32

‡I am grateful to Daniela Andrén, Tim Barmby, Amitabh Chandra, Eberhard Feess, Joachim R. Frick, John P. Haisken-DeNew, Guido Imbens, Per Johansson, Martin Karlsson, Sonja Kassenböhmer, Henning Lohmann, Bruce D. Meyer, Andrew Newell, Hendrik Schmitz, Martin Peitz, Per Pettersson-Lidbom, Tom Siedler, Arne Uhlendorff, Hans-Martin von Gaudecker, Gert G. Wagner, the participants in a seminar in Mannheim as well as at the 2nd IZA/IFAU Conference on Labour Market Policy Evaluation, the 20th Annual Conference of the European Association of Labour Economists (EALE), the 13th Annual Meeting of the Latin American and Caribbean Economic Association (LACEA), and the 25th Annual Congress of the European Economic Association (EEA) for their helpful comments and discussions. A special thank 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, German Socio-Economic Panel Study (SOEP), Mohrenstraße 58, D-10117 Berlin, Germany, and Berlin University of Technology (TUB) e-mail: nziebarth@diw.de

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).

At the same time, legislative frameworks differ widely from one country to the next. In Europe, the statutory sickness absence insurance is integral part of the social insurance system. Typically, employers are obliged to provide sick pay for short-term absences, whereas health insurance providers or taxpayers compensate wage losses for the long-term sick. The U.S. do not know a statutory sickness insurance for short-term absentees on the federal level. However, the U.S. and Canada know the *workers' compensation insurance (WCI)* that is administered on a state-by-state basis and covers incomes losses due to *work-related* sickness or injury. On the federal level, the *disability insurance (DI)* replaces income losses stemming from a permanent labor market withdrawal due to work disability.

The literature on sickness absence in general is quite rich. It has been found that workplace conditions determine sick leave behavior (Dionne and Dostie, 2007) as well as probation periods and economic upswings or downturns (Ichino and Riphahn, 2005; Askildsen et al., 2005). However, empirical evidence concerning the relationship between the design of the sickness insurance scheme and sick leave behavior is scanty at best, especially as compared to other fields like the vast literature on unemployment benefits and unemployment duration. Some studies from Sweden have shown that employees adapt their sick leave behavior to changes in replacement levels (Johansson and Palme, 2002; Henrekson and Persson, 2004; Johansson and Palme, 2005). Moreover, Puhani and Sonderhof (2010) have shown that changes in statutory short-term sick pay affected the sick leave behavior in Germany. All studies cited above explicitly analyze the effects on short-term sickness absences within the European statutory sickness insurance. There is also empirical evidence from North America on the workers' compensation insurance (WCI) and the disability insurance (DI), althought the findings are inconclusive. While Meyer et al. (1995) find that an increase in WCI benefits in 1987 has led to increased injury duration, the results from the Curington (1994) study using data from the 1960s and 1970s are mixed. Besides the WCI, the DI has attracted a lot of attention among economists. Many studies find that the generosity of the DI affects labor supply behavior on the extensive margin (Bound, 1989; Gruber, 2000; Chen and van der Klaauw, 2008), although there is also convincing evidence that this might not be the case (Campolieti, 2004). Researchers have also studied the DI application process (e.g. de Jong et al. (2010)).

It is very important to keep in mind that the empirical findings concerning the DI and the WCI are unlikely to be directly transferable to the sickness absence insurance. While the European statutory sickness insurance covers all types of work-related *and* work-unrelated illnesses, the WCI solely covers the special case of a work-related illness or injury. On the other hand, both social insurances have in common that employees are still employed while being on sick leave. Thus, both focus on labor supply behavior on the *intensive* margin. In contrast to that, the DI deals with labor supply behavior on the *extensive* margin and hence a complete withdrawal from the labor market.

This is one of the very few existing papers that explicitly analyzes the impact of cuts in statutory long-term sick pay on long-term absenteeism, i.e., on sickness spells of more than six weeks. In Germany, statutory long-term sick pay is provided by the Statutory Health Insurance (SHI) system. In 1996, the total benefit sum amounted to €9.3 billion, comprising 7.3 percent of all expenditures by the SHI system. At that time, two health reforms were implemented, both of which cut the level of paid sick leave. I theoretically and empirically analyze the effects of both reforms on long-term absenteeism. Additionally, I calculate the reform-induced SHI savings and redistributional effects.

In the remainder of this manuscript, 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. Analogously, statutory sick pay during the first six weeks of a spell is defined as statutory short-term sick pay, while statutory long-term sick pay refers to episodes of more than six weeks.

The first reform cut statutory short-term sick pay from 100 to 80 percent of foregone gross wages, whereas the second reform cut statutory long-term sick pay from 80 to 70 percent of foregone gross wages. Both reforms generate exogenous sources of variation and yield testable implications.

To theoretically predict the effects of both reforms on long-term absenteeism, I 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, the cut in long-term sick pay should lead to a decrease in long-term absenteeism as the direct costs of being on long-term sick leave unambiguously increase. However, short-term sick pay was likewise cut at the same time. Since the cut in short-term sick pay was stronger than the cut in long-term sick pay, this might have triggered an indirect effect. Hence, second, from a theoretical point of view, the two reforms jointly may have affected long-term absenteeism in a positive way since the costs of long-term absences decreased relative to the costs of short-term absences. In other words, the gap in the replacement levels between short-term and long-term sick leave decreased as a consequence of the reforms. Later on, in Section 3, I derive the direct and the indirect effect by means of a simple theoretical model. However, under the assumption that employees on long-term sick leave are seriously sick, the incentive structure of the sick pay scheme would break down and individuals would not adapt their labor supply behavior to moderate cuts in sick pay.

Since Germany has two independent health care systems existing side by side, I am able to identify subsamples that were affected by none, one, or both of the reforms (see next section). Then, using data from the German Socio-Economic Panel Study (SOEP) and difference-in-differences methods, I can estimate the net and the direct effect of the two reforms on the incidence and duration of long-term absence spells.

My empirical findings in Section 6 indicate that, on population average, the cut in replacement levels did not affect the incidence and duration of long-term sickness spells, either directly or indirectly. This result is in line with my model predictions under the assumption that employees on long-term sick leave are indeed seriously sick. However, I find evidence of heterogeneity in the effects. For the poor as well as for middle-aged persons employed full-time, the duration of long-term absenteeism decreased significantly. Overall, my findings suggest that employees who have been on certified sick leave for more than six weeks are not very responsive to moderate monetary incentives, which implies that, in contrast to shorter absence spells, moral hazard is of less importance in the upper end of the sickness spell distribution. In the last subsection before I conclude, I calculate that from 1997 to 2006, the cut in statutory long-term sick pay redistributed five billion Euros from the long-term sick to the statutory health insurance pool for the benefit of lower contribution rates.

2 The German Health Care System and The Policy Reforms

2.1 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 or a private health insurance provider. 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 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. Privately insured people pay risk-related insurance premiums determined by an initial health checkup. The premiums exceed the expected expenditures in younger age brackets, since health insurance providers build up reserves for rising expenditures with increased age. Coverage is provided under a range of different health plans, and insurance contracts are subject to private law. Consequently, in Germany, public health care reforms apply only to the SHI, not to the PHI.

It is important to keep in mind that compulsorily insured persons have no right to choose the health insurance system or benefit package. They are compulsorily insured under the standard SHI insurance scheme. Once an optionally insured person (a high-income earner, self-employed person, or civil servant) opts out of the SHI system, it is practically impossible to switch back into it. Employees above the income threshold are legally forbidden from switching back, while employees who fall below the income threshold in subsequent years may do so under certain conditions, but are not able to carry along the reserves that their PHI providers have built up since these are not portable (neither between PHI and SHI, nor between the different private health insurance providers).[1] In reality, switching to a private health insurance provider may be regarded as a lifetime decision, and switching between the SHI system and PHI – as well as

[1] Until 2009, accrued reserves for rising health expenditures with increased age were not portable at all. From January 1, 2009 on, portability of accrued reserves between PHI providers has been made compulsive to a strictly defined extent.

between PHI providers – is therefore very rare.

2.2 The German Statutory Sick Pay Scheme

If an employee falls sick, a certificate from a physician is required from the fourth day of sick leave. The employer is legally obliged to provide statutory short-term sick pay 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 long-term sick pay is provided 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). Employees insured under the PHI insure the risk of falling long-term sick individually.

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 *Medizinischer Dienst der Gesetzlichen Krankenversicherungen* (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, 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.[2] In 2007, 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, 2008).

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

2.3 The Policy Reforms

Two health reforms were implemented at the end of 1996. First, from October 1996 on, the replacement level during the first six weeks of a sickness episode (i.e., statutory short-term sick pay) was reduced from 100 to 80 percent of foregone gross wages.[3] This reform had, at least theoretically, an indirect impact on sickness spells of more than six weeks and should therefore be considered. Second, the replacement level for absence spells of more than six weeks (i.e., statutory long-term sick pay) was cut from 80 to 70 percent of foregone gross wages for those insured under the SHI. This health reform act became effective on January 1, 1997. [4] Figure 1 illustrates how the two reform worked.

[Insert Figure 1 about here]

SHI statutory long-term sick pay is 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 (2009: 0.7*€ 3,675 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 taxes and social contributions. For example, a worker might earn € 2,500 gross per month and € 1,800 net per month. Then, the basic rule implied statutory long-term sick leave that amounted to 0.8*€ 2,500=€ 2,000. However, the second benefit cap limited the benefit to € 1,800 per month before the reform. The cut in statutory long-term sick pay decreased the replacement level to 70 percent of the gross wage (i.e., 0.7*€ 2,500=€ 1,750) and the benefit cap to 90 percent of the net wage (i.e., 0.9*€ 1,800=€ 1,620). As can be seen by means of this little example, benefit caps were also decreased in the course of the reform, depending on the relation between gross and net wages – which in turn is determined by the income level, the marital status, and the number of children – employees insured under the SHI were affected differently by the cut in long-term sick pay. This introduces additional exogenous variation which allows me to generate an index that mirrors the cut in statutory long-term sick pay for each individual on a continuous scale from

[3] 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 I am not able to precisely identify those employees who were effectively affected by this law, as employers and unions voluntarily agreed in some collective wage agreements to continue the old sick pay scheme. However, in principle, the law applied to all private sector employees whom I define below as being treated by the cut in statutory short-term sick pay. Using all private sector employees jointly as treatment group, Ziebarth and Karlsson (2010) have shown that the cut in statutory short-term sick pay reduced short-term absenteeism.

[4] 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.*

zero percent of the gross wage up to 10 percent of the gross wage.

Independent from the reforms analyzed in this paper, the German sick pay scheme exerts an incentive to substitute a long-term spell by several short-term spells since statutory sick pay for the latter is higher. However, German social legislation explicitly forbids such substitution of spells: If employees repeatedly call-in sick due to the same illness, they are no longer entitled to employer-provided statutory short-term sick pay.[5]

I 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 long-term sick leave replacement levels are subject to individual insurance contracts.

[**Insert Table 1 about here**]

Table 1 shows 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 cut in statutory long-term sick pay but not by the cut in statutory short-term sick pay due to political decisions. The same holds for SHI-insured trainees (subsample (3)). While subsample (1) is defined as *Treatment Group 1*, subsamples (2) and (3) are called *Treatment Group 2*. The last two subsamples, PHI-insured public-sector employees and PHI-insured self-employed persons, were not affected by any of the reforms and are called *Control Group*.

3 A Dynamic Model of Absence Behavior

In the following, I analyze the absence behavior of an individual i within a two-period model. I 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:

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

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.

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

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, I assume that $f(\sigma_t)$ follows a uniform distribution:

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

Hence 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, I define the statutory long-term sick pay as r_l with $0 < r_l < 1$ and the statutory short-term sick pay 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 I am primarily interested in the reform effects on long-term absenteeism, I assume that the 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:

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

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.[6]

[6] I assume a rigid employment contract without the possibility of working overtime or less than the contracted

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 I can define $E(U_{t+1}^{absent})$ which is the expected utility in $t+1$ conditional on having been absent at time t:

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

As can be seen from (3), 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 on having worked in t:

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

hours h.

Finally, I 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, I 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 I 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.

I 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 I 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.

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

I 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 the 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 6, I empirically estimate the effects which I derived above theoretically .

4 Data and Variable Definitions

The data set that I use in the empirical part 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).

For the core analyses, I use data of the years 1994 to 1999. As my goal is to evaluate a reduction in wage compensation levels, I drop non-working respondents and those who are not eligible for long-term sickness compensation (i.e., people who earn less than €400 per month and working students). Furthermore, I drop observations with item-non response and restrict the sample to respondents aged 18 to 65.

4.1 Dependent Variables and Covariates

The SOEP contains various questions about the usage of health services and health insurance. I generate my first dependent dummy variable, *longabs*, 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 in 19XX?"* Since sick pay decreases after six weeks, since 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, I use the following SOEP question: *"How many days were you not able to work in 199X because of illness?"* I generate my second dependent variable by subtracting, for those who had a long-term absence spell, the number of employer-paid sick days – namely 30 for the first six weeks – from the total number of days absent. This variable is called *longabsdays* and measures the duration of long-term absenteeism.[7] [8] Clearly, this duration indicator is subject to measurement errors as I assume that the respondents had no other absence spells. Moreover, comparing the mean value of *longabsdays* 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,

[7] Public sector employees enjoy special privileges. In contrast to private sector employees, they receive 100 percent sick pay up to 26 weeks depending on seniority. Since I have detailed information about the seniority levels, I am able to identify these privileged public sector employees. For them, I redefine long-term absence spells as sickness spells for which they receive the lower SHI statutory sick pay.

[8] 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, I replace the values on *longabsdays* with a one. By estimating a Zero-inflated Negbin-2 model and predicting the total number of benefit days, I impute missing values for respondents with item-non response on the variable about total sick leave days. I impute the values only for respondents who indicated that they were on long-term sick leave and who had no missings on the other covariates.

as long as the cut in statutory 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 should be sufficient to assess the reform effects. While the former assumption is likely to hold, one could argue that the latter is more problematic. Those who were only affected by the cut in statutory long-term sick pay had an increased incentive to substitute long-term spells by short-term spells. However, according to German law, the eligibility for employer-provided statutory short-term sick pay expires in case of such sickness spell substitutions (see Section 2.3 for more details). Once more, the importance of having various treatment groups is emphasized here. By comparing *Treatment Group 1* with the *Control Group*, I cannot unambiguously identify reform effects on the duration of long-term workplace absences, since a negative effect on *longabsdays* might have been triggered by the cut in short *or* long-term sick pay. However, 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 no issue in this setting, I can estimate the impact of the cut in statutory long-term sick pay on the length of long-term sickness spells.

Since both questions on absenteeism, and thus both dependent variables, refer to the last calendar year, I use 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, I use the current values of their covariates and assume that they did not change since the onset of the long-term sick leave episode.

The whole set of explanatory variables can be found in the Appendix. It is categorized as follows: A first group of covariates incorporates variables on personal characteristics, like the dummies *female, 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 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*.

4.2 Treatment Indicators and Treatment Intensity Indices

As described in Section 2.3 and visualized in Table 1, I define different subsamples as *Treatment Group 1*, *Treatment Group 2*, and *Control Group*. Since the SOEP is very detailed about the insurance status and the workplace of the respondents, I can precisely assign respondents to the different 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. Since I am unable to identify respondents with such contracts, I drop them.

An can be seen in Table 2, I generate three treatment dummy indicators that I use below in my empirical models to estimate the direct, indirect, and net effect of the two sick pay reforms on long-term absenteeism. *T1* has a one for all employees who were affected by both reforms (*Treatment Group 1*) and a zero for all those who were affected by none (*Control Group*). I use *T1* to estimate the reforms' net effet on long-term absenteeism. To disentangle the direct effect, I employ *T2* which has a one for all respondents who were solely affected by the cut in statutory long-term sick pay (*Treatment Group 2*) and a zero for the *Control Group*. In contrast, *T3* has a one for *Treatment Group 1* and a zero for *Treatment Group 2*, helping me in assessing the indirect effect.

[Insert Table 2 about here]

As discussed in Section 2.3, not only was statutory long-term sick pay cut from 80 to 70 percent of foregone gross wages but likewise was its benefit cap decreased from 100 to 90 percent of the net wage after taxes and social contributions. Depending on the relation between gross and net wage, this reform element generated an additional source of exogenous variation in terms of treatment intensity. As a result, individuals experienced cuts in their statutory long-term sick pay from zero up to ten percent of their gross wage. Thus, I calculate for each individual his or her (potential) reform induced decrease in statutory long-term sick pay relative to the gross wage. This is feasible since the SOEP samples data on gross wages, net wages, and other income components such as Christmas or vacation bonuses. The SOEP group deals precisely with the problem of item-non response and imputes missing values thoroughly (Frick and Grabka, 2005).

Then, in addition to the three treatment dummy indicators, I generate two continuous treatment intensity indices. Both sample the same individuals as *T1* and *T2* and are called *T1index* and *T2index*. *T1index* has the value 0 for those in the *Control Group* and values from 0.57 up to 10.00 for those in *Treatment Group 1*, meaning that the decrease in statutory long-term sick pay varied between 0.57 and 10 percent of the employees' gross wage. Equivalently built is *T2index*. Everyone in the *Control Group* has a zero on *T2index* and employees in *Treatment Group 2* have positive values up to 10.00. The density of *T1index* and *T2index* peaks around six and ten. About 80 percent of the treated faced a cut in statutory long-term sick pay between 4 and 8 percent of their gross wage and about 12 percent experienced a cut of 10 percent of their gross wage.

5 Estimation Strategy and Identification

5.1 Probit Specification

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

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

where y_{it} stands for the incidence of long-term absenteeism, *longabs*, for individual i in year t. The dummy $p97_t$ has a one for post-reform years and a zero for pre-reform years. Depending of the empirical specification, the treatment variable D_{it} stands representative for *T1*, *T2*, *T3*, *T1index*, or *T2index* (see Section 2.3 and Section 4.2). My variable of interest, DiD_{it}, can be interpreted as the interaction term between D_{it} and $p97_t$ and takes on positive values for treated individuals in post-reform years. By including time dummies ρ_t I control for common time shocks that might affect long-term absenteeism. State dummies ϕ_s account for permanent differences across the 16 German states along with the annual state unemployment rate that controls for changes in the tightness of the regional labor market and that is included in the $K \times 1$ column vector s'_{it}. The other $K - 1$ regressors are made up of personal controls including health status, educational controls, and job-related controls as shown in the Appendix.

Should the assumptions discussed below hold, the marginal effect of the interaction term DiD_{it} gives us the causal reform effect and is henceforth always displayed when output tables are presented.[9]

5.2 Count Data Specification

To estimate how the policy reforms affected the duration of long-term absence spells in post-reform periods, I fit count data models. Since the second dependent variable *longdaysabs* is a count with excess zero observations and overdispersion, i.e., the conditional variance exceeding the conditional mean, count data models should capture these distributional properties appropriately. Based on the Akaike (AIC) and Bayesian (BIC) information criteria as well as on Vuong tests, I

[9] Puhani (2008) has shown that the advice of Ai and Norton (2004) to compute the discrete double difference $\frac{\Delta^2\Phi(.)}{\Delta \text{p97} \Delta \text{D}}$ is not of relevance in nonlinear models when the interest lies in the estimation of a treatment effect in a difference-in-differences model. Using treatment dummy indicators, the average treatment effect on the treated is given by $\frac{\Delta\Phi(.)}{\Delta(\text{p97*D})} = \Phi(\alpha + \beta p97_t + \gamma D_{it} + \delta DiD_{it} + s'_{it}\psi + \rho_t + \phi_s) - \Phi(\alpha + \beta p97_t + \gamma D_{it} + s'_{it}\psi + \rho_t + \phi_s\zeta)$ which is exactly what I calculate and present throughout the paper.

found two model specifications to be well suited.

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. 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.[10] The Negative Binomial (NegBin) Model model is a special case of a continuous mixture model. In the notation of Cameron and Trivedi (2005), the negative binomial distribution can be described as a density mixture of the following form:

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

where $f\left(y|\mu,\nu\right)$ is the conditional Poisson distribution and $\gamma(\nu|\alpha)$ is assumed to be gamma-distributed with ν as an unobserved parameter with variance $\alpha = 1/\delta$. $\Gamma(.)$ denotes the gamma integral and $\mu = exp(\boldsymbol{X\beta})$ where the matrix $\boldsymbol{X}$ incorporates the same variables as the probit

[10] The unobserved heterogeneity allowed for in the NegBin-2 is based on functional form and does not capture unobserved heterogeneity which is correlated with explanatory variables.

model in equation (11). The Negative Binomial Model 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.

5.3 Identification

In every difference-in-differences (DiD) model, the main identification assumption is the common time trend assumption. It assumes that, for both groups – treatment and control group – the trend of the outcome variable would have developed parallely in the absence of the policy intervention. In other words, after having conditioned on all available covariates, unobservables should not have a differential impact on treatment and control group with respect to changes in the dependent variable over time. Depending on the context, this may be a more or less strong assumption. My identification strategy is based on various pillars, making me confident that I am able to identify true causal reform effects.

First, I use three different subsamples that were differently affected by the two reforms. In my empirical specifications, I employ three distinct models, all of which compare these mutually exclusive subsamples to one another. The first two models contrast *Treatment Group 1* as well as *Treatment Group 2* separately to the *Control Group*, and the third model compares *Treatment Group 1* to *Treatment Group 2* (see Section 4.2 and Table 2). By this means, I esimate the net, direct, and indirect effect of the two reforms on long-term absenteeism. Comparing the findings from these three distinct models allows me to cross check the plausibility and coherence of my results.

Second, I 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. Working with survey data makes it possible to take a rich set of background variables into account – at the cost of having no detailed spell data. In Section 4.1, I have discussed why, nevertheless, the available work absence information is sufficient to measure the direct reform effect on the duration of long-term absenteeism. Moreover, I exploit an additional source of exogenous variation which allows me to distinguish effects by treatment intensity (see Section 2.3 for more details): The main replacement level of statutory long-term sick pay was cut along with a decrease in the upper limit of this benefit. Depending on the ratio of net to gross wages, treated employees experienced cuts of between one and ten percent of their gross wage. By using SOEP income information, I am able to calculate the individual reform induced decrease in statutory long-term sick pay remarkably exactly. I use this information in extended analyses that differentiate by treatment intensity.

Third, the implementation of the reform and the variation in the treatment intensity were clearly exogenous to the individuals and politically determined. I 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 to demonstrate strength and capacity to act.

Fourth, as in almost every study that builds upon natural experiments, the three distinct groups that I use as control and treatment groups differ significantly in terms of their 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 regulation of the German health insurance. However, the differences in characteristics are not the result of treatment-related self-selection but politically determined. Moreover, I adjust the sample composition with respect to all of these observed characteristics. Most importantly, I use various measures of the respondents' health status which is clearly the key determinant of long-term absenteeism. Please note 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 unlikely that unobservables have a diverging effect on the dynamic of the outcome – 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. 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**]

Fifth, to prove the consistency of the results, I perform various robustness checks. Thanks to the panel structure of my data, I am able to control for labor force and panel attrition by using balanced panels. Moreover, I experiment with different pre- and post-reform years. Additionally, to assess whether effect heterogeneity plays a role, I restrict the sample 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, I focus on short time horizons. As Bertrand et al. (2004) have shown, one 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. I 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] Besides focusing on short time spans to resolve serial correlation concerns, I use robust standard errors and correct for clustering at the individual level throughout the analysis.

Finally, as an important feature of this study, I can exclude that selection into or out of the treatment drives the results, which is a central issue in other settings, e.g., when labor market programs are evaluated. The reason lies in the institutional setting: Switching between the two diverse health care systems – remember that only employees insured with the SHI were affected by the cut in statutory long-term sick pay – is not allowed for the great majority. I am able to identify the only subsample that has this right to opt out of the SHI and exclude it in my

[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 proposed by Donald and Lang (2007).

robustness checks.[12]

My basic empirical strategy is to pool the data for the years 1995 to 1998 and to estimate various diffference-in-differences models. As explained above, using different subsamples which I compare against each other, I run three main models to estimate the net, the direct, and the indirect effect of the sick pay reforms on long-term absenteeism. In addition, in extended models, I differentiate by treatment intensity. Moreover, I do not only estimate the effects on the incidence of long-term absenteeism but also on the duration of long-term absenteeism.

6 Results

Table 5 provides the unconditional DiD estimates of the reforms' net and direct effects on the incidence of long-term absenteeism. The unconditional long-term absence incidence for *Treatment Group 1* decreased from 6.16 percent in the pre-reform years 1995/1996 to 5.92 percent in the post-reform years 1997/1998. The incidence for *Treatment Group 2* decreased 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 incidence for the *Control Group* also decreased from 3.49 to 3.11 percent in the same time period, resulting in overall difference-in-differences (DiD) 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 long-term sick leave benefit days decreased between the pre- and the post-reform period 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 sets of covariates and measure the reforms' net effect on the incidence of long-term absenteeism. Each specification represents a probit model equivalent to equation (11). The dependent variable *longabs* is 1 if the respondent had a long-term sickness spell and zero otherwise. The variable of interest is displayed

[12] Only employees who are optionally insured with the SHI (self-employed, civil servants, and high-income earners above the income threshold) have the right to opt out of the SHI and to become part of the PHI (see Section 2). However, it is very unlikely that employees opted out of the SHI as a reaction to the cut in statutory long-term sick pay. Opting out is a lifetime decision since switching back to the SHI system is almost impossible. Moreover, the elderly would have to pay extremely high premiums and it makes no sense for the young either, since they are very likely to be unaffected by long-term absenteeism anyway.

as *DiD1* and is one for employees in *Treatment Group 1* in the post-reform period. In every specification, marginal effects are calculated and displayed. In none of the model specifications is the *DiD1* estimate statistically different from zero. The estimated coefficients are very close to zero, 0.0063 in the preferred specification, and positive. The standard error in the preferred specification is 0.0086. Note that the *DiD1* coefficients are robust to the inclusion of sets of covariates and close to the unconditional DiD estimate, which reinforces the plausibility of the common time trend assumption.

[**Insert Table 7 about here**]

In the next step, I 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. A priori, one would expect the sign of the direct effect to be negative since it assesses the impact of the cut in statutory long-term sick pay on long-term absenteeism. The indirect effect stems from the fact that the gap in the replacement levels between statutory short- and long-term sick leave shrank due to the reform, which might have had a positive impact on long-term absenteeism. As has been shown theoretically in Section 3, being able to disentangle these potentially diverging effects is important since it may be that the indirect reform effect compensated the direct effect.

[**Insert Table 8 about here**]

Column (1) in Table 8 once again displays the net effect; the regression model equals Model 6 in Table 7. Column (2) estimates the effect of the cut in statutory long-term sick pay on the incidence of long-term absenteeism, i.e., the direct effect. In contrast to column (1), *Treatment Group 2* – those *only* affected by the cut in statutory long-term sick pay – is contrasted with the *Control Group*. The regressor of interest is now *DiD2*. The *DiD2* estimate is again positive and statistically not different from – but close to – zero. The findings from column (1) and (2) are confirmed in column (3). Here, I compare those who were affected by both reforms to those who were only affected by the cut in statutory long-term sick pay, i.e., *Treatment Group 1* to *Treatment Group 2*. Again, point estimate and standard error are close to zero in magnitude and the indirect reform effect on the incidence of long-term absenteeism is not statistically different from zero.

[**Insert Table 8 about here**]

T1index and *T2index* represent the treatment intensity of the reform, which I define as the cut in statutory long-term sick pay relative to the individual's gross wage (see Section 2.3 and 4.2).

By interacting these continuous variables with the post-refom dummy *p1997*, I estimate the net effect and the direct effect on the incidence of long-term absenteeism in Table 9. As above, I am unable to reject the hypothesis that the reforms have induced any significant behavioral changes, which is illustrated by the *DiD1index* and *DiD2index* coefficients that are very close to zero in size and not significantly different from zero.

[Insert Table 9 about here]

Table 10 uses the number of days that long-term sick leave benefits were received (*longabsdays*) as dependent variable and estimates count data models as explained in Section 5.2. I always focus on the direct effect and differentiate by treatment intensity, i.e., I use *T2index* and its interaction with *p1997* (*DiD2index*). 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).

[Insert Table 10 about here]

6.1 Robustness Checks and Heterogeneity in Effects

Until now my estimation strategy was to pool the data over four years, which means that I 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, I should check whether this selection out of the labor market drives my 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). I 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. Especially if I had found reform effects that suggested a significant reduction in long-term absenteeism, the estimate might have been driven by selection out of the labor market. In the following, I discuss why illness-related selection out of the labor market is no source of serious concern in this setting.

First, 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 refreshment sample, 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).[13]

Second, the availability of a control group allows me to control for treatment-independent selection out of the labor market.[14] 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, since I use panel data, in addition to correcting the sample composition by observables, I use a balanced sample in one of the robustness checks shown in Table 11 and 12.

Table 11 and 12 report results for the *direct* effect specification on the incidence and duration of long-term absenteeism. Both tables use *T2index* and *DiD2index*, meaning that I always differentiate by treatment intensity. As a first test, I center the data two years around the reform (column (1)). Afterwards, I restrict my 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 who answered the SOEP questionnaire in both years (column (2)). An alternative robustness check would be to take 1995 as reference year and contrast it with 1997 and 1998. It might have been the case that anticipation effects played a role and that employees already adapted their behavior in 1996, when the reform plans were made public (column (3)). This is, however, not very probable as many catalysts of long-term absences, like cancer diagnosis, 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, I contrast the years 1995/1996 with 1998 in column (4).

To test effect heterogeneity, I 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 had an impact on the exposure to treatment. On the household level, the relevant parameter might be the decrease in total household income rather than in individual wages. Since optionally SHI insured could have switched to the PHI system as a reaction to the reform, I exclude all optionally insured people in column (7). I also split the sample at the median gross wage (columns (8) and

[13] For the other employment groups like the part-time employed, the deviation between 1996 and 1997 was less than 1.6 percent.

[14] I cannot, however, entirely exclude the possibility that the reform had an effect on the decision to leave the labor market voluntarily. I am 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, reform-induced selection out of the labor market is likely to play a negligible role.

(9)).

Table 11 shows the results when I use the incidence of long-term sick leave, *longabs*, as dependent variable. None of the *DiD2index* coefficients is statistically different from zero but all are very close to zero in magnitude, which reinforces my main findings above. Note that although all coefficients are practically zero, they all have positive signs.

[Insert Table 11 and 12 about here]

In Table 12, where I use the number of long-term sick leave benefit days (*longdaysabs*) as dependent variable, I do not find significant reform effects for most of the specifications either. The coefficients are close to zero in magnitude, and columns (2) and (6) even have positive signs. However, I find significantly negative reform effects for middle-aged full-time employed and the poor (columns (5) and (8)), which suggests heterogeneity in the reform effects on benefit duration. According to the estimates, a one unit increase in *T1index*, which equals an increase in the absence costs of about 5 percent, led to a decrease in the average number of long-term sick leave benefit days of around 0.04 and 0.11, respectively. Middle-aged full-time employed people most likely need to feed 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. Besides the notion that these subsamples have reacted to monetary incentives, another explanation is possible: Although *Treatment Group 2*, which I use in these specifications, was solely affected by the cut in statutory long-term sick pay, there might have been spillover effects from the cut in statutory short-term sick pay. Since Puhani and Sonderhof (2010) has shown that the cut in statutory short-term sick pay clearly induced reductions in short-term sick pay, it is at least imaginable that public sector employees and trainees insured with the SHI were not fully aware of their priviledges. If the cut in statutory short-term sick pay reduced short-term sickness spells that these groups might have had *in addition* to their long-term spell, my estimates here would be contaminated. Moreover, it might have been the case that spillover effects were induced if employees in *Treatment Group 1* had partners working in the private sector who reacted to the cut in short-term sick pay.

[Insert Table 13 about here]

A conventional 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 1995 and 1996. All placebo estimates turn out to be insignificant.

6.2 Calculation of SHI Reform Savings

Statutory long-term sick pay amounted to 80 percent of the monthly gross wage before the reform and was reduced to 70 percent after the reform. The benefit cap decreased from 100 percent of the monthly net wage to 90 percent of the wage after taxes and social contributions. I calculate the total price adjusted SHI reform savings from 1997 to 2006, reflecting the redistributional effect of the reform. Reducing the sick pay level for the long-term sick benefited the rest of the statutory health insurance pool through lower contribution rates.

As a first estimate, I calculate statutory long-term sick pay according to the old and the new regulations for every eligible individual and the years 1997 to 2006, take the difference, and sum over the frequency-weighted number of long-term absences for the whole period.

Through the reform, statutory 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 long-term sick leave benefit days equals about 2.5 months, this translates into a benefit cut of about € 100 per month. The decrease represents about seven percent of the average monthly net wage.

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

Consequently, I make use of administrative data from the German Ministry of Health on the total number of SHI long-term sick pay cases and the average number of long-term sick leave benefit days for SHI insured. Unfortunately, no personal data and no income information are collected by official statistics. Hence, I combine administrative data with the SOEP data set, which contains very detailed income information. By this means, I estimate that, between 1997 and 2006, the total sum that the SHI saved due to the reform amounted to around € 5.5 billion.[15] Considering social contributions, this translates into an accumulated net loss for the long-term sick of about five billion euros during that period of time.

[15] In the working paper version, I present a more detailed analysis of the redistributional effects. Various specifications and different scenarios are discussed.

7 Discussion and Conclusion

This study explicitly analyzes how cuts in statutory long-term sick leave affect long-term absenteeism in the context of the European statutory sickness absence insurance. In the first part of the paper, by means of a simple dynamic model of absence behavior, I analyze the different incentive effects on long-term absenteeism that were triggered by two cuts in the German statutory sick pay scheme. However, under the assumption that employees on long-term sick leave are seriously sick, the incentive structure of the sick pay scheme breaks down and employees would not react to monetary incentives.

In the second part of the paper, I use SOEP panel data to empirically estimate the reform effects on long-term absenteeism. This is feasible by means of conventional difference-in-differences models since the cut in stautory long-term sick pay applied universally to all employees insured with the public health insurance, but not to respondents insured with the private health insurance. In Germany, the two health care systems co-exist independently. Since switching between the two systems is almost impossible due to federal regulations, I can exclude that treatment-related selection drives my results. Moreover, the reform was clearly exogenous to the individual and a fairly random instrument of the ruling administration to cut health expenditures and to demonstrate capacity to act.

I run three distinct main difference-in-differences models that all contrast mutually exclusive subsamples, which were affected differently by the reforms, with one another. Moreover, I do not only estimate the effects on the incidence of long-term absenteeism but also on the number of long-term sick leave benefit days. In addition, I am able to differentiate by treatment intensity since one element of the reform induced additional exogenous variation such that employees were not affected equally by the reform.

The consistency of the findings from this variety of approaches, together with the results from various robustness checks, makes me confident of having identified true causal reform effects which are not driven by diverging time trends or selection effects. All empirical models consistently suggest that the incidence of long-term absence spells was not affected by the cut in statutory long-term sick pay. All effects on the incidence are close to zero in magnitude and even have positive signs. This suggests that not only imprecision in the estimates leads to the conclusion that employees have not adapted their long-term sick leave behavior. As for the effects on the duration of long-term absenteeism, I also find mostly insignificant reform effects but the sign of the effects is negative, as expected. However, for two subsamples – the poorer half of the sample as well as middle-aged employees working full-time – I find that the cut in statutory long-term sick pay reduced the length of these spell significantly. This suggests heterogeneity in the reform effects on the number of benefit days.

My empirical results are in line with the findings from Campolieti (2004) who convincingly showed that benefit recipients of the Candadian disability insurance (DI) have not adapted their labor supply behavior as a reaction to changes in benefits. The results are also partly in line with the findings from Curington (1994), who used U.S. data from the 1970s on the workers' compensation insurance (WCI). However, the European statutory sickness absence insurance is not directly comparable to the North American DI and WCI.

I have several explanations for my main finding that the long-term sick have not significantly adapted their sick leave behavior to benefit changes: First, the result is in line with my 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). The causes for long-term absenteeism differ substantially from those for short-term abenteeism. Short-term sick leave is mostly determined by flus and light illnesses which clearly leave more space for moral hazard, especially when physicians' certificates are not required during the first days of a spell.

Second, the stringency of the sick leave monitoring and screening process is a potentially important determinant of labor supply reactions and moral hazard. In Germany, certification requirements increase with the length of spells. After six weeks of continuous sick leave, physicians need to issue different certificates in regular time intervals. Moreover, as discussed in Section 2.2, German social legislation explicitly requires the Medical Service of the Statutory Health Insurance (SHI) to take measures that prevent long-term absenteeism and the risk of patients descending the social ladder through long-term illness. Likewise, both employers and sickness funds have clear incentives to avoid unnecessary and overlong sickness episodes. They are encouraged by law to cooperate with the Medical Service which employed about 2,000 independent physicians and examined 1.7 million cases of absenteeism in 2007 (Medizinischer Dienst der Krankenversicherung, 2008).

Third, relative to short absence periods, the incentive structure of the German statutory sick leave scheme makes long absence periods unattractive. The replacement level does not increase with the duration of a spell – as in other European countries like Spain, Czech Republic, or Portugal – but decreases. As has also been shown theoretically in Section 3, decreasing benefits yields an incentive to accumulate shirking behavior in the lower end of the sickness spell distribution rather than in the upper end.

Finally, given that sickness episodes of more than six weeks are typically triggered by serious sickness, the cut in long-term sick pay may have been too moderate to induce changes in the labor supply behavior. My calculations suggest that, on average, the long-term sick have lost €250 per spell or €100 per month – the latter figure represents seven percent of the monthly net wage.

By combining SOEP income data with administrative data, I estimate that the cut in statutory long-term sick pay redistributed five billion Euros from the long-term sick to the SHI insurance pool in order to achieve lower contribution rates. This was the reform's main objective: cutting health expenditures in order to achieve lower contribution rates and to stimulate job creation.

Various pieces of evidence throughout this study suggest that moral hazard is of minor importance when sickness spells of more than six weeks are considered. Consequently, health reforms like the German one do not lead to more efficient sickness insurance markets by decreasing the degree of moral hazard but are merely an instrument to cut health expenditures. On the other hand, when cuts in replacement levels are moderate, this cost containment instrument seems to be economically efficient in the sense that it induces no major behavioral reactions that might lead to undesirable equilibria. Policy makers should be aware of the redistributional consequences. It is simply a normative question whether such an instrument to cut health expenditures should be applied.

The U.S. and Canada have no social insurance comparable to the European statutory sickness insurance. However, the number of DI recipients is growing steadily in these countries, imposing large economic and social costs since DI recipients usually completely withdraw from the labor market. Moreover, various studies suggest that moral hazard plays a crucial role in the U.S. DI insurance program. The main idea of the European statutory sickness insurance is to provide relatively generous benefits even for work-*un*related sickness while keeping employees employed, together with a stringent monitoring system. This might be a more promising approach to maintain employees' working capacity in the long run. Further research on this topic is needed.

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.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *American Economic Review 79*(3), 482–503.

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.

Campolieti, M. (2004). Disability insurance benefits and labor supply: Some additional evidence. *Journal of Labor Economics 22*(4), 863–890.

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.

Chen, S. and W. van der Klaauw (2008). The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics 142*(2), 757–784.

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.

de Jong, P., M. Lindeboom, and B. van der Klaauw (2010). Screening disability insurance applications. *Journal of the European Economic Association forthcoming.*

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.

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.

Gruber, J. (2000). Disability insurance benefits and labor supply. *Journal of Political Economy 108*(6), 1162–1183.

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.

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*(9-10).

Medizinischer Dienst der Krankenversicherung (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*.

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.

Puhani, P. A. and K. Sonderhof (2010). The effects of a sick pay reform on absence and on health-related outcomes. *The Journal of Health Economics 29*(2), 285–302.

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.

Ziebarth, N. R. and M. Karlsson (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics 94*(11-12), 1108–1122.

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

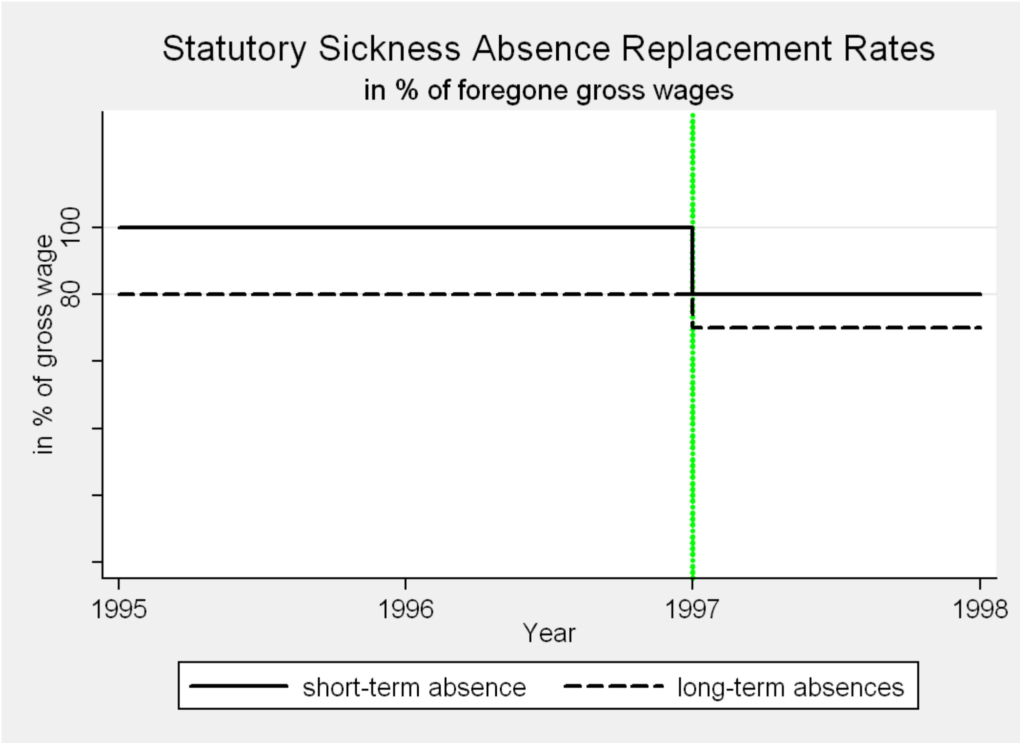

Table 1: Definition of Subsamples

	Cut statutory short-term sick pay $<$ 6 weeks (employer)	**Cut statutory long-term sick pay $>$ 6 weeks (SHI)**
Private sector employees with SHI (1)	yes	yes
(*Treatment Group 1*)		
Public sector employees with SHI (2)	no	yes
Trainees with SHI (3)	no	yes
(*Treatment Group 2*)		
Public sector employees with PHI (4)	no	no
Self-employed with PHI (5)	no	no
(*Control Group*)		

Table 2: Definition of Treatment Indicators to Estimate Reform Effects

Effect to be estimated	**Treatment Indicator**	**=1**	**=0**
Net effect	T1	subsample (1) (*Treatment Group 1*)	subsamples (4) + (5) (*Control Group*)
Direct effect	T2	subsamples (2) + (3) (*Treatment Group 2*)	subsamples (4) + (5) (*Control Group*)
Indirect effect	T3	subsample (1) (*Treatment Group 1*)	subsamples (2) + (3) (*Treatment Group 2*)

Table 3: Variable Means by Treatment and Control Groups

Variable	Control Group	Treatment Group 1	Treatment Group 2	Min	Max
Incidence of long-term absenteeism (*longabs*)	0.033	0.060	0.026	0	1
Duration of long-term absenteeism (*longabsdays*)	1.965	3.392	2.249	0	365
Personal characteristics					
Female	0.410	0.366	0.587	0	1
Age	40.57	39.86	37.48	18	65
Age square/100	17.58	17.01	15.60	3.24	42.25
Immigrant	0.097	0.215	0.112	0	1
East Germany	0.166	0.258	0.378	0	1
Partner	0.762	0.803	0.650	0	1
Married	0.673	0.696	0.569	0	1
Children	0.483	0.470	0.435	0	1
Disabled	0.033	0.052	0.053	0	1
Good health	0.648	0.607	0.604	0	1
Bad health	0.080	0.099	0.104	0	1
No sports	0.287	0.409	0.331	0	1
Educational characteristics					
Dropout	0.021	0.050	0.044	0	1
Certificate after 8 years of schooling	0.230	0.357	0.271	0	1
Certificate after 10 years of schooling	0.290	0.330	0.438	0	1
Certificate after 12 years of schooling	0.051	0.035	0.035	0	1
Certificate after 13 years of schooling	0.363	0.115	0.162	0	1
Other degree	0.046	0.112	0.051	0	1
Work in job trained for	0.608	0.545	0.511	0	1
New job	0.204	0.179	0.179	0	1
No. of years in company	10.29	9.04	8.79	0	47.9
Job characteristics					
No tenure	0.106	0.051	0.273	0	1
One-man company	0.099	0.000	0.000	0	1
Small company	0.327	0.274	0.169	0	1
Medium-sized company	0.179	0.312	0.281	0	1
Large company	0.126	0.221	0.290	0	1
Very large company	0.268	0.193	0.260	0	1
Self employed	0.308	0.000	0.000	0	1
Blue collar worker	0.112	0.528	0.190	0	1
White collar worker	0.150	0.472	0.579	0	1
Public sector	0.493	0.000	0.829	0	1
Civil servant	0.395	0.000	0.031	0	1
Self employed	0.307	0.000	0.000	0	1
High job autonomy	0.506	0.160	0.152	0	1
Gross income per month	2,383	2,013	1,675	204	40903
Regional unemployment rate	11.49	12.04	13.07	7	21.7
N	2,693	16,006	6,500		

Table 4: Probit Model: Determinants of the Incidence 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		
Certificate after 8 years' of schooling (d)	-0.006	0.006
Certificate after 10 years' of schooling (d)	-0.008	0.007
Certificate after 12 years' of schooling (d)	-0.018***	0.007
Certificate after 13 years' of schooling (d)	-0.013**	0.006
Other certificate (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-sized company (d)	0.0012***	0.004
Large company (d)	0.015***	0.004
Very large 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	

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

Table 5: Unconditional DiD Estimates 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 (*longabs*) is displayed.
Standard errors in parentheses.

Table 6: Unconditional DiD Estimates on the Average Number of Long-Term Sick Leave 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 (*longabsdays*) is displayed.
Standard errors in parentheses.

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

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
DiD1	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	-0.0012	-0.0123	-0.0133	-0.0102	-0.0117	-0.0102
(p1997)	(0.0124)	(0.0140)	(0.0140)	(0.0135)	(0.0127)	(0.0123)
Year 1996	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	-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	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

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

Standard errors in parentheses are adjusted for clustering on person identifiers.

Marginal effects are displayed and calculated at the means of the covariates except for *Treatment Group 1* (=1), *p1997* (=1), *Year 1996* (=0), and *Year 1997* (=1).

Dependent variable: dummy that is 1 if respondent had long-term absence spell (*longabs*).

Every column represents one probit model as in equation 11.

DiD1 is the DiD indicator. It has a one for respondents in *Treatment Group 1* in post-reform years. *DiD1* estimates the net reform effect.

Table 8: DiD Estimation on Incidence: Disentangling the Direct from the Indirect Reform Effect

Variable	**Net effect**	**Direct effect**	**Indirect effect**
DiD1	0.006		
	(0.009)		
T1	0.012**		
	(0.006)		
DiD2		0.010	
		(0.010)	
T2		-0.015	
		(0.012)	
DiD3			-0.000
			(0.004)
T3			-0.021***
			(0.006)
Post-reform dummy	-0.010	0.007	-0.000
(p1997)	(0.012)	(0.012)	(0.004)
Year 1996	0.000	0.016*	0.002
	(0.005)	(0.009)	(0.003)
Year 1997	-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

* p<0.1, ** p<0.05, *** p<0.01
Standard errors in parentheses are adjusted for clustering on person identifiers.
Marginal effects are displayed and calculated at the means of the covariates except for *T1 (2, 3)* (=1), *p1997* (=1), *Year 1996* (=0), and *Year 1997* (=1).
Dependent variable: dummy that is 1 if respondent had long-term absence spell (*longabs*).
Every column represents one probit model as in equation 11.
T1 is one for respondents who were affected by both cuts, in statutory short- *and* long-term sick pay (*Treatment Group 1*), and is zero for those who were not affected at all by the reforms (*Control Group*). *T2* contrasts those only affected by the cut in statutory long-term sick pay (*Treatment Group 2*) to the *Control Group*, and *T2* compares the incidence of *Treatment Group 1* with the incidence of *Treatment Group 2*. *DiD1 (DiD2, DiD3)* is the DiD indicator and has a one for *T1 (T2, T3)* =1 and post-reform years. *DiD1 (DiD2, DiD3)* estimates the net (direct, indirect) reform effect. More information about the treatment indicators can be found in Section 4.2 and in the Appendix.

Table 9: DiD Estimation on Incidence with Varying Treatment Intensity

Variable	**Net effect**	**Direct effect**
DiD1index	0.000	
	(0.001)	
T1index	0.003***	
	(0.001)	
DiD2index		0.000
		(0.001)
T2index		0.000
		(0.002)
Post-reform dummy	-0.005	0.011
(p1997)	(0.010)	(0.012)
Year 1996	0.000	0.016*
	(0.005)	(0.009)
Year 1997	-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

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

Standard errors in parentheses are adjusted for clustering on person identifiers.
Marginal effects are displayed and calculated at the means of the covariates except for *p1997* (=1), *Year 1996* (=0), and *Year 1997* (=1).
Every column represents one probit model as in equation 11.
Dependent variable: dummy that is 1 if respondent had long-term absence spell (*longabs*). *T1index (T2index)* is the treatment intensity index for the same subsamples as *T1 (T2)*. It takes on positive values on a continuous scale up to 10.00 for *Treatment Group 1* and is zero for the *Control Group*. *DiD1index (DiD2index)* is the DiD intensity index and has positive values for *Treatment Group 1 (2)* and post-reform years. *DiD1index (DiD2index)* estimates the net (direct) reform effect.
More information about the treatment intensity indices can be found in Section 4.2 and in the Appendix.

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**
DiD2index	-0.041	-0.904
	(0.058)	(1.915)
T2index	0.043	1.188
	(0.044)	(1.006)
Post-reform dummy	-0.402	-16.524
(p1997)	(0.642)	(24.307)
Year 1996	-0.064	1.509
	(0.275)	(10.047)
Year 1997	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

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

Standard errors in parentheses are adjusted for clustering on person identifiers.

Marginal effects are displayed and calculated at the means of the covariates except for *p97* (=1), *Year 1996* (=0), and *Year 1997* (=1).

Every column represents one count data model as in equation 12.

Dependent variable: Number of long-term sick leave benefit days (*longabsdays*).

T2index is the treatment intensity index for the same subsample as *T2*.

It takes on positive values on a continuous scale up to 10.00 for *Treatment Group 2* and is zero for the *Control Group*. *DiD2index* is the DiD intensity index and has positive values for *Treatment Group 2* and post-reform years.

DiD2index estimates the direct reform effect. More information about the treatment intensity indices can be found in Section 4.2 and in the Appendix.

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**
DiD2index	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

* $p<0.1$, ** $p<0.05$, *** $p<0.01$
Standard errors in parentheses are adjusted for clustering on person identifiers.
Marginal effects are calculated at the means of the covariates except for *p1997* (=1), *Year 1996* (=0), and *Year 1997* (=1).
Every column represents one probit model as in equation 11.
Dependent variable: dummy that is 1 if respondent had long-term absence spell (*longabs*).
DiD2index is the DiD intensity index and has positive values for *Treatment Group 2* and post-reform years. It is zero for respondents in the *Control Group*.
DiD2index estimates the direct reform effect.
More information about the treatment intensity indices can be found in Section 4.2 and in the Appendix.

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**
DiD2index	-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

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

Standard errors in parentheses are adjusted for clustering on person identifiers.

Marginal effects are calculated at the means of the covariates except for *p1997* (=1), *Year 1996* (=0), and *Year 1997* (=1).

Every column represents one Zero-Inflated NegBin-2 model as in equation 12.

Dependent variable: number of long-term sick leave benefit days (*longabsdays*).

DiD2index is the DiD intensity index and has positive values for *Treatment Group 2* and post-reform years. It is zero for respondents in the *Control Group*.

DiD2index estimates the direct reform effect.

More information about the treatment intensity indices can be found in Section 4.2 and in the Appendix.

Table 13: Placebo Estimates Using Treatment Index 2

Variable	**Direct effect (Incidence)**	**Direct effect (Duration)**
DiD2index96	0.001	-0.042
	(0.003)	(0.159)
DiD2index95	-0.003	-0.171
	(0.005)	(0.277)
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

* p<0.1, ** p<0.05, *** p<0.01
Standard errors in parentheses are adjusted for clustering on person identifiers.
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).
Column (1) estimates one probit model as in equation 11 and column (2) estimates one Zero-Inflated NegBin-2 model as in equation 12.
Dependent variable in column (1): incidence of long-term absenteeism (*longabs*).
Dependent variable in column (2): number of long-term benefit days (*longabsdays*).
DiD2index96 (95) is the DiD intensity index for a pseudo-reform in 1996 (1995) and has positive values for pseudo-*Treatment Group 2* and pseudo-post-reform years. It is zero for respondents in the pseudo-*Control Group*.
DiD2index96 (95) estimates the pseudo direct reform effect. More information about the treatment intensity indices can be found in Section 4.2 and in the Appendix.

Appendix

Table 14: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Longabs	0.049	0.215	0	1	25199
Longabsdays	2.967	19.449	0	335	25199
T1	0.856	0.351	0	1	18699
T2	0.707	0.455	0	1	9193
T3	0.289	0.453	0	1	22506
T1index	5.699	2.755	0	10	18699
T2index	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
Certificate after 8 years' of schooling	0.321	0.467	0	1	25199
Certificate after 10 years' of schooling	0.354	0.478	0	1	25199
Certificate after 12 years' of schooling	0.037	0.188	0	1	25199
Certificate after 13 years' of schooling	0.154	0.361	0	1	25199
Other certificate	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
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 company	0.253	0.435	0	1	25199
Medium-sized company	0.289	0.454	0	1	25199
Large company	0.229	0.42	0	1	25199
Very large 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