Forward-looking moral hazard in the social insurance system: evidence from a natural experiment*

Marcus Eliason[a], Per Johansson[a,b,c], Martin Nilsson[a,d]

[a]*Institute for Labour Market Policy Evaluation (IFAU), Box 513, SE–751 20 Uppsala, Sweden*

[b]*Department of Statistics, Uppsala University, Box 513, SE–751 20 Uppsala, Sweden*

[c]*The Institute for the Study of Labor IZA, Bonn, Germany*

[d]*Department of Economics, Uppsala University, Box 513, SE–751 20 Uppsala, Sweden*

Abstract

This study tests for forward-looking moral hazard in the social insurance system by exploiting a reform as a natural experiment. The replacement rate was reduced for short but not for long absences, which introduced a potential future cost of returning to work. Using this exogenous variation in the replacement rate and controlling for dynamic selection, we found that the potential future cost decreased the outflow by 13 percent and prolonged long absences by 10 days. This suggests that temporarily disabled people are forward-looking and highlights the importance of taking forward-looking behavior into account when designing and evaluating social insurance programs.

Keywords: Sickness absence; Temporary disability; Moral hazard; Forward-looking agents; Sickness Insurance, Natural Experiment

JEL classification: H55; I12; J22

1 Introduction

The large number of people being absent from work due to sickness or disability has been regarded as "Sweden's single biggest economic problem" by the OECD. This figure can be broken down into the number of people being temporarily absent due to sickness or disability in a given point in time, the length of such spells, and the number of people who has permanently left the labor force through disability retirement. In an international comparison Sweden has fallen out poorly with respect to all three components. For example, an overview by OECD (2009) showed that in 2007, Sweden had the second highest number of lost working days due to sickness and disability among the OECD countries: more than 25 days per employee and year, as compared to 9 days for the U.S.. However, during the last decades several changes have been introduced to combat the high sickness and disability rates, mostly by changing economic incentives and reducing moral hazard: the benefit levels has changed rather frequently, with large cuts in the early 1990s, time limits for benefit receipt have been introduced,[1] the

*Corresponding author: Per Johansson, e-mail: per.johansson@statistik.uu.se

[1]However, the time limits for benefit receipt introduced in 2008 were abolished in 2016.

screening of new applicants has become more stringent, and the monitoring of absentees more careful. While moral hazard is an inevitable problem in all social insurance schemes, the degree varies between different insurance programs and also within the same type of insurance program depending on the particular design. Hence, the policymaker faces the problem of balancing the benefits of risk-sharing, allowing for intertemporal consumption smoothing, associated with a generous insurance program with the disincentive, or moral hazard, effects of the same program.

In this study, we test for the presence of "forward-looking moral hazard" in the Swedish social insurance system, by using a reform, that introduced a potential future cost of returning to work following longer sickness absences, as a natural experiment. The reform reduced the replacement level from 90 percent to 65 percent of foregone earnings during the first three days of absence, to 80 percent during days 4–90, while leaving it unchanged at 90 percent for longer absences. Hence, for absences longer than 90 days there were no direct costs associated with the reform, since the replacement rate was unchanged for such absences, but an indirect cost was introduced by the lower replacement rate in case of a relapse that required a new absence after having returned to work. A rational forward-looking individual would take into consideration not only any direct gains from leaving a benefit state but also potential future costs of having to re-enter the same state. If the replacement rate is high enough, such forward-looking behavior can create a "locking-in-effect", i.e., that the individual remains in the benefit state longer than necessary.

The problem of forward-looking moral hazard is more likely to be important in social insurance programs where: (i) the eligibility criteria is difficult to verify, (ii) the benefits do not diminish with the duration in the state, and/or (iii) there is no time limits for benefit receipt. Contrary to the unemployment insurance (UI), this situation often prevails in sickness insurance (SI), disability insurance (DI), and workers' compensation insurance (WCI) programs (cf., Krueger and Meyer, 2002). There is vast empirical evidence of quite substantial effects of changes in the replacement rate on the transition rate (out of insured unemployment) in the UI system, while corresponding findings for the DI, WCI and SI systems provide a less coherent picture (see, e.g., Fevang et al., 2017; Krueger and Meyer, 2002). Given that the populations within the various programs differ in many respects, and especially in terms of health, these seemingly contradictory findings may not be surprising. However, another potential explanation is the more frequent occurrence of benefit receipt without time limits, and potential costs of re-entering the benefit state, in the latter programs.[2] That is, dynamic incentives, or forward-looking moral hazard, might be more important aspects in the DI, WC, and SI programs than in the UI program.

The present study contributes to two different fields. First, it contributes to the literature on on the presence of moral hazard in the social insurance system, more specifically the importance of economic incentives for absence behaviour. Most of the evidence on DI and WCI programs comes from the U.S. and Canada and the findings are somewhat inconclusive. While several studies do find that increased benefit levels are associated with reduced labour supply (e.g., Curington, 1994; Meyer et al., 1995; Gruber, 2000; Neuhauser and Raphael, 2004), there are also studies that have found no, or only modest, effects (e.g., Campolieti, 2004; Chen and van der Klaauw, 2008).[3] The evidence on SI programs instead comes from the European

[2]The costs of re-entering the benefit state can be in terms of, for example, a lower replacement rate or a lengthy application process.

[3]Two notable exceptions to the studies on North American data found significant return-to-work, from the Norwegian temporary disability insurance (TDI) and DI program, induced by financial incentives (Kostøl and Mogstad, 2014; Fevang et al., 2017).

countries. On the one hand, these studies consistently show that, in accordance with economic theory, cuts in the replacement rate for short-term sickness absence increase the outflow from (or decrease the inflow to) the benefit state (e.g., Johansson and Palme, 2002, 2005; Ziebarth and Karlsson, 2010; Pettersson-Lidbom and Thoursie, 2013; De Paola et al., 2014; Ziebarth and Karlsson, 2014; Aaviksoo and Kiivet, 2016). On the other hand, the evidence on the importance of economic incentives for the outflow from long-term sickness absence or temporary disability is much weaker. Some have found an increased outflow due to lower replacement rates (e.g., Markussen et al., 2011; Puhani and Sonderhof, 2010),[4] while others have found negligible effects (e.g., Ziebarth, 2013; Aaviksoo and Kiivet, 2016).[5] Yet others have found that the replacement rate for short-term absences is positively related to the outflow from longer-term absences (e.g., Johansson and Palme, 2005; Pettersson-Lidbom and Thoursie, 2013; De Paola et al., 2014; Pollak, 2017).[6] [7]

Most closely related to the present study is Johansson and Palme (2005). Using a survey sample of 1396 blue collar workers they investigated the impact of the same reform on incidence and duration of absence (both short- and long-term absence). For shorter absences than seven days they found that the reduced replacement rate increased the hazard of leaving the benefit state, while for absences longer than seven days there was a reverse response. These potential effects were not statistical significant when they controlled for the compositional changes caused by the reform (see, Johansson and Palme, 2004) nor could they account for the probable dynamic selection into absences longer than three days.[8] That is, since the reform both reduced the inflow to sickness absence and increased the outflow from shorter absences, those with longer absences are likely to be more negatively selected in terms of health than before the reform. We are able to circumvent both problems of compositional changes and dynamic selection, by limiting the analysis to a cohort of individuals with absence periods reaching 90 days no later than the day before the reform, and exploit the fact that the new insurance scheme did not apply to ongoing absence periods. Hence, we are the first to study how increased potential future costs of returning to work causally affects absence behaviour among long-term sickness absentees (or temporarily disabled individuals) using exogenous variation in the replacement rate and simultaneously avoiding bias due to dynamic selection.

Second, the present study also contributes to a rapidly growing literature on the empirical testing for forward-looking behavior. This research spans over different fields such as the demand for college textbooks (Chevalier and Goolsbee, 2009), cigarettes and alcohol (e.g., Gruber

[4]Markussen et al. (2011) found a “dramatic” increase in return-to-work when approaching the one-year limit in the Norwegian SI program, where absentees are transferred from a sickness benefit, with a replacement rate of 100 percent, to a rehabilitation benefit with a replacement rate of 66 percent. Puhani and Sonderhof (2010) reported that the reduced absence following a cut in sick-pay (from 100 to 80 percent) in Germany mainly was the result of longer absences becoming shorter.

[5]Ziebarth (2013) found that cuts in sick-pay (from 80 to 70 percent) for long-term absence in Germany neither affected the incidence nor the duration of long-term sickness absence. Aaviksoo and Kiivet (2016) reported that cuts in sick-pay (from 80 to 70 percent), together with an extension of the waiting period from one to three days, in Estonia had negligible impact on longer-term sickness absence (but large effects on short-term absence).

[6]Some of these studies are on the impact of waiting periods, but a waiting period can be viewed as a replacement rate of zero for very short absences.

[7]De Paola et al. (2014) found that cuts in sick-pay (from 100 to 80–90 percent) for short-term absence, together with stricter monitoring, in Italy, increased the duration of longer (more than 10 days) absences. Pollak (2017) found that the workers who were compensated by supplementary sick pay during the three-day waiting period in France had shorter absence periods, and Pettersson-Lidbom and Thoursie (2013) found that the abolishment of a waiting period of one day and an increase in the benefit levels for sickness absences shorter than 14 days increased the outflow among long absences.

[8]Neither do Johansson and Palme (2002); Pettersson-Lidbom and Thoursie (2013); De Paola et al. (2014); Pollak (2017).

and Köszegi, 2001; Tiezzi, 2005; Pierani and Tiezzi, 2011), and medical care/drugs (e.g., Long et al., 1998; Dalton et al., 2015; Abaluck et al., 2015; Einav et al., 2015; Alpert, 2016; Kaplan and Zhang, 2017). The studies on the consumption of alcohol and cigarettes are generally supportive of consumers being forward-looking (e.g., Gruber and Köszegi, 2001; Tiezzi, 2005; Pierani and Tiezzi, 2011) rather than myopic as some models of addiction suggest. Chevalier and Goolsbee (2009) also found that students are forward-looking in their demand for college textbooks. However, the studies on the utilization of medical care, mostly utilizing the dynamic pricing incentives in the Medicare Part D, instead report conflicting results. While several studies have found support for forward-looking behavior in drug demand (e.g., Einav et al., 2015; Alpert, 2016; Kaplan and Zhang, 2017), others report results that instead are supportive of myopia (e.g., Long et al., 1998; Dalton et al., 2015; Abaluck et al., 2015).

Most closely related to the present study is Autor et al. (2014) that investigated dynamic incentives in a private long-term DI program in the U.S.. Autor et al. (2014) reported that a 90-day reduction of the elimination (or waiting) period, from 180 to 90 days, approximately doubled the incidence of accessions and that disability spells within plans with longer elimination periods had substantially longer durations. Because longer elimination periods will reduce the average daily replacement rate more for shorter, than for longer, disability periods, the former are more likely to be deterred. Hence, workers seems to account for the expected duration of their disability in their decision whether to seek benefits for impairments or not. Albeit this suggest that the workers are forward-looking, the considerably larger economic costs associated with a 90 day elimination period, compared to the reduced replacement rate in the present study, implies that binding liquidity constraints is a much more likely alternative explanation.[9]

By exploiting the policy reform described above as a natural experiment we can add to both strands of the literature. We show, using a linear difference-in-difference model and a two-period Cox PH model, that the the expected cost of returning to work decreased the transition back to work by 13 percent and prolonged long absences by 10 days among the temporarily disabled (i.e., those with at least 90 days of absence). This suggests that forward-looking moral hazard, indeed, is present in the Swedish social insurance system. A placebo test also supports our claim of a causal interpretation. Moreover, while most previous studies have found that men react more strongly to changes in replacement levels than women do (e.g., Henrekson and Persson, 2004; Johansson and Palme, 1996; Ziebarth and Karlsson, 2014), we do not. A potential explanation is that women are more risk averse than are men (e.g., Byrnes et al., 1999; Croson and Gneezy, 2009; Eckel and Grossman, 2002, 2008). In previous studies of cuts in the replacement rates the focus has been on the effects of changed direct costs (i.e., the replacement rate for the current absence period), which are not associated with any uncertainty. The indirect cost (i.e., the potential future cost associated with relapse that requires a new absence period with a lower replacement rate), on the contrary, is realized if, and only if, one has to begin a new absence period after having returned to work. Hence, one would expect the response to be stronger among more risk averse individuals. A stronger responsiveness to economic incentives among men might therefore be outweighed by a greater risk averseness among women. Furthermore, we also found that those with more absence periods in the past seem to have responded more strongly to the reform. This finding should be expected, given that they were forward-looking and that the perceived risk of relapse increases with the number of past absence periods .

The rest of the paper is structured as follows. In the next section we provide an overview of

[9] It should be noted, however, that Autor et al. (2014) performed a test using monthly earnings as a proxy for liquidity constraints. The test neither pointed to binding liquidity constraints being the main explanation nor could it reject the hypothesis.

the Swedish institutions, and discuss the reform and its predicted impact on absence behavior. In Section 3 we describe the data and the empirical strategy. The results are presented in Section 4, and finally Section 5 concludes.

2 Institutions, the reform, and theoretical predictions

2.1 The Swedish social insurance system

All workers (employed and unemployed) are covered by the public SI and DI programs administrated by the Swedish Social Insurance Agency (SIA). Until July 2008, there was no formal time limit for benefit receipt in either of the programs. The screening and monitoring in both programs has been quite lax. Furthermore, since the replacement rate in the SI program is higher than in the DI program, there are economic incentives to remain on sickness benefits rather than on disability benefits. The implication is that the SI and DI programs are intimately related and that most or all individuals on disability benefits have had a long history of sickness absence. As a consequence, we will henceforth refer to individuals on long-term sickness absence (i.e., more than 90 days) as temporarily disabled.

During the first seven days of sickness absence, the individual him-/herself is decisive of whether being sick and to which extent it warrants absence from work. The individual merely has to inform the employer (or the SIA if being unemployed) that he or she is sick. As of the eighth day, a medical certificate is required, stating the length and extent of sick leave that is (expected to be) necessary.

Both the replacement rate and the employer's responsibility for sickness benefits have changed on several occasions during the last few decades. During the time period studied here, there were neither a qualifying period nor a period of employer provided sick-pay. Both before and after the reform, the insurance replaced a part (see Table 1 in the next subsection) of the earnings up to the social security ceiling of 7.5 price-base amounts. In 1991, this amounted to SEK 32,200 (appr. EUR 3,300) and only about 7 percent of the labor force had labor earnings above this ceiling.

2.2 The reform

The reform that we have exploited took effect on March 1, 1991. It was one of several budget cuts proposed by the Swedish government in early 1991 in response to the deep economic crisis in Sweden at the time. The reform implied that the insurance scheme changed from a flat replacement rate of 90 percent of foregone earnings, to a scheme where the replacement rate varied with the duration of the absence (see Table 1). More precisely, the replacement rate was reduced to 65 percent during the first three days of absence, and to 80 percent during days 4–90, while it remained unchanged at 90 percent from day 91. This design of the new SI scheme was motivated by a desire to cut public spending without adversely affect the financial situation further for an already disadvantaged group. The scheme applied to all new absence periods, but not to already ongoing ones.

Table 1: Replacement rates before and after the reform for new and ongoing absence periods

	New absences		Ongoing absences	
Absence length	Before reform	After reform	Before reform	After reform
1–3 days	90%	65%	90%	90%
4–90 days	90%	80%	90%	90%
$>$ 90 days	90%	90%	90%	90%

2.3 Theoretical predictions

The reform affected both the direct and indirect costs of absence as illustrated in Table 2, where we have defined the direct cost as 100% minus the replacement rate in percentages (i.e., the part not replaced by the insurance) and the indirect cost as the difference between the replacement rates if remaining absent and if starting a new absence period. The lower replacement rates (under the new scheme) during the first 90 days of absence implied increased (compared to the old scheme) direct costs of absence. These increased direct costs of absence are expected to have reduced the length of absence. For absence periods longer than 90 days there was no change in direct costs of absence, since the replacement rate remained unchanged at 90 percent. However, that the replacement rate was increasing by absence length introduced an indirect, or potential future, cost because of the risk of having to start a new absence period with a lower replacement rate in case of a relapse after having returned to work. That is, under the new scheme a temporarily disabled individual (i.e., someone with an absence longer than 90 days) would receive 90 percent of foregone earnings, but after having returned to work the replacement rate would be 65 percent (to start with) if starting a new absence period. A forward-looking individual would consider both the direct and indirect costs, while a myopic individual would consider only the direct costs. Hence, if temporarily disabled individuals behaved in a forward-looking manner, on average, they would have reacted to the reform by prolonged absences, while if they behaved in a myopic manner there would be no response to the reform.[10]

Table 2: Direct and indirect costs associated with the two regimen for new and ongoing absence periods

	New absences						Ongoing absences					
	Direct cost			Indirect cost			Direct cost			Indirect cost		
Absence length	Before reform	After reform	Diff.	Before reform	After reform	Diff.	Before reform	After reform	Diff.	Before reform	After reform	Diff
1–3 days	10%	35%	25%	0%	0%	0%	10%	10%	0%	0%	25%	25%
4–90 days	10%	20%	15%	0%	15%	15%	10%	10%	0%	0%	25%	25%
$>$ 90 days	10%	10%	0%	0%	25%	25%	10%	10%	0%	0%	25%	25%

Note: The direct cost is defined as 100% minus the replacement rate in percentages (i.e., the part not replaced by the insurance). The indirect cost is defined as the difference between the replacement rate for the particular absence length and the replacement rate if starting a new absence period.

[10] A simple theoretical model of how a forward-looking – in comparison to a myopic – individual will react to the reform is outlined in Appendix A

3 Data and empirical strategy

3.1 Empirical strategy

In this subsection, we will outline our empirical strategy to test for forward-looking moral hazard, among temporary disabled individuals, stemming from the indirect cost introduced by the reform. Ideally, we would like to compare the situation where only the indirect cost (not the direct cost) was changed to the situation with no changes in either direct or indirect costs.

A temporarily disabled individual (i.e., one having at least 90 days of absence) who began the absence period after the reform took effect would (as can be seen in Table 2) experience unchanged direct costs (i.e., 10 percent), but a changed indirect cost (i.e., from 0 to 25 percent). However, during the first 90 days of absence s/he would have experienced a larger direct cost (35 and 20 percent compared to 10 percent). This would create a dynamic selection problem (cf., Johansson and Palme, 2002, 2005), where those becoming temporarily disabled under the new scheme can be expected to be a more selected group than those becoming temporarily disabled under the old scheme. To circumvent this, we exploited the fact that the new scheme applied only to new entrants. Most importantly, however, ongoing absences would still be affected by the indirect cost (see Table 2). By limiting the analyses to those who (potentially) had reached 90 days of absence by the day before the reform took effect (i.e., February 28, 1991) we obtained a reform cohort (or study group) comprising those who (1) began an absence period between January 1 and November 30, 1990,[11] and (2) whose absence period were at least 90 days (i.e., they were temporary disabled). We selected a comparison cohort (or control group) similarly, but for the preceding year. Hence, the comparison cohort comprised those who (1) began an absence period between January 1 and November 30, 1989, and (2) whose absence period were at least 90 days.[12]

Using this sampling scheme we observed for each individual i in the two cohorts a pre-February 28 duration T_{i0} and potentially a post-February 28 duration T_{i1}.[13] For those leaving the absence before February 28 we only observed T_{i0}, while for those who had not returned to work by February 28, T_{i0} was censored by this date, and we also observed the post-February 28 duration T_{i1}. These latter durations (T_{i1}) were instead censored by the end of the year.[14]

In the following empirical analysis we have estimated both a linear difference-in-difference (DiD) model $T_{ij} = \alpha + \beta \mathbb{1}(j=1) + \gamma \mathbb{1}(D_i = 1) + \delta \mathbb{1}(j=1, D_i=1) + \varphi \mathbf{X_i} + \varepsilon_{ij}$, and a two-period Cox Proportional Hazard (Cox PH) model $h(t) = h_0(t) \exp(\beta \mathbb{1}(j=1) + \gamma \mathbb{1}(D_i=1) + \delta \mathbb{1}(j=1, D_i=1) + \varphi \mathbf{X_i})$, which handles the censoring. In these models, $\mathbf{X_i}$ is a vector of individual background characteristics measured the year(s) preceding the absences,[15] $\mathbb{1}(j=1)$ is an indicator function being equal to one for the post-February 28 period (i.e., $j=1$) and equal to zero for the pre-February 28 period (i.e., $j=0$), and $\mathbb{1}(D_i=1)$ is an indicator function being equal to one for the reform cohort (i.e., $D_i=1$) and equal to zero for the comparison cohort (i.e., $D_i=0$). The interaction between the two, represented by the indicator function $\mathbb{1}(j=1, D_i=1)$, is the variable of interest and its associated parameter δ is our estimate of the impact of the reform. A positive (negative in the Cox PH model) estimate of the coefficient δ would suggest that the reform had an unintended effect and that there exist forward-looking

[11] Excluding absences starting in December ascertains that all absences would have reached 90 days if lasting until February 28, 1991.

[12] For a placebo analysis presented in Section 4.2 we sampled an additional cohort (1988) in the same way.

[13] Hence, the total duration of an ongoing absence is $T_i = T_{i0} + T_{i1}$.

[14] The censoring is due to a new reform of the SI system by January 1, 1992.

[15] We will estimate both models both including $\mathbf{X_i}$ (i.e., covariate adjusted) and excluding $\mathbf{X_i}$ (i.e., unadjusted).

moral hazard behavior in the social insurance system.[16]

3.2 The data sources

The data used in the present study originate from Swedish administrative registers with universal coverage. Linkage between them is possible because of the 10-digit personal identity number maintained by the National Tax Board and unique to each Swedish resident. Specifically, three registers/databases were used to create the data set: First, to identify all periods of temporary disability we used the Sickness Benefit Register. This register is administered by the Swedish Social Insurance Agency (SIA) and contains information on sickness benefit payments for each individual. Most importantly, for this study, from 1986 the register contains both start and end date for each and every insured absence period.[17] Second, background characteristics were drawn from Statistics Sweden's longitudinal database LOUISE. This database contains comprehensive annual information from 1990, drawn from a number of administrative registers, for the nationally registered population aged 16–64 years by the end of each year. The aim of LOUISE is to enhance the conditions for research on sickness insurance and labor market issues requiring longitudinal individual data. However, that the database does not cover any years before 1990 is a limitation, since the comparison cohort is 1989 (and the additional cohort used in the placebo analysis is 1988). Therefore, we had to draw information also from the more limited Employment Register. The background information available both in the Employment Register and in LOUISE was: age, sex, immigration status, education, county of residence, and annual earnings. From the Sickness Benefit Register we could obtain also information on sickness absence during two preceding years.

3.3 Summary statistics and descriptive analysis

Before turning to the results from our estimations, we present some summary statistics and a descriptive analysis. Table 3 contains summary statistics for the pre-February 28 duration (T_{i0}) and the post-February 28 duration (T_{i1}), for both the reform cohort (1990) and the comparison cohort (1989).[18] The reform and comparison cohorts contain 158,460 and 153,500 absences, respectively. The pre-February 28 durations (T_{i0}) were on average 176.2 and 176.5 days. Somewhat less than half of these absences (i.e., 74,011 and 72,545, respectively) lasted until February 28. The durations in this latter period (T_{i1}) were on average 197.7 and 188.2 days in the reform and comparison cohort, respectively. Hence, these summary statistics show that (1) as expected there was no difference (i.e., less than -0.4 days) in pre-February 28 durations (T_{i0}), but (2) there was a considerable difference (i.e., 9.8 days) in post-February 28 durations (T_{i1}), between the two cohorts. This strongly suggests that the reform increased the duration of absence among temporarily disabled individuals, which in turn suggests that these individuals behaved as rational forward-looking agents.

[16]For the Cox PH model we will, in the following, present hazard ratios instead of the coefficients. A hazard ratio less than one is equivalent of a negative coefficient.

[17]Until the end of 1991 this covered basically all absence periods; Thereafter only absence periods longer than 14 days were recorded because of the introduction of a two-week period of employer-provided sick-pay at the beginning of each absence.

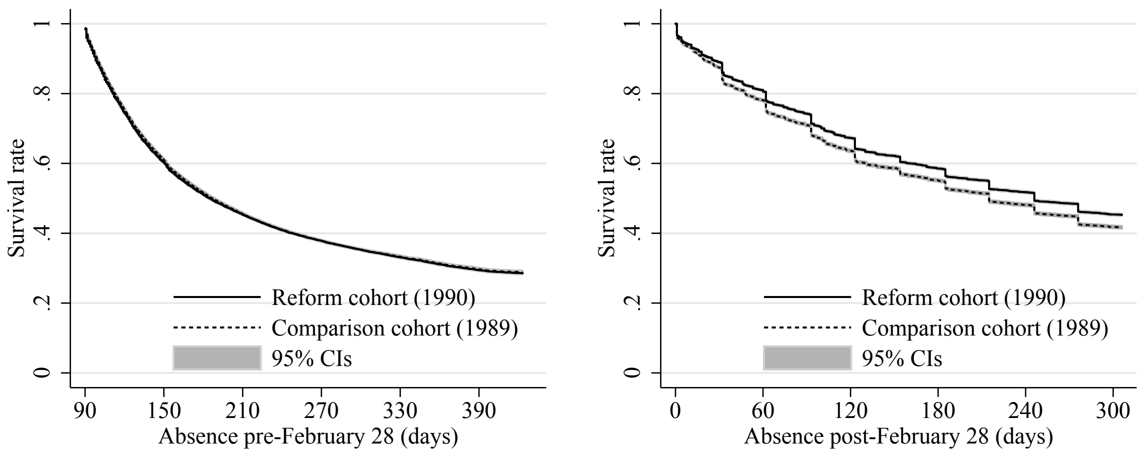
[18]The same statistics, but also including the additional cohort 1988 used in the placebo analysis, is found in Tables B1 and B1 together with summary statistics for all background characteristics.

Table 3: Summary statistics for the pre-February 28 duration (T_{i0}) and the post-February 28 duration (T_{i1}), for the reform cohort (1990) and the comparison cohort (1989)

	Reform cohort (1990)				Comparison cohort (1989)			
	Pre-February 28		Post-February 28		Pre-February 28		Post-February 28	
Absence	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)
T_{i0}	176.163	(85.421)	N/A		176.513	(84.858)	N/A	
T_{i1}	N/A		197.737	(116.765)	N/A		188.080	(118.654)
No obs.	158,460		74,011		153,680		72,611	

To give a more complete depiction of each duration, we have also plotted the Kaplan-Meier survival functions for the pre- and post-February 28 periods, conditional on being temporarily disabled (i.e., having an absence period longer than 90 days), in Figure 1. For the pre-February 28 period, the survival curves for the two cohorts are not distinguishable by the eye (Figure 1; left), while for the post-February 28 period, there is a clear gap between the two curves (Figure 1; right).[19]

Figure 1: Kaplan-Meier survival functions with 95 percent confidence intervals (95% CIs), for the pre-February 28 period (left) and the post-February 28 period (right)



4 Results

In this section, we present, in a first subsection, our main estimates of the behavioral response to the reform among temporarily disabled individuals. In the following subsection, we present the results from a placebo analysis to support our argument that our estimates can indeed be regarded as causal effects of the reform. This is followed by the results from repeating the main analysis for various subgroups (i.e., divided by sex and sickness absence history) for which different responses to the reform could be expected.

[19]The step-wise shape of the survival functions in the post-February 28 period is due to that all durations started the same day (i.e., March 1) and that more absence periods end the last day of each calendar month. While the latter is true also in pre-February 28 period, any such pattern is concealed by the varying start dates.

4.1 Main results

If temporarily disabled individuals are forward-looking, the indirect cost of returning to work due to the reform is expected to have prolonged the duration of absence (reduced the transition rate back to work) for the reform cohort relative the comparison cohort. The descriptive analysis in the previous section also suggested that the temporarily disabled individuals behaved in this way. In Table 4, we present the results from formally testing this by estimating (unadjusted and adjusted) linear DiD and two-period Cox Proportional Hazard (Cox PH) models (see Section 3.1).

Although it is the coefficient for the indicator function $\mathbb{1}(j=1, D_i=1)$ (i.e., the interaction between the post-February 28 indicator and the reform cohort indicator) that is of main interest, we also present the estimated coefficients for the post-February 28 indicator $\mathbb{1}(j=1)$ and the reform cohort indicator $\mathbb{1}(D_i=1)$. For the Cox PH model we present the hazard ratios (HRs) instead of the coefficients.

Table 4: Main analysis: Unadjusted and covariate adjusted coefficient estimates from a linear difference-in-difference (DiD) and hazard ratios (HRs) from a two-period Cox proportional hazard (Cox PH) model, with 95 percent confidence intervals (95% CIs)

	Linear DiD				Cox PH			
	Unadjusted		Adjusted		Unadjusted		Adjusted	
Covariate	Coef.	95% CI	Coef.	95% CI	HR	95% CI	HR	95% CI
$\mathbb{1}(j=1)$	11.567	(10.662, 12.472)	7.172	(6.267, 8.077)	1.027	(1.014, 1.040)	1.107	(1.093, 1.121)
$\mathbb{1}(D=1)$	-0.350	(-0.950, 0.250)	-0.327	(-0.767, 0.412)	1.015	(1.001, 1.024)	1.011	(1.003, 1.020)
$\mathbb{1}(j=1, D=1)$	10.007	(8.737, 11.278)	10.049	(8.789, 11.309)	0.869	(0.860, 0.890)	0.875	(0.854, 0.884)
No obs.	458,762		458,762		458,762		458,762	

Note: $\mathbb{1}(j=1)$ is an indicator function being equal to one for the post-February 28 period (i.e., $j=1$) and equal to zero for the pre-February 28 period (i.e., $j=0$), and $\mathbb{1}(D=1)$ is an indicator function being equal to one for the reform cohort (i.e., $D=1$) and equal to zero for the comparison cohort (i.e., $D=0$). The interaction between the two, is represented by the indicator function $\mathbb{1}(j=1, D=1)$. The adjusted models also include age (9 categories), female, foreign born, attained education (4 categories), previous earnings (lagged 2 and 3 years), county of residence, absence days (lagged 2 and 3 years), absence periods (lagged 2 and 3 years), and long absence periods (lagged 2 and 3 years). Robust standard errors are clustered at the individual level.

A first observation is that whether we adjust for potential differences in observed characteristics, or not, have virtually no influence at all on our estimate of the effect of the reform. Therefore, we will henceforth refer only to the estimates from the adjusted models unless otherwise stated. We find that the reform increased the duration of absence among temporarily disabled individuals by 10 days (95% CI: 8.8, 11.3; see the linear DiD model in Table 4). While this estimate does not take censoring into account, the hazard ratio of 0.87 (95% CI: 0.85, 0.88), i.e., a relative decrease in the hazard of leaving absence by 13 percent, obtained by the Cox PH model do. Hence, it seems that the indirect cost, that was introduced by the reform, indeed made temporarily disabled individuals more reluctant to return to work.

It is also evident from Table 4 that there were no differences between the two cohorts during the pre-February 28 period. That is, the estimated coefficient for $\mathbb{1}(D=1)$ is statistically insignificant and close to zero in the linear DiD model, and although it is marginally significant in the Cox PH model, it is also close to one.

4.2 Placebo Analyses

As a test of our identification strategy we have performed a placebo analysis, where the reform date was artificially changed to March 1, 1990 (i.e., one year before the actual reform). The cohort that hitherto has been used as the comparison cohort in the analyses then became the (placebo) reform cohort, and the preceding cohort became the new comparison cohort. Using these two cohorts we repeated the analyses of Section 4.1. If we were to find a statistically significant placebo effect, this would cast serious doubts on whether the previously reported estimates represent causal (indirect) effects of the actual reform.

However, the resulting estimates from this exercise, presented in Table 5, are statistically insignificant both in the linear DiD and in the Cox PH model.[20] We interpret these findings as supportive of our claim that the estimation strategy provides causal estimates of the (indirect) effect of the reform.

Table 5: Placebo analysis: Unadjusted and covariate adjusted coefficient estimates from a linear difference-in-difference (DiD) and hazard ratios (HRs) from a two-period Cox proportional hazard (Cox PH) model, with 95 percent confidence intervals (95% CIs)

	Linear DiD				Cox PH			
	Unadjusted		Adjusted		Unadjusted		Adjusted	
Covariate	Coef.	95% CI	Coef.	95% CI	HR	95% CI	HR	95% CI
$\mathbb{1}(j=1)$	12.217	(11.323, 13.111)	7.663	(6.768, 8.557)	1.046	(1.033, 1.059)	1.123	(1.109, 1.137)
$\mathbb{1}(D=1)$	-1.673	(-2.280, -1.066)	-0.928	(-1.523, -0.333)	1.030	(1.021, 1.039)	1.017	(1.008, 1.026)
$\mathbb{1}(j=1, D=1)$	-0.650	(-1.922, 0.622)	-0.759	(-2.019, 0.502)	0.979	(0.963, 0.996)	0.984	(0.968, 1.001)
No obs.	455,705		455,705		455,705		455,705	

Note: $\mathbb{1}(j=1)$ is an indicator function being equal to one for the post-February 28 period (i.e., $j=1$) and equal to zero for the pre-February 28 period (i.e., $j=0$), and $\mathbb{1}(D=1)$ is an indicator function being equal to one for the reform cohort (i.e., $D=1$) and equal to zero for the comparison cohort (i.e., $D=0$). The interaction between the two, is represented by the indicator function $\mathbb{1}(j=1, D=1)$. The adjusted models also include age (9 categories), female, foreign born, attained education (4 categories), previous earnings (lagged 2 and 3 years), county of residence, absence days (lagged 2 and 3 years), absence periods (lagged 2 and 3 years), and long absence periods (lagged 2 and 3 years). Robust standard errors are clustered at the individual level.

4.3 A subgroup analysis

In this subsection, we present the results from analyses of various subgroups (i.e., the sample has been divided by sex and various sickness absence histories) for which different responses to the reform could be expected. For brevity, we only report the estimates of coefficient associated with our variable of interest (i.e., $\mathbb{1}(j=1, D=1)$).

First, we have repeated the analysis, separately, for men and women, whose absence behaviour is known to differ. On the one hand, it is well known that women, on average, are more absent from work for health reasons than are men (e.g., Paringer, 1983; Broström et al., 2004; Mastekaasa and Olsen, 1998; Angelov et al., 2013). On the other hand, men have been found to react more strongly to cuts in replacement rates (e.g., Henrekson and Persson, 2004; Johansson and Palme, 1996; Ziebarth and Karlsson, 2014). The results from this analysis are reported in Table 6.

[20] Although the unadjusted estimate in the Cox PH model is statistically significant, the corresponding estimate in the linear DiD model is not.

Table 6: A subgroup analysis by previous sickness: Unadjusted and covariate adjusted coefficient estimates from a linear difference-in-difference (DiD) and hazard ratios (HRs) from a two-period Cox proportional hazard (Cox PH) model, with 95 percent confidence intervals (95% CIs)

	Linear DiD				Cox PH				
	Unadjusted		Adjusted		Unadjusted		Adjusted		
Covariate	Coef.	95% CI	Coef.	95% CI	HR	95% CI	HR	95% CI	No obs.
Sex									
Men	8.793	(6.821, 10.765)	8.881	(6.918, 10.845)	0.888	(0.866, 0.912)	0.882	(0.859, 0.906)	189,269
Women	10.978	(9.315, 12.640)	10.903	(9.261, 12.545)	0.865	(0.846, 0.885)	0.860	(0.841, 0.880)	269,493
Absence days									
0–23	8.414	(5.824, 11.005)	8.622	(6.062, 11.181)	0.888	(0.858, 0.919)	0.875	(0.845, 0.906)	113,022
23–72	9.702	(7.137, 12.267)	9.650	(7.119, 12.181)	0.874	(0.844, 0.905)	0.866	(0.836, 0.897)	115,112
72–176	10.801	(8.256, 13.346)	10.830	(8.305, 13.356)	0.866	(0.837, 0.896)	0.862	(0.832, 0.892)	115,239
176–	10.813	(8.336, 13.290)	10.908	(8.436, 13.379)	0.876	(0.847, 0.906)	0.872	(0.843, 0.903)	115,389
All absences									
0–3	6.845	(4.256, 9.435)	6.986	(4.415, 9.557)	0.912	(0.881, 0.945)	0.907	(0.875, 0.939)	107,292
3–6	9.592	(7.068, 12.115)	9.671	(7.169, 12.174)	0.882	(0.852, 0.913)	0.872	(0.843, 0.903)	115,150
6–10	10.157	(7.625, 12.689)	10.108	(7.599, 12.617)	0.876	(0.847, 0.906)	0.872	(0.842, 0.903)	115,982
10–	13.702	(11.167, 16.237)	13.662	(11.146, 16.179)	0.838	(0.810, 0.867)	0.834	(0.806, 0.863)	120,388
Long absences									
0–1	9.849	(8.401, 11.298)	9.908	(8.475, 11.342)	0.878	(0.860, 0.894)	0.869	(0.852, 0.886)	355,033
1–	10.599	(7.945, 13.252)	10.636	(7.990, 13.282)	0.869	(0.838, 0.901)	0.867	(0.836, 0.899)	103,729

Note: The presented estimates are the the estimated coefficient associated with $\mathbb{1}(j=1, D=1)$. All other coefficient estimates are suppressed for brevity. The estimated models include age (9 categories), female, foreign born, attained education (4 categories), previous earnings (lagged 2 and 3 years), county of residence, absence days (lagged 2 and 3 years), absence periods (lagged 2 and 3 years), and long absence periods (lagged 2 and 3 years). Robust standard errors are clustered at the individual level.

The sample of women is considerably larger than the sample of men (269,382 compared to 189,134), which is in line with that women generally are more absent from work for health reasons than are men. However, the estimates actually suggest that the effect of the reform was somewhat larger among women than among men, contrary to what could be expected based on previous studies showing that men react more strongly to changes in replacement rates.[21]

Second, we have also repeated the analysis for subgroups based on their sickness absence histories. Because the indirect cost of absence was realized only if starting a new absence period, the temporarily disabled individuals' history of sickness absence may have affected the response to the reform. It is plausible to assume that individuals with a history of repeated and/or longer absence periods, on average, have poorer health. Given this, they would likely have perceived a higher risk of actually encountering the potential future cost of returning to work (i.e., having another absence period with a lower replacement rate). If this was the case, we should expect to find a larger effect among the temporarily disabled individuals' with a history of sickness absence.

We present in Table 6 also the estimates from this analysis, where the sample has been divided in quarters (if possible) based on the number of previous absence days, absence periods, and long absence periods (i.e., more than 90 days), respectively. It does not seem to be the case that the estimated reform effect is consistently larger among those with a history of sickness but it seems to depend on how we measure it. Neither the number of previous days of sickness absence nor the number of previous long absence periods largely affected the response to the reform, whereas the number of previous absence periods (short and long) did. In the linear DiD model the reform effect is twice as large among those with the most previous absence periods (i.e., those with at least 10 periods during a two-year period) as among those with the fewest previous absence periods (i.e., those with less than 3 during a two-year period). The corresponding hazard ratios from the Cox PH model are 0.83 (95% CI: 0.81, 0.86) and 0.91 (95% CI: 0.88, 0.94), respectively. Because the indirect cost introduced by the reform was realized only if the return to work was followed by a new absence period this makes sense, given that the subjective risk of relapse is positively related to previous absences.

5 Discussion and conclusions

Policymakers face the problem of balancing the benefits of risk-sharing, allowing for intertemporal consumption smoothing, associated with a generous insurance program with the disincentive, or moral hazard, effects of the same program. There is strong empirical evidence of quite substantial effects of changes in the replacement rate on the transition rates out of insured unemployment and short-term sickness absence. However, the studies on how economic incentives affect the inflow to, and outflow from, temporary and permanent disability in the sickness, disability and worker compensation insurance programs are much fewer and also provide a less coherent picture. One potential, and obvious, explanation to the weaker evidence on temporarily and permanently disabled peoples' behavioural responses to economic incentives, within respective program, is that their sickness/disability leaves little room for such adjustments. Another potential explanation, however, is that there are differences in expected future costs across various states and programs. In social insurance programs where the eligibility criteria is difficult to verify, and the benefits do not diminish with the duration in the state or there are no

[21] To investigate whether the difference between men and women is statistically significant, we have also estimated pooled models fully interacted with the indicator for being a woman. The interaction with $\mathbb{1}(j=1, D=1)$ was statistically insignificant in both the linear DiD and the Cox PH model.

time limits for benefit receipt, it is likely that individuals, if being forward-looking, will not only respond to the direct cost/benefit but also take into account any expected future costs of re-entering the benefit state.

In this study we have tested for the presence of such forward-looking moral hazard behavior by exploiting a reform of the Swedish SI program as a natural experiment. Prior to the reform, the insurance replaced 90 percent of foregone earnings from the first day of absence and onward (i.e. there was no time limit of benefit receipt). Following the reform, the replacement rate was reduced to 65 percent during the first three days of absence and to 80 percent during days 4–90. From day 91 onward it remained unchanged at 90 percent. The reason behind maintaining the compensation level for absences longer than 90 days was to avoid to adversely affect the financial situation further for an already disadvantaged group. The reduced replacement rate for shorter absences did not only imply an increased direct costs of being absent for those with shorter absences but among those with longer absences it also introduced an indirect, or potential future, cost associated with ending an absence period: that is, after having returned to work a relapse requiring a new absence period would imply receiving 65 percent (to start with) of foregone earnings instead of 90 percent had they remained absent.

We show, by applying a linear difference-in-difference model (and a two-period Cox PH model) to administrative register data on the complete account of all absence periods during the period of interest, that the potential future cost decreased the transition back to work by 13 percent and prolonged long absences by 10 days among the temporarily disabled. This suggests that temporarily disabled individuals, indeed, take potential future costs into account, and that forward-looking moral hazard is present in the Swedish social insurance system. A placebo test also supports our claim of a causal interpretation. Moreover, contrary to most previous studies we do not find that men reacted more strongly than women to the changes in the replacement rate. A potential explanation is that women are more risk averse than are men. In previous studies the focus has been on the effects of changes in the direct costs (i.e., the replacement rate for the current absence period), which are not associated with any uncertainty. However, the indirect cost (i.e., the potential future cost associated with relapse that requires a new absence period with a lower replacement rate) studied here was realized if, and only if, a new absence period was started after having returned to work. Hence, all else equal, one would expect the response to be stronger among more risk averse individuals. A potentially stronger responsiveness to economic incentives among men might then be outweighed by a greater risk averseness among women. Furthermore, we also found that those with more absence periods in the past seem to have responded more strongly to the reform. Once again, since the indirect cost was realized only if a new absence period was started after having returned to work, this finding should be expected, given forward-looking behaviour and that the perceived risk of relapse increases with the number of past absence periods.

To conclude, our findings do not only show that, in line with economic theory, temporarily disabled individuals respond to economic incentives but also that they are forward-looking. We believe that these findings contribute to two different fields of the literature. First, our (arguably) convincing causal estimates add to the sparse literature on the importance of economic incentives for absence behaviour among temporarily disabled individuals. Second, we also add to a rapidly growing literature on the empirical testing for forward-looking behavior. The latter is not only of academic interest but from a policy perspective it is important to know whether absentees take potential future costs into account, and how this affects their absence behaviour, when designing and evaluating social insurance programs. Moreover, our findings also suggest that, in order to reduce moral hazard, the replacement rates in the temporary disability insurance should decrease with the duration or there should be a time limit for benefit receipt. In the latter case, individuals should be screened and those no longer deemed able to work should be granted permanent disability insurance with a replacement rate lower than in the end of the period of temporary disability.

References

Aaviksoo, E. and Kiivet, R. A. (2016), 'Influence of the sickness benefit reform on sickness absence', *Health Policy* **120**(9), 1070–1078.

Abaluck, J., Gruber, J. and Swanson, A. (2015), Prescription drug use under medicare part d: a linear model of nonlinear budget sets, Technical report, National Bureau of Economic Research.

Alpert, A. (2016), 'The anticipatory effects of medicare part d on drug utilization', *Journal of Health Economics* **49**, 28 – 45.

Angelov, N., Johansson, P. and Lindahl, E. (2013), Gender differences in sickness absence and the gender division of family responsibilities, Working Paper 2013:9, IFAU - Institute for Evaluation of Labour Market and Education Policy, Uppsala.

Autor, D., Duggan, M. and Gruber, J. (2014), 'Moral hazard and claims deterrence in private disability insurance', *American Economic Journal: Applied Economics* **6**(4), 110–141.

Broström, G., Johansson, P. and Palme, M. (2004), 'Economic incentives and gender differences in work absence behavior', *Swedish Economic Policy Review* **11**, 33–63.

Byrnes, J. P., Miller, D. C. and Schafer, W. D. (1999), 'Gender differences in risk taking: A meta-analysis.', *Psychological Bulletin* **125**(23), 367–383.

Campolieti, M. (2004), 'Disability insurance benefits and labor supply: Some additional evidence', *Journal of Labor Economics* **22**(4), 863–889.

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

Chevalier, J. and Goolsbee, A. (2009), 'Are durable goods consumers forward-looking? evidence from college textbooks', *The Quarterly Journal of Economics* **124**(4), 1853–1884.

Croson, R. and Gneezy, U. (2009), 'Gender differences in preferences', *Journal of Economic literature* **47**(2), 448–474.

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

Dalton, C. M., Gowrisankaran, G. and Town, R. (2015), Myopia and complex dynamic incentives: Evidence from medicare part d, Technical report, National Bureau of Economic Research.

De Paola, M., Scoppa, V. and Pupo, V. (2014), 'Absenteeism in the italian public sector: The effects of changes in sick leave policy', *Journal of Labor Economics* **32**(2), 337–360.

Eckel, C. C. and Grossman, P. J. (2002), 'Sex differences and statistical stereotyping in attitudes toward financial risk', *Evolution and Human Behavior* **23**(4), 281–295.

Eckel, C. C. and Grossman, P. J. (2008), 'Men, women and risk aversion: Experimental evidence', *Handbook of experimental economics results* **1**, 1061–1073.

Einav, L., Finkelstein, A. and Schrimpf, P. (2015), 'The response of drug expenditure to non-linear contract design: evidence from medicare part d', *The quarterly journal of economics* **130**(2), 841–899.

Fevang, E., Hardoy, I. and Røed, K. (2017), 'Temporary disability and economic incentives', *The Economic Journal* **126**(603), 1410–1432.

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

Gruber, J. and Köszegi, B. (2001), 'Is addiction "rational"? theory and evidence*', *The Quarterly Journal of Economics* **116**(4), 1261–1303.

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

Johansson, P. and Palme, M. (1996), 'Do economic incentives affect work absence? empirical evidence using swedish micro data', *Journal of Public Economics* **59**(2), 195–218.

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

Johansson, P. and Palme, M. (2004), Moral hazard and sickness insurance: Empirical evidence from a sickness insurance reform in sweden, Working Paper 2004:10, IFAU - Institute for Evaluation of Labour Market and Education Policy, Uppsala.

Johansson, P. and Palme, M. (2005), 'Moral hazard and sickness insurance', *Journal of Public economics* **89**(9), 1879–1890.

Kaplan, C. M. and Zhang, Y. (2017), 'Anticipatory behavior in response to medicare part d's coverage gap', *Health Economics* **26**(3), 338–351.

Kostøl, A. R. and Mogstad, M. (2014), 'How financial incentives induce disability insurance recipients to return to work', *The American Economic Review* **104**(2), 624–655.

Krueger, A. B. and Meyer, B. D. (2002), 'Labor supply effects of social insurance', *Handbook of public economics* **4**, 2327–2392.

Long, S. H., Marquis, M. and Rodgers, J. (1998), 'Do people shift their use of health services over time to take advantage of insurance?', *Journal of Health Economics* **17**(1), 105–115.

Markussen, S., Røed, K., Røgeberg, O. J. and Gaure, S. (2011), 'The anatomy of absenteeism', *Journal of health economics* **30**(2), 277–292.

Mastekaasa, A. and Olsen, K. M. (1998), 'Gender, absenteeism, and job characteristics a fixed effects approach', *Work and Occupations* **25**(2), 195–228.

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

Neuhauser, F. and Raphael, S. (2004), 'The effect of an increase in worker's compensation benefits on the duration and frequency of benefit receipt', *The Review of Economics and Statistics* **86**(1), 288–302.

Paringer, L. (1983), 'Women and absenteeism: health or economics?', *The American Economic Review* **73**(2), 123–127.

Pettersson-Lidbom, P. and Thoursie, P. S. (2013), 'Temporary disability insurance and labor supply: evidence from a natural experiment', *The Scandinavian Journal of Economics* **115**(2), 485–507.

Pierani, P. and Tiezzi, S. (2011), 'Infrequency of purchase, individual heterogeneity and rational addiction in single households' estimates of alcohol consumption', *Giornale degli Economisti e Annali di Economia* **70**(2), 93–116.

Pollak, C. (2017), 'The impact of a sick pay waiting period on sick leave patterns', *The European Journal of Health Economics* **18**(1), 13–31.

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

Tiezzi, S. (2005), 'An empirical analysis of tobacco addiction in italy', *The European Journal of Health Economics* **6**(3), 233–243.

Ziebarth, N. R. (2013), 'Long-term absenteeism and moral hazard—evidence from a natural experiment', *Labour Economics* **24**, 277–292.

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

Ziebarth, N. R. and Karlsson, M. (2014), 'The effects of expanding the generosity of the statutory sickness insurance system', *Journal of Applied Econometrics* **29**(2), 208–230.

Appendix A Theoretical framework

Here we present a simple model to show how a forward-looking – in comparison to a myopic – individual will react to a change in the social insurance scheme such as the one described in the previous section.[22] Assume that the individual's instantaneous utility (u_t) is a weighted sum of consumption (c_t) and leisure (l_t):

$$u_t = (1-\sigma_t)c_t + \sigma_t l_t \tag{A.1}$$

where σ_t denotes the sickness level (related to his/her work), which for simplicity is assumed to be uniformly distributed over [0,1]. The larger σ_t, the sicker is the individual, and the larger weight is given to leisure (or recuperation time) and the less weight is given to consumption. A working individual receives a (fixed) net wage income w (normalized to one) from h hours of work, which is used to consumption, and enjoys $T-h$ hours of leisure, where T is total time available. Hence, the instantaneous utility is $u_t = (1-\sigma_t)w + \sigma_t(T-h)$. An individual that is absent due to health problems instead receives sickness benefits equal to a fraction r of the net wage income w and enjoys T hours of leisure (or recuperation time). Hence, the instantaneous utility is $u_t = (1-\sigma_t)rw + \sigma_t T$. To model the social insurance scheme described in the previous section, we need a model with (at least) two periods. In a first period of work absence the replacement rate is r_s, while in case of long-term sickness or temporary disability (i.e., the individual is absent for more than one period) the replacement rate is r_l, where $0 < r_s \leq r_l < 1$. As our focus is on long-term sickness or temporary disability, we assume that the individual is absent at the start and decides in period 1 whether to remain absent or to return to work. Using this model we can determine the "reservation sickness level" σ_1^{a*} (i.e., the analogue to the reservation wage in job search models) at which the individual is indifferent between returning to work and remaining absent. That is, it is the level of sickness that equalizes the utility from returning to work in period 1 plus the discounted expected utility in period 2 (conditional on working in the previous period, denoted by superscript w) and the utility from remaining absent in period 1 plus the discounted expected utility in period 2 (conditional on being absent in the previous period, denoted by superscript a). More formally:

$$(1-\sigma_1^{a*})w + \sigma_1^{a*}(T-h) + \tfrac{1}{1+\rho}E[u_2^w] = (1-\sigma_1^{a*})r_l w + \sigma_1^{a*}T + \tfrac{1}{1+\rho}E[u_2^a], \tag{A.2}$$

where ρ denotes the individuals time preference rate. An individual who is absent in period 1 will remain absent in period 2 if $\sigma_2 \geq \sigma_2^{a*}$, where σ_2^{a*} is the reservation sickness level in period 2 conditional on being absent in period 1. An individual who instead is working in period 1 will become absent in period 2 if $\sigma_2 > \sigma_2^{w*}$, where σ_2^{w*} is the reservation sickness level in period 2 conditional on working in the previous period.

Using that σ_2 is uniformly distributed over $[0,1]$, we can define $E[u_2^w]$ and $E[u_2^a]$ as:[23]

$$E[u_2^w] = \sigma_2^{w*}\left[(1-\tfrac{\sigma_2^{w*}}{2})w + \tfrac{\sigma_2^{w*}}{2}(T-h)\right] + (1-\sigma_2^{w*})\left[(1-\tfrac{1+\sigma_2^{w*}}{2})r_s w + \tfrac{1+\sigma_2^{w*}}{2}T\right] \tag{A.3}$$

and

$$E[u_2^a] = \sigma_2^{a*}\left[(1-\tfrac{\sigma_2^{a*}}{2})w + \tfrac{\sigma_2^{a*}}{2}(T-h)\right] + (1-\sigma_2^{a*})\left[(1-\tfrac{1+\sigma_2^{a*}}{2})r_l w + \tfrac{1+\sigma_2^{a*}}{2}T\right] \tag{A.4}$$

[22] The model follows Ziebarth (2013), but will differ in that the replacement rate for short-term absence is lower than the replacement rate for long-term absence and not the other way around.

[23] Specifically, we use that $\Pr[\sigma_2 \leq \sigma_2^{s*}] = \sigma_2^{s*}$, $E[\sigma_2|\sigma_2 \leq \sigma_2^{s*}] = \sigma_2^{s*}/2$, and $E[\sigma_2|\sigma_2 > \sigma_2^{s*}] = (1+\sigma_2^{s*})/2$, where $s = a, w$.

Note that the reservation sickness levels σ_2^{w*} and σ_2^{a*} equalize the terms within the left and right brackets of equations (A.3) and (A.4), respectively. We can derive the following explicit solutions:

$$\sigma_2^{w*} = \frac{w - r_s w}{w - r_s w + h} \tag{A.5}$$

$$\sigma_2^{a*} = \frac{w - r_l w}{w - r_l w + h} \tag{A.6}$$

Taking the partial derivative of σ_2^{w*} and σ_2^{a*} with respect to r_s and r_l, respectively, it is evident that the threshold or reservation sickness levels are decreasing with the replacement rates:

$$\frac{\partial \sigma_2^{w*}}{\partial r_s} = \frac{wh}{(w - r_s w + h)^2} < 0$$

$$\frac{\partial \sigma_2^{a*}}{\partial r_l} = \frac{wh}{(w - r_l w + h)^2} < 0$$

Hence, a higher replacement rate will increase the probability of both becoming and remaining absent.

By substituting equations (A.3)–(A.6) into (A.2) and solving for σ_1^{a*}, we obtain:

$$\sigma_1^{a*} = \sigma_2^{a*} + \tfrac{1}{(1+\rho)}\kappa, \tag{A.7}$$

where

$$\kappa = \frac{(r_s - r_l)wh^2}{2(w - r_s w + h)(w - r_l w + h)^2} > 0. \tag{A.8}$$

Because σ_1 is uniformly distributed over $[0,1]$ the transition rate back to work in period 1, i.e., $E[\sigma_1 < \sigma_1^{a*}]$, is directly obtained from equation A.7. The discounted term in equation A.7 can be viewed as the indirect effect of the reform on the transition rate or, in other words, the impact on the transition rate of the expected cost of having to return to sickness absence, at a lower replacement rate, in the future. To determine how a change in the replacement rate for short-term absence affects the transition rate for temporarily disabled individuals, we take the partial derivative of equation A.7 with respect to r_s:

$$\frac{\partial \sigma_1^{a*}}{\partial r_s} = \frac{\partial \sigma_2^{a*}}{\partial r_s} + \frac{1}{1+\rho}\frac{\partial \kappa}{\partial r_s} = 0 + \frac{1}{1+\rho}\frac{wh^2}{2(w - r_l w + h)(w - r_s w + h)^2} > 0 \tag{A.9}$$

Equation A.9 shows that – given that individuals are forward-looking – lowering the replacement rate for short-term absence will actually decrease the transition rate back to work among the temporarily disabled. For the myopic individual, σ_1^{a*} instead equals σ_2^{a*} and a change in the replacement rate for short-term absence will not affect the transition rate at all.

The implication from this simple model is that in a system with a cost of, or a risk of not being able to, reenter the benefit state, we will in general for a given (exogenous) reduction of r_s (r_l) underestimate (overestimate) the effect of economic incentives on the transition rate out of the benefit state given that individuals are forward-looking. In our empirical analysis, we can then test whether temporarily disabled individuals behave in a forward-looking or myopic manner. In the theoretical model we extracted from the more realistic situation where health in period 2 is correlated with health in period 1 and also can be affected by whether being absent or working. This would introduce dynamic selection into the model. Hence, a direct empirical comparison of the populations in respective regime would be biased. However, by exploiting that the reform affected only new entrants allow us to circumvent the bias associated with dynamic sorting. Controlling for dynamic selection on σ_t and using the exogenous variation in r_s provided by the reform, we are able to estimate the transition rate back to work for temporary disabled individuals under the two different regimes when $r_s < r_l$ and $r_s = r_l$, respectively. A difference between the two estimated transition rates will suggest both that the reform had an unintended effect and that there exist forward-looking moral hazard behavior in the SI.

Appendix B Additional descriptive statistics

Table B1: Descriptive statistics

	Cohort 1991				Cohort 1990				Cohort 1989			
	Pre-February 28		Post-February 28		Pre-February 28		Post-February 28		Pre-February 28		Post-February 28	
Characteristics	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)
$T_{i0}{}^a$	176.163	(85.421)	N/A		176.513	(84.858)	N/A		178.186	(86.228)	N/A	
$T_{i1}{}^b$	N/A		197.737	(116.765)	N/A		188.080	(118.654)	N/A		190.403	(116.820)
Female	0.594	(0.491)	0.572	(0.495)	0.595	(0.491)	0.573	(0.495)	0.593	(0.491)	0.571	(0.495)
Foreign born	0.159	(0.366)	0.165	(0.372)	0.162	(0.369)	0.167	(0.373)	0.160	(0.367)	0.166	(0.372)
Age												
20–24 years	0.065	(0.246)	0.045	(0.207)	0.068	(0.252)	0.048	(0.214)	0.066	(0.249)	0.048	(0.214)
25–29 years	0.106	(0.308)	0.075	(0.264)	0.106	(0.308)	0.077	(0.266)	0.099	(0.298)	0.072	(0.258)
30–34 years	0.106	(0.308)	0.086	(0.281)	0.108	(0.311)	0.087	(0.281)	0.105	(0.306)	0.082	(0.275)
35–39 years	0.106	(0.307)	0.094	(0.292)	0.106	(0.308)	0.094	(0.292)	0.104	(0.305)	0.092	(0.289)
40–44 years	0.116	(0.321)	0.115	(0.319)	0.117	(0.321)	0.114	(0.318)	0.117	(0.322)	0.112	(0.315)
45–49 years	0.127	(0.334)	0.130	(0.337)	0.122	(0.327)	0.124	(0.329)	0.115	(0.319)	0.115	(0.320)
50–54 years	0.116	(0.321)	0.127	(0.334)	0.114	(0.318)	0.125	(0.331)	0.114	(0.317)	0.124	(0.330)
55–59 years	0.123	(0.328)	0.148	(0.355)	0.125	(0.331)	0.151	(0.358)	0.134	(0.341)	0.159	(0.366)
60–64 years	0.134	(0.341)	0.179	(0.383)	0.134	(0.340)	0.182	(0.386)	0.146	(0.353)	0.196	(0.397)
Attained education												
Compulsory schooling	0.379	(0.485)	0.409	(0.492)	0.401	(0.490)	0.434	(0.496)	0.418	(0.493)	0.456	(0.498)
Upper secondary schooling	0.420	(0.494)	0.405	(0.491)	0.407	(0.491)	0.391	(0.488)	0.395	(0.489)	0.376	(0.484)
College/university	0.131	(0.337)	0.117	(0.321)	0.119	(0.324)	0.105	(0.307)	0.115	(0.320)	0.100	(0.300)
Unknown	0.071	(0.256)	0.069	(0.253)	0.072	(0.259)	0.069	(0.254)	0.071	(0.257)	0.069	(0.253)

[a] Refers to the pre-February 28 duration.
[b] Refers to the post-February 28 duration.

Table B1: Continued.

	Cohort 1991				Cohort 1990				Cohort 1989			
	Pre-February 28		Post-February 28		Pre-February 28		Post-February 28		Pre-February 28		Post-February 28	
Characteristics	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)
Previous earnings and sickness[a]												
Earnings (SEK 1000s)	145.817	(92.820)	145.116	(93.670)	140.309	(89.875)	140.137	(90.070)	138.758	(87.898)	139.124	(88.577)
# Absence days	63.187	(74.808)	69.653	(80.607)	61.803	(72.221)	67.589	(77.817)	59.619	(70.612)	64.721	(75.063)
# Absence periods	3.621	(2.966)	3.480	(2.977)	3.358	(2.727)	3.222	(2.685)	3.045	(2.548)	2.940	(2.491)
# Long absence periods	0.136	(0.279)	0.145	(0.279)	0.131	(0.274)	0.138	(0.273)	0.128	(0.269)	0.136	(0.270)
County of residence												
Stockholm	0.196	(0.397)	0.191	(0.393)	0.192	(0.394)	0.185	(0.388)	0.196	(0.397)	0.186	(0.389)
Uppsala	0.032	(0.177)	0.032	(0.175)	0.031	(0.173)	0.029	(0.167)	0.031	(0.175)	0.029	(0.168)
Södermanland	0.032	(0.175)	0.034	(0.182)	0.031	(0.173)	0.032	(0.177)	0.030	(0.171)	0.032	(0.175)
Östergötland	0.044	(0.205)	0.045	(0.208)	0.045	(0.207)	0.046	(0.208)	0.047	(0.212)	0.049	(0.216)
Jönköping	0.032	(0.176)	0.033	(0.178)	0.032	(0.175)	0.031	(0.175)	0.032	(0.177)	0.033	(0.179)
Kronoberg	0.018	(0.132)	0.018	(0.133)	0.018	(0.133)	0.018	(0.133)	0.017	(0.131)	0.017	(0.131)
Kalmar	0.026	(0.159)	0.026	(0.160)	0.025	(0.155)	0.026	(0.158)	0.024	(0.152)	0.025	(0.155)
Gotland	0.006	(0.077)	0.006	(0.079)	0.006	(0.078)	0.006	(0.077)	0.006	(0.079)	0.006	(0.076)
Blekinge	0.017	(0.131)	0.018	(0.133)	0.017	(0.130)	0.019	(0.138)	0.016	(0.125)	0.017	(0.130)
Kristianstad	0.030	(0.170)	0.029	(0.169)	0.031	(0.175)	0.031	(0.174)	0.031	(0.172)	0.033	(0.178)
Malmöhus	0.088	(0.283)	0.085	(0.279)	0.091	(0.287)	0.091	(0.288)	0.089	(0.284)	0.090	(0.286)
Halland	0.024	(0.154)	0.025	(0.156)	0.024	(0.154)	0.025	(0.156)	0.023	(0.150)	0.023	(0.151)
Göteborg	0.100	(0.300)	0.102	(0.303)	0.100	(0.301)	0.102	(0.302)	0.101	(0.302)	0.100	(0.300)
Älvsborg	0.046	(0.210)	0.044	(0.204)	0.048	(0.214)	0.046	(0.210)	0.049	(0.215)	0.047	(0.211)
Skaraborg	0.026	(0.160)	0.025	(0.157)	0.028	(0.164)	0.026	(0.161)	0.026	(0.159)	0.025	(0.156)
Värmland	0.033	(0.179)	0.034	(0.181)	0.033	(0.178)	0.033	(0.178)	0.033	(0.179)	0.034	(0.182)

[a] Annual averages over the two calendar years preceding selection.

Table B1: Continued.

	Cohort 1991				Cohort 1990				Cohort 1989			
	Pre-February 28		Post-February 28		Pre-February 28		Post-February 28		Pre-February 28		Post-February 28	
Characteristics	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)	Mean	(Std.dev.)
Örebro	0.029	(0.167)	0.030	(0.170)	0.027	(0.163)	0.028	(0.165)	0.025	(0.157)	0.025	(0.155)
Västmanland	0.028	(0.165)	0.028	(0.166)	0.027	(0.162)	0.026	(0.159)	0.028	(0.166)	0.029	(0.167)
Dalarna	0.032	(0.176)	0.032	(0.176)	0.032	(0.176)	0.032	(0.177)	0.031	(0.173)	0.032	(0.175)
Gävleborg	0.040	(0.195)	0.041	(0.199)	0.039	(0.194)	0.042	(0.200)	0.039	(0.194)	0.042	(0.200)
Västernorrland	0.030	(0.171)	0.031	(0.173)	0.031	(0.174)	0.032	(0.176)	0.032	(0.176)	0.033	(0.179)
Jämtland	0.019	(0.136)	0.019	(0.135)	0.018	(0.132)	0.018	(0.134)	0.019	(0.135)	0.018	(0.134)
Västerbotten	0.038	(0.190)	0.039	(0.193)	0.037	(0.188)	0.037	(0.189)	0.037	(0.189)	0.038	(0.192)
Norrbotten	0.034	(0.181)	0.033	(0.178)	0.037	(0.190)	0.037	(0.189)	0.036	(0.187)	0.038	(0.191)
No obs.	158,460		74,011		153,680		72,611		154,712		74,702	