IFAU – INSTITUTE FOR LABOUR MARKET POLICY EVALUATION

Moral hazard and sickness insurance:

Empirical evidence from a sickness insurance reform in Sweden

Per Johansson
Mårten Palme

WORKING PAPER 2004:10

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Industry, Employment and Communications, situated in Uppsala. IFAU's objective is to promote, support and carry out: evaluations of the effects of labour market policies, studies of the functioning of the labour market and evaluations of the labour market effects of measures within the educational system. Besides research, IFAU also works on: spreading knowledge about the activities of the institute through publications, seminars, courses, workshops and conferences; creating a library of Swedish evaluational studies; influencing the collection of data and making data easily available to researchers all over the country.

IFAU also provides funding for research projects within its areas of interest. There are two fixed dates for applications every year: April 1 and November 1. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The authority has a traditional board, consisting of a chairman, the Director-General and eight other members. The tasks of the board are, among other things, to make decisions about external grants and give its views on the activities at IFAU. A reference group including representatives for employers and employees as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Moral hazard and sickness insurance: Empirical evidence from a sickness insurance reform in Sweden*

Per Johansson†and Mårten Palme‡

23rd August 2004

Abstract

We use a reform of Sweden's sickness insurance system as a source of exogenous variation to analyse the presence of moral hazard. As a result of the reform, the replacement level was reduced from 90 percent of forgone earnings to 65 percent for the first three days; to 80 percent between day 4 and 90; and remained at 90 percent after 90 days. We find that the incidence of work absence decreased due to the decrease in compensation level and that effect on duration is in accordance with moral hazard in the sickness insurance. We estimate the elasticities of the incidence with respect to forgone earning to -1 for males and -0.70 for females.

Keywords: Worker absenteeism, Cox proportional hazard models, regression-discontinuity.

JEL: C41, J22, J28, H53

*We thank Patrik Hesselius, Bertil Holmlund, Laura Larsson Assar Lindbeck, Erik Mellander, Mats Persson, Per Petterson Lidbom and participants in seminars at IFAU, Uppsala University and at the Research Department of Statistics Norway in Oslo for useful comments. We have received financial support from the Swedish Council for Social Research.

†Department of Economics, Uppsala University and Institute for Labour Market Policy Evaluation, SE-751 20 Uppsala, Sweden. E-mail: Per.Johansson@ifau.uu.se.

‡Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: Marten.Palme@ne.su.se.

Contents

1 Introduction **3**

2 Sickness insurance in Sweden **6**

3 Testable implications and empirical modelling **7**

4 Data **11**

5 Results **12**

5.1 Incidence 12
5.2 Duration 17
5.3 Sensitivty analysis 20
5.3.1 The effect of the reform on the prevalence of work absence 20
5.3.2 Effects of persistent heterogeneity 22
5.3.3 Effects of time-varying heterogeneity 23
5.4 Cost elasticities 24

6 Discussion and conclusions **26**

1 Introduction

"Regarding the high absence rate at the Department: Acquiring minor diseases, such as colds or flu, is an act of choice."

Ragnar Frisch, Nobel laureate and founder of the Econometric Society, in an announcement in 1962 on the noticeboard at the Department of Economics, University of Oslo, Norway.

The sickness insurance system (SI), which provides compensation for lost earnings due to work absence on the grounds of temporary illness, is one of the largest social welfare programs in Sweden. In 2002, the expenditure amounted to SEK 48.3 billion, or about 2 percent of GDP. Between 1998 and 2002, the work-absence, and hence the costs of the insurance system, increased by almost 75 percent (see eg The Economist 26 October 2002). This rapid cost increase and subsequent changes in the labor force have dominated the welfare state debate in Sweden for some time. Among the main themes in this debate has been the issue of to what extent insured workers adapt their work-absence behavior, and their work effort, to the generosity of the sickness-insurance scheme, ie, the classical welfare state dilemma of moral hazard.

The requirement for the presence of moral hazard in the sickness insurance system is that the insured workers can affect their utilization of sickness insurance, ie their work absence, and that they moreover act on the basis of economic incentives - the generosity of the sickness insurance - in this respect. This requirement is, however, not as strong as the above quotation by Ragnar Frisch might indicate, ie, that the workers can in fact affect their health status, or that they are engaged in abuse in the sense that they use the sickness insurance rather than taking vacation, although this is sufficient. Since eligibility for compensation from the sickness insurance only requires that the health status of the insured worker is such that it prohibits the worker from doing his or her regular work, and certification from a physician is not needed until the eight day in a work-absence spell, the only requirement is that the workers' *perception* of their health status and their ability to do their regular work is affected by economic incentives.

There are two problems in empirically studying moral hazard in sickness insurance.[1] The first is unobserved heterogeneity. It is well known that

[1] For a general overview of methodological problems in empirical studies on behavior in insurance markets, see Chiappori & Salanié (2000).

workers with preferences for not being absent from work, or with good health and who therefore have no need for frequent absences, tend to be more productive and earn higher wages. Since the sickness insurance system, like most income security programs, replaces a fraction of forgone labor earnings, it is more expensive for high-income workers to be absent from work. In measuring the relationship between the cost of being absent and work-absence, unobserved heterogeneity will, therefore, lead to a downward bias of the estimated effect of the program generosity. Second, adverse selection may result in workers with preferences favoring absences choosing more generous insurance programs. This may, on the other hand, lead to an upward bias of the estimated moral hazard effect.

In this study, we use a reform of Sweden's sickness insurance system which came in force 1 March, 1991 as a source of exogenous variation in the cost of being absent from work. In this reform, the compensation level was reduced from 90 percent of forgone earnings below the social security ceiling to 65 percent during the first three days in a spell of absence and to 80 percent between day 4 and day 89. After day 90, in order to avoid disadvantageous effects of the reform on income distribution, the compensation level remained at 90 percent. We are, thus, able to study how the insured workers change their behavior when a new, less generous policy is introduced. In other word, we can study the moral hazard problem directly.

The design of the reform has at least two testable implications with respect to moral hazard. First, the cost of beginning a work absence period unambiguously increased for all workers due to the reform. To the extent that workers change their work absence behavior as a result of the change in the insurance contract, ie, moral hazard is present, this effect implies that the *incidence* of work absence (ie, the frequency of spells) will decrease. Second, there is an unambiguous increase in the cost of *returning* to work after day 90 as a result of the possibility of commencing a new absence spell with an initially lower compensation level after returning to work. If moral hazard is present, workers on long work absence spells will prolong the duration of these spells. Thus, the two testable implications work in different directions with respect to the overall *prevalence* of work absence.

Our data is based on a random sample of 1,396 blue-collar workers obtained from the Swedish Level of Living Survey (SLLS), matched with information on work absence for each of the 730 days in the years 1990 and 1991. The data on work absence was obtained from registers compiled by the National Social Insurance Board. Since most work-absence periods consist of only a few days, the data set contains rich information on repeated absence spells for most workers in the sample.

The main empirical analysis is based on non-parametric hazard estimates and discrete-time Cox proportional hazard regression models. The data allows us to study incidence and duration separately, which is an essential requirement for analysis of all aspects of the reform mentioned above. Since we compare the behavior of the same individuals before and after the reform, we are able to control for unobserved heterogeneity. An additional advantage is that we do not need to consider potential compositional effects, for example due to the increase in the unemployment rate. Finally, the fact that all workers are covered by the same compulsory insurance programme means that we are able to isolate the effect of moral hazard from adverse selection.

The results are robust with respect to the different empirical methods used and show a remarkable resemblance between what might be expected from the cost changes implied by the reform and actual behavioral change. Firstly, the incidence of work absence declined significantly after the reform. In other words, economic incentives affect work-absence behavior. Secondly, the impact of the reform decreased significantly after seven days in a spell when a certificate from a physician was required for receipt of compensation. This result can be interpreted as evidence for that control is also effective in dealing with moral hazard in the sickness insurance system. Thirdly, the hazard rate from longer then 90 days durations in absence decreased significantly after the reform. Again, this suggests that changes in costs affect work-absence behavior. The point estimates of the elasticities for the incidence of work absence with respect to the compensation level in the sickness insurance system are -1 for males and -0.7 for females.

Previous empirical studies of absenteeism in economics have primarily been oriented towards the conventional labor supply models, considering the labor supply decision conditional on the contracted number of hours of work (see eg Allen, 1981 or Johansson & Palme, 1996 and 2002). Alternatively, previous studies have been oriented towards empirical personnel economics, that studies the effect of different incentive contracts between the employer and the employee (see eg Barmby, Orme & Treble, 1991 and 1995). In this study, however, we explicitly analyze absenteeism as a moral hazard problem in the sickness insurance system. Henreksson & Persson (2003) have a similar focus, but there are several differences between their study and the present one. First, Henreksson & Persson's study uses aggregate data and studies several reforms of the sickness insurance over a long time period. We use individual data and study individual behavior under two different policy regimes. This enables us to isolate the behavioral effect from the composition of the labor force. Secondly, since we have longitudinal data covering individuals, we are able to distinguish between the effect of the reform on the incidence of work absence from the duration of the spells of absence. This distinction proves to be useful for evaluation of the reform. Finally, we are able to quantify the behavioral response in elasticities with respect to the compensation level - an important policy parameter.

The rest of the paper is organized as follows. Section 2 gives a brief description of Sweden's sickness insurance and income tax systems. Section 3 describes the data. Section 4 specifies the testable implications on moral hazard and sickness insurance as a result of the reform and the empirical model. Section 5 presents the results of the estimations. It also contains a sensitivity analysis. First, we analyze the overall change in the prevalence of work absence as a result of the reform. To distinguish the effect of the reform from any trend, or seasonal variation in the work absence rate, we use a differences-in-differences estimator. We compare the change between January/February and March/April 1991 with the corresponding change in 1990. The second difference removes the effect of any seasonal pattern, while the first difference removes any pre-existing trend in work absence, but not the effect of the reform. We also discuss two different sources of change in composition of the sample after the reform. These sources may, in turn, cause bias as a result of unobserved heterogeneity. We use two different stratified estimators to check for these biases, and find that the previous estimates, based on the Cox regressions, are robust. Section 6 provides a discussion of the results and conclusions.

2 Sickness insurance in Sweden

Sweden has a compulsory national sickness insurance. It is financed by a proportional payroll tax and replaces earnings forgone due to temporary health problems that prevent the insured worker from doing his regular job. Sickness insurance is administered by local insurance offices. Since it is very hard to judge whether or not a worker is able to perform his or her regular job, monitoring abuse is very light during the first seven days of a sickness period. However, a certificate from a physician is required for entitlement to sickness insurance payments as from the eigths day in a sickness period.

The compensation level - the proportion of earnings paid to the worker by the insurance system - has been modified on several occasions in recent years. In the major reform covered by our longitudinal data - which was implemented on 1 March 1991 - the compensation level was reduced from 90 percent of labor earnings below the social security ceiling[2] from the first day in a sickness spell, to 65 percent in the first three days in a spell and to 80 percent from day four to day 90.

The 1991 reform of the sickness insurance system was the most important of several budget cuts proposed in the early spring of 1991. The reason for these cuts was an increasing budget deficit, excess aggregate demand and a political ambition to defend the fixed exchange rate for the Swedish currency (SEK). The increase in the unemployment rates and the general downturn in the Swedish economy commenced several months later, in the autumn after the 1991 parliamentary elections. The background to the new compensation rates was that, in the 1988 parliamentary election campaign, the Social Democratic government party had guaranteed not to introduce a "waiting day"[3] in the sickness insurance system.

In addition to the compulsory national sickness insurance, most Swedish workers are covered by negotiated sickness insurance programs regulated in agreements between the labor unions and the employers confederations. In general, these insurances replaced about 10 percent of forgone earnings. We have, however, chosen to not consider these programs in our analysis. The reason for excluding them is that they differ somewhat between different groups of workers and it is complicated to assign the right insurance to each worker in the sample. In addition, our policy variable is the compensation level in the national sickness insurance and since the negotiated insurances were not affected by the reform excluding them should not affect our results.

3 Testable implications and empirical modelling

As noted in the introduction, the design of the reform of the sickness insurance implies that the change in the cost of absence from work depends on the state of the insured worker, ie, whether he is absent or not and, if absent, for how long the work absence period has lasted. We define the *direct* cost of being absent from work (ie entering into absence or remaining in absence) as the percentage share of earnings not replaced by the sickness insurance. This direct cost changed from 10 to 35 percent during the first

[2] In 1995, about 6.7 percent of all insured workers had labor earning above the social security ceiling. For a description of the construction and indexation of the social insurance ceiling see e.g. Palme & Svensson (1999).

[3] The first day in each sickness spell without any compensation from sickness insurance.

three days in a spell, but remained at 10 percent from day 91 in a absence spell.

We define the cost of *returning* to work as the difference in the percentage share of earnings not replaced by the sickness insurance between remaining in an ongoing absence period and the corresponding share in a new absence spell after returning to work. Since the replacement level is 90 percent irrespective of spell length before the reform this cost is zero in all states. After the reform it is zero only if the duration of the absence spell is less than 3 days; if the duration is between 3 days and 90 days the cost of returning is 15 percent; and in spells longer than 90 days it is 25 percent.

A for practical purposes attractive feature of the reform is that these cost changes are independent of the income tax rate. This means that the costs defined above are not affected by the 1991 income tax reform. The reform has the following implications for the direct cost and cost of returning to work:

i. There is an unambiguous *direct* cost increase in beginning a work absence period due to the reduction in the compensation level for the first 90 days of the period.[4]

ii. For work absence spells of less than 91 days, there is an ambiguous effect of the reform. First, there is the increased *direct* cost of continuing a spell. Second, there is an increased cost of *returning* to work.

iii. For absence periods longer than 90 days, there is no change in the direct cost of absence. There is, however, a 25 percent point increase in the cost of returning to work. The reform, thus, implied an unambiguous *decrease* in the relative cost to remain absent in such spells.

Since the unambiguous implications (i) and (iii) work in opposite directions the a priori effect of the reform on the prevalence of work absence is ambiguous.

We use different strategies for testing the behavioral response to these implications empirically. The first strategy is based on non-parametric estimates of hazard rates and the survival functions (see eg Kaplan & Meier, 1958). Although these estimates have the advantage of not imposing any functional-form assumptions, the technique has a disadvantage when there is a need to control for "confounders" (for example, it does not allow us

[4] One can note that the reform, not only gave an unambiguous relative cost increase, but also gave an unambiguous cost increase absolutely.

to control for the increase in the unemployment rate which occurred in the autumn of 1991 or controlling for unobserved heterogeneity). In order to do this, we will use (discrete time) Cox regression models.

For the incidence we use the following specification:

$$\lambda_1(t) = \lambda_0(t)e^{\delta I^R}, \tag{1}$$

where $\lambda_0(t)$ is the baseline hazard (ie the hazard before the reform) and I^R is a step function, taking the value 1 after the reform and 0 before.

For the duration (ie absence spells) we use the following specification:

$$\lambda_1(t) = \lambda_0(t)e^{I^R(1-3)\beta_1 + I^R(4-7)\beta_2 + I^R(8-90)\beta_3 + I^R(91-)\beta_4}, \tag{2}$$

where $\lambda_0(t)$ is the baseline hazard, $I^R(j-k)$ are impulse functions, such that $I^R(j-k) = I^R\mathbf{I}(j \leq t \leq k)$ where $\mathbf{I}(.)$ takes the value one if the argument within the parenthesis is true, and $I^R(91-) = I^R\mathbf{I}(91 \leq t)$ is a step function. We also include $\mathbf{I}(1-3)$, $\mathbf{I}(4-7)$ and $\mathbf{I}(8-90)$ in the specification. The interpretation of β_1, β_2, β_3 and β_4 is thus the change in hazard rate caused by the reform. If moral hazard is present then these coefficients can be expected to differ from zero. Because of the control of absence from work we expect that $\beta_3 < \beta_2$, even though the cost change due to the reform is the same. Also, β_4 is expected to be less than zero.

As described in the previous section, the unemployment rate increased substantially in the autumn of 1991. Figure 1 shows the weighted average of the county-level unemployment rate, together with the aggregate work-absence rate in the sample. There is a literature in the field which suggests a correlation between general conditions in the labor market and absence rate as a result of the disciplining effect of high unemployment (see eg Arai & Skogman Thoursie, 2001, Henreksson & Persson, 2003, Lantto, 1991). To control for this potential effect we use monthly data on unemployment levels for each of Sweden's 24 counties (local labor markets) matched with the data.

Figure 1 also reveals a seasonal pattern in work absence and the immediate drop in the prevalence just after the reform. Since the 1991 reform of the sickness insurance system did not occur until March 1, this is a potential problem. To balance the pre-reform and post-reform samples, we in the estimations exclude spells beginning in January or February in both 1990 and 1991.[5]

[5] The literature on the labor supply effects of income tax reform contains some discus-

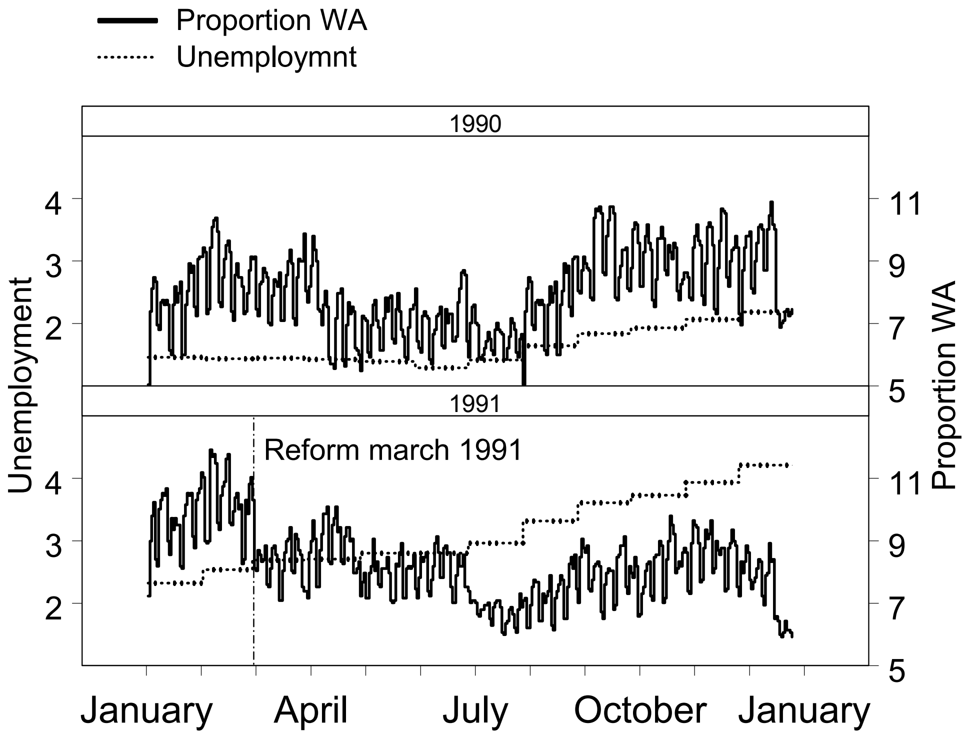


Figure 1: Prevalence of work absence in the sample and the monthly average unemployment rate in 1990 and 1991.

In the empirical analysis, we use an exact maximum likelihood estimator in discrete time (see Kalbfleich & Prentice, 1980, Chapter 4). In addition, to control for individual heterogeneity we use a stratified partial maximum likelihood estimator (see eg Lancaster, 1990, Chapter 9 or van den Berg 2000, Section 6).[6]

For the exact maximum likelihood estimator in discrete time, the baseline hazard for the incidence is specified using a dummy variable for each day in a work spell. In the case of the duration in work absence, the baseline hazard is piecewise constant.[7]

4 Data

We use the 1991 Swedish Level of Living Survey (SLLS). The SLLS is a micro data set that contains information compiled from interviews and official public registers for a random sample of about 6,000 individuals. This survey is described in detail in Fritzell & Lundberg (1994). Data on the dependent variable - absence from work compensated by the sickness insurance - was obtained from the National Social Insurance Board by matching with the SLLS sample. As the data was collected from registers for actual transactions to insured individuals, there are likely to be much fewer measurement errors as compared with self-assessed data.

We restricted the sample to blue-collar workers aged between 20 and 64 who were employed during 1990 and 1991. The final sample consisted of 1,396 individuals (738 males and 658 females).

Table 1 provides descriptive statistics on the incidence and duration of work absence, subdivided into the pre-reform and post-reform period,

---

sion of potential "dynamic effects" in the sense that households may indulge in "intertemporal substitution" and adjust their labor supply behavior to the most favorable income tax regime (see e.g. Hausman & Poterba, 1987). One might conceive a similar behavioral response to the reform studied in this paper, since the compensation level did not change for spells initiated before March 1 and the reform was announced about two weeks before it was implemented. This could imply that some of the behavioral changes recorded in the data would not be permanent but rather short-term effects of anticipated changes in compensation levels. However, the prevalence of absence from work shown in Figure 1 does not reveal any increased absence in February 1991 in addition to seasonal pattern. As a result, we will not consider this potential effect in the empirical analysis.

[6] Note that we do not have the common problem of using the partial maximum likelihood estimator with tied duration here, since no individuals have the same duration before and after the reform.

[7] The following specification is used: a dummy variable for each day for days 1 to 3 and dummy variables for days 4 -7, 8-14, 14-21, 22-28, 29-42, 43-56, 57-70 and 71-84.

Table 1: Descriptive statistics. Before and after the reform of sickness insurance. Male and female workers separately.

	Males		Females	
	Before Reform	After Reform	Before Reform	After Reform
Average spell length:				
Work	93	114	87	108
Work Absence	8.8	11	9.7	12
Number of work spells	2,227	1,812	2,069	1,666
Number of absence spells	1,605	1,182	1,550	1,160
Proportion 1-3 days absence	0.48	0.48	0.54	0.57
Proportion 4-6 days absence	0.24	0.24	0.21	0.16
Proportion 7-89 days absence	0.26	0.25	0.23	0.23
Proportion 90- days absence	0.02	0.03	0.02	0.04

respectively. It is apparent from the table that the mean work spells of females are shorter than the mean work spells of the males. The opposite is true for the work absence spells. We can also see that both means are higher after the reform, hence there is a reduction in incidence and a countervailing effect on prevalence, due to an increased duration of absence from work. When it comes to the distribution of the days of absence, we can see that the female workers seem to change their work absence behavior more than the males and that the longer spells (ie, more than 90 days) are more prevalent for both men and women after the reform.

5 Results

5.1 Incidence

Let us begin by analyzing the first implication of the reform - the increased cost of leaving a work spell. Figure 2 shows Kaplan-Meier estimates of the survival functions in the work spells before and after the reform respectively. These estimates show that there is a distinct impact of the reform on decreasing survival rate for both males and females. The change is highlighted in Figure 3, which shows the difference between the post-reform and pre-reform survival functions along with a 95 percent confidence interval. This figure shows that there is a statistically significant reduction (ie an increase

Table 2: Discrete-time Cox proportional hazard model estimates ($Est.$) and standard errors (se) of the effect of the sickness insurance reform on incidence of work absence.

	Males				Females			
	$Est.$	se	$Est.$	se	$Est.$	se	$Est.$	se
I^R	-0.316	0.005	-0.310	0.007	-0.211	0.005	-0.240	0.007
Unemployment	-		-0.114	0.010	-		-0.077	0.003
Unemployment^2	-		0.022	0.002	-		0.018	0.001
County factor	No		Yes		No		Yes	
Log likelihood	-9668.6		-9646.6		-9929.2		-9892.9	
$\chi^2(25)$; $p - value$	44.0; 0.01				72.6; <0.001			

Note: χ^2 statistics and $p - value$ for likelihood ratio test of joint significance of local labor market unemployment rate and county factors.

in the incidence) in the survival rate for both males and females.

A further non-parametric analysis is presented in Figure 4, where the log difference of the pre-reform and post-reform hazards are plotted for different work durations, along with predictions from a linear regression. The intercepts from these regressions are -0.36 for male and -0.34 for female workers respectively, ie, a 36 and 34 percent increase in average incidence (i.e and increase in the for males and females, respectively).[8]

Table 2 shows the results from the estimates of the Cox regressions. We use two different specifications - one which only includes indicators related to the reform and one which also includes a quadratic specification for the unemployment rate in the county (local labor market), as well as indicators for each county (ie a county factor).[9] The χ^2 statistics and the p-value provided in the table show that local labor market unemployment rate together with the county factors are jointly significant. The moral hazard effect estimates are statistically significant and the estimates are -0.310 and -0.240 for the males and females, respectively.

These estimates confirm the results of the non-parametric analysis: There is a significant moral hazard effect on incidence. This effect is robust, although the magnitude changes somewhat when we include controls for counties and the unemployment rate in the local labor market. Males react significantly more strongly to the reform than females.[10]

[8] The intercepts (se) are -0.36 (0.08) for males and -0.34 (0.10) for females. The corresponding slopes (se) are 0.001 (0.001) and 0.000 (0.001), respectively.

[9] Sweden is administratively divided into 24 counties.

[10] A 95 percent confidence interval for the difference is $[0.05, 0.09]$.

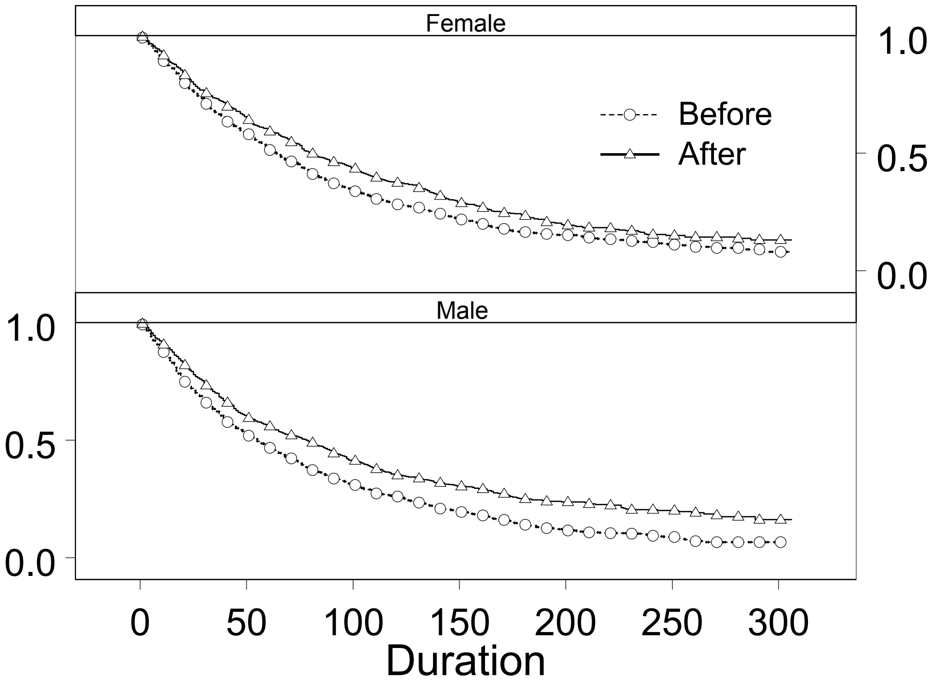


Figure 2: Kaplan-Meier survival function estimates of incidence of work absence before and after the sickness insurance reform.

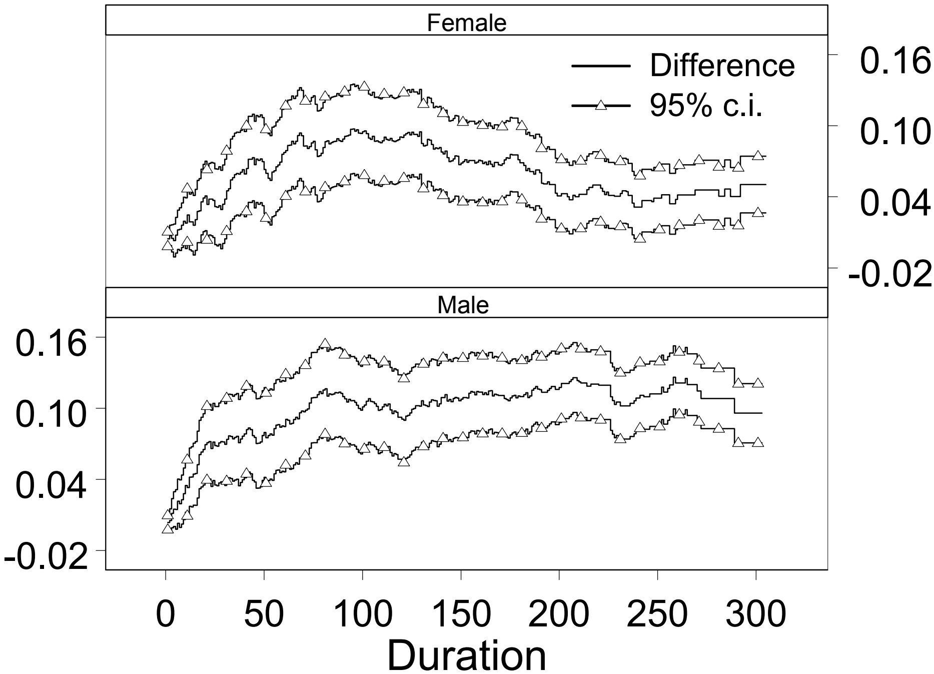
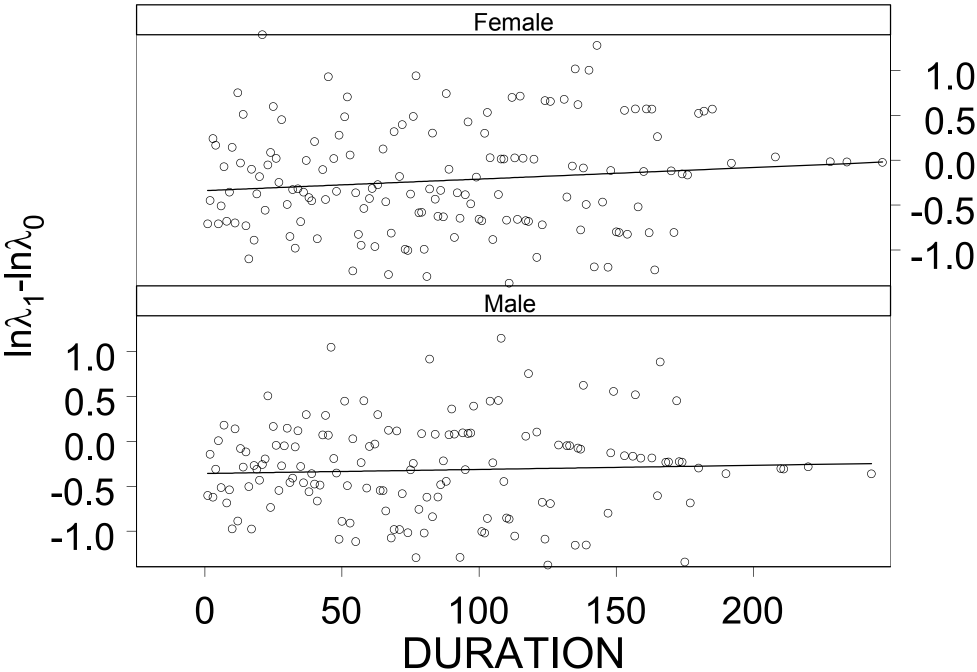


Figure 3: Before and after the reform differences of survival functions estimates. Incidence of work absence



`

Figure 4: Scatter plots of the differences between the log incidence of work absence before and after the reform and also prediction of these differences based on a linear regression model.

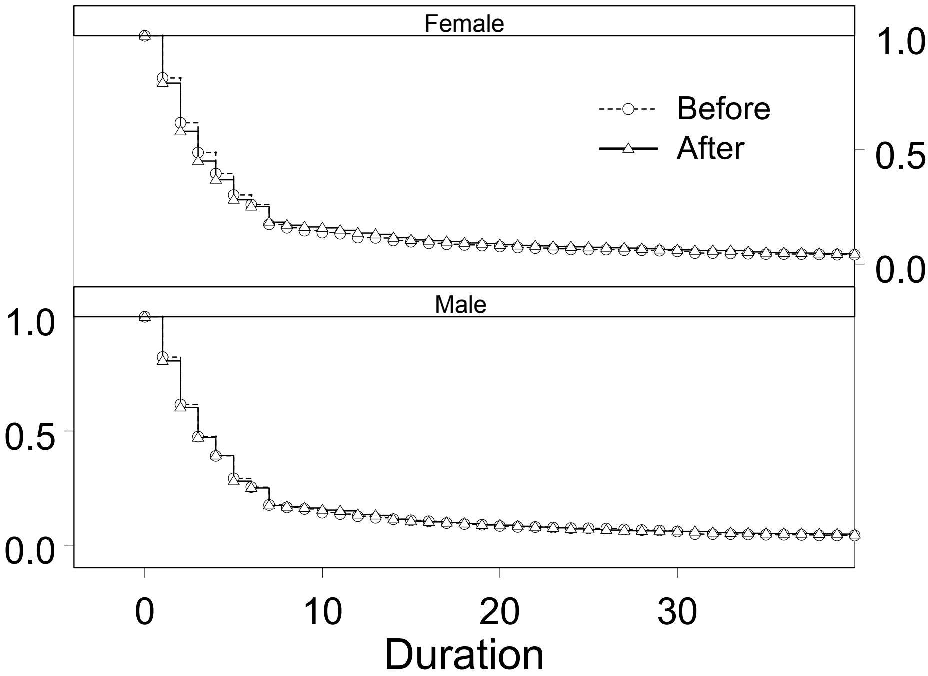


Figure 5: Kaplan-Meier survival estimates of the duration of work absence spells before and after the sickness insurance reform.

5.2 Duration

Turning to the duration in work absence, we expect a more complex behavioral response to the reform. Firstly, there is an ambiguous effect for short durtions. This is a result of the increased cost of remaining absent from work versus a possibly decreased income for returning to work due to the possibility of beginning a new spell, during the entire initial 90 day period. Secondly, whatever the net effect during the first 7 days, we expect a decreased effect after day 7 due to the requirement of a certificate from a physician. Finally, we expect a decreased hazard for long durations due to the unambiguous decrease in expected income of returning to work after

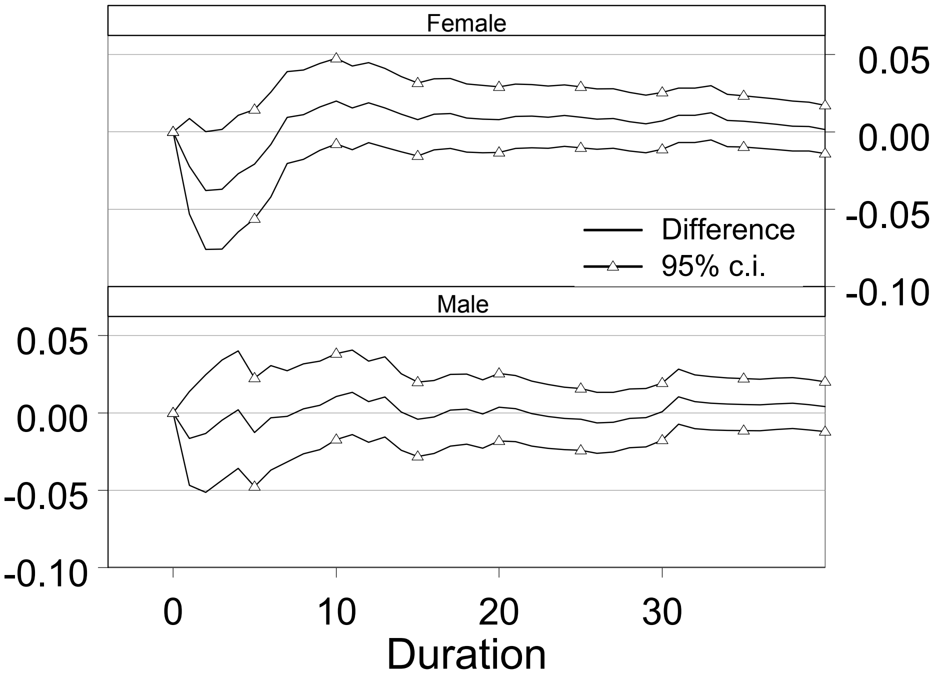


Figure 6: Kaplan-Meier estimates of the difference between the duration of the work absence spells before and after the reform.

day 90.

The estimated survival functions for the first 40 days in a spell of absence are displayed in Figure 5 and the pre-reform and post-reform difference (including a 95 percent confidence interval) is shown in Figure 6. As is apparent in Figure 6, there is no significant change in the duration of the absence spell for either males or females.[11] In the case of females, however, there is a marked decrease the survival function up to day 6 in a spell.

The results of the Cox regressions are shown in Table 3. Again, we use

[11] There is no significant impact of the reform on survival functions after 40 either, when males and females are considered separately. However, when the samples are pooled, there is a significant increase in the duration of the spells 150 days after the reform.

Table 3: Discrete-time Cox proportional hazard regression estimates (*Est.*) and standard errors (*se*) of the effect of the reform on the duration in work hazard (hazard of ending a work absence spell).

	Males				Females			
	Est.	*se*	*Est.*	*se*	*Est.*	*se*	*Est.*	*se*
$I^R(1-3)$	0.062	0.029	0.099	0.033	0.068	0.028	0.077	0.032
$I^R(4-7)$	0.015	0.300	0.040	0.045	0.050	0.031	0.074	0.046
$I^R(8-90)$	-0.022	0.021	0.006	0.029	-0.101	0.020	-0.095	0.027
$I^R(91-)$	-0.127	0.030	-0.100	0.036	-0.639	0.027	-0.591	0.035
Unemployment			0.150	0.028			-0.013	0.027
Unemployment^2			-0.032	0.004			0.017	0.042
County factor	No		Yes		No		Yes	
Log likelihood	-4362.5		-4350.8		-4256.7		-4241.9	
$\chi^2(25)$; $p-value$	23.32; 0.57				29.6 ;0.27			

Note: The baseline hazard is specied as picewise constant. Indicators for 1-3, 4-7 and 8-90 in a spell are also included in the specification. χ^2 statistics and $p-value$ for likelihood ratio test for joint significance of *local labor* market unemployment rate and county factors.

two different specifications - with and without controls for the county and county level unemployment rate. However, unlike the model for incidence of work absence, the χ^2 statistics and the p-values show that we cannot reject exclusion of the unemployment rate and county effects from the specification. The results confirm that the hazard from short absence spells (shorter than 8 days) increased after the reform. This means, following the anticipated behavioral effects outlined in Section 3, that the effect of the increased cost dominates the effect of decreased expected incomes of going back to work and being exposed to the risk of embarking on a new spell of absence. This result applies to both gender groups and is robust to alternative specifications. The effect is, however not statistically significant for the 4-7 days for the males.

For spells of between 8 and 90 days we find (as expected) reduced effects of the reform on the hazard from work absence. In the female sample, this effets is significantly negative. Finally, Table 3 shows that the reform had a significantly negative effect on the hazard on spells of more than 90 days - this applies to both gender groups and is robust in the alternative specifications. This result gives support to the hypothesis that the decrease in the expected income of returning to work has an impact on worker behavior.

5.3 Sensitivty analysis

5.3.1 The effect of the reform on the prevalence of work absence

To be able to distinguish the effect of the reform from any trend in the work absence rate we exploit the discontinuous nature of the reform. The idea behind this strategy is that behavioral changes that can be referred to trends, or gradual changes in the society, cannot be recorded as "jumps" or discontinuous changes in the data,[12] as opposed to the effect of policy interventions that are implemented at a particular date in time. To measure this effect it would be possible to take the difference between the prevalence immediately before and after the reform, eg during January and February 1991 and March and April the same year. However, since there is an apparent seasonal pattern in the work absence rate (see Figure 1) such estimate would be distorted. To deal with this problem, we use a differences-in-differences estimator, where the difference between January/February and March/April 1990 is used as a "control". Formally, the estimator is defined as

$$DD = (\overline{m}_{jf}^{91} - \overline{m}_{ma}^{91}) - (\overline{m}_{jf}^{90} - \overline{m}_{ma}^{90}), \tag{3}$$

where $\overline{m}_{ma}^{91}$ and $\overline{m}_{ma}^{90}$ are the mean prevalence in March and April 1991 and 1990, respectively and $\overline{m}_{jf}^{91}$ and $\overline{m}_{jf}^{90}$ are the mean prevalence in January and February 1991 and 1990, respectively.

Figure 7 shows the samples day-by-day work absence rate for the period January to May 1990 (left panel) and the corresponding period in 1991 (right panel) including the date of the reform in March 1. It is apparent that there is a discontinuous shift in the work absence rate around the date of the reform, which did not take place around that date the previous year. The overall level of work absence is, however, somewhat higher in 1991.

Table 4 shows the estimated components and the overall result from the differences-in-differences estimator specified in equation (3) both for the males and females and separately for the genders. The results show a significant decrease in the work absence rate in the sample as a result of the reform. The effect is much stronger for the males than for the females.

[12] Angrist & Krueger (1999) uses the famous quote "Natura non facit saltum" or "Nature does not make jumps" from Marshall's *Principles of Economics*, as a motivation for this observation.

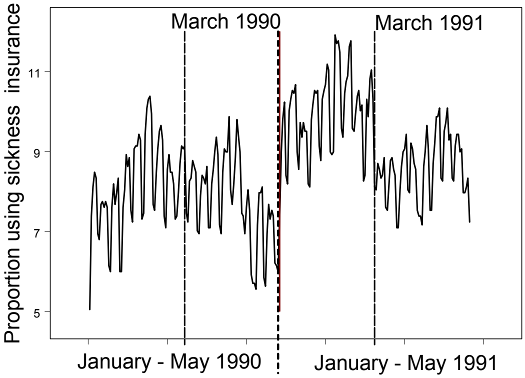


Figure 7: Daily work absence rate January 1 to April 30, 1990 and January 1 to April 30, 1991. Reform date in March 1, 1991 is marked. Males and Females.

Table 4: Estimated means and standard errors (se) for prevalence of work absence for different sub-periods.

Jan/Feb 1990	Mar/Apr 1990	Jan/Feb 1991	Mar/Apr 1991	DD
		Males and Females		
8.13	7.75	9.94	8.62	0.93
(0.15)	(0.14)	(0.14)	(0.11)	(0.27)
		Males		
7.31	7.16	8.40	6.69	1.56
(0.16)	(0.13)	(0.14)	(0.11)	(0.27)
		Females		
9.07	8.41	11.68	10.81	0.21
(0.16)	(0.18)	(0.18)	(0.15)	(0.33)

5.3.2 Effects of persistent heterogeneity

As discussed in the introduction, work absence is likely to be heterogenous in any group of workers. It is also likely that individual incidence of work absence and the duration in absence are correlated. This means that people with high incidence of work absence also may have long work absence durations. If this is the case and if moral hazard exists the composition of the work and work-absence samples may be different between the two periods even though our sample of individuals is the same in the two periods.

If moral hazard exists then would the decreased cost of terminating a work-absence spell imply, due to heterogeneity, that low-incidence workers will be overrepresented in the post-reform sample of work spells, compared to the pre-reform situation, and also that low-hazard workers will be overrepresented in the sample of work-absence spells. These compositional effects would weaken the moral hazard effect for the incidence and would lead to an exaggeration of the moral hazard effect for the duration in work absence. Since we have data on repeated spells for individuals, we are able to evaluate whether this neglected heterogeneity is a serious problem.

Adding fixed effects to the above models would results in biased estimates (see eg Topel & Ward, 1992). To avoid obtaining biased estimates, we use the stratified partial maximum likelihood estimator. This type of estimator has been used in several previous empirical studies (see eg Lindeboom & Kerkhofs, 2000, or Ridder & Tunelı, 1999), but here we use the individual as a stratification unit rather than group assignment, which is the stratification unit used in the previous empirical studies mentioned above.[13]

The results of this analysis, both for the incidence and the duration, are shown in Table 5. The estimates of the reform effect on the incidence of work absence are very similar to those obtained in the Cox regression models, although the precision is, as expected, somewhat inferior. For the duration, the pattern is similar to the one found in the Cox regression analysis. The precision of the estimates is very low, however, and the estimates from the two estimators are never statistically significantly different from zero. We conclude that the estimates from the discrete-time Cox regressions are robust with respect to this source of compositional change.

[13] The stratified approach needs at least one spell before the reform and one spell after the reform. If this selected sample reacts different to the change in cost than the original sample of blue collar workers also this estimator would yield biased estimates of the moral hazard effect for the blue collar workers.

Table 5: Estimates (*Est.*) and standard errors (*se*) for the stratified proportional hazard analysis of the effect of the reform on incidence and duration using the individual as stratification unit.

	Males		Females	
	Est.	*se*	*Est.*	*se*
			Incidence	
$\mathbf{I}^R$	−0.342	0.059	−0.258	0.060
$\chi^2(1)$; $p-value$	34.3; < 0.00			18.8; < 0.00
			Duration	
$I^R(1-3)$	0.057	0.074	0.060	0.075
$I^R(4-7)$	0.066	0.117	−0.046	0.115
$I^R(8-90)$	−0.115	0.198	−0.392	0.235
$I^R(91-)$	−0.115	30.667	0.747	28.036
$\chi^2(7)$; $p-value$	2470; < 0.00			2364; < 0.00

Note: Factors for spells of 1-3, 4-7 and 8-90 days are included in the specification. χ^2 statistics and p-value for likelihood ratio test for joint significance of all variables included.

5.3.3 Effects of time-varying heterogeneity

The bias discussed above relates to unobserved and heterogenous propensity to be absent that is *persistent* over the time period under study (eg, permanent differences in health status). However, *time-varying* heterogeneity may also affect our estimates. This second form of compositional effect concerns unobserved and time-varying health effects. Work absence may be seen as a form of "investment" in one's health.[14] The decreased incidence after the reform may, following this logic, have a detrimental effect on the general health status and result in longer work-absence spells. This argument has been used extensively in the Swedish public policy debate on the sickness insurance system.

Since we have information on in which month each of the work-absence spell in our data begun we have the possibility of evaluating this last concern regarding biased inference. The logic is as follows: if there is a detrimental effect on health status due to the reform, it should become apparent at some time after the reform. We should not expect a detrimental effects

[14] See Paringer (1983) for a discussion on work absence as an investment in individual health.

Table 6: Estimates ($Est.$) and standard errors (se) from discrete time Cox proportional hazard model of the effect of the reform on the duration in work absence using work absence spells that starts in March and April in either 1990 and 1991.

$I^R(.)$	$1-3$		$4-7$		$8-90$		$91-$	
	$Est.$	se	$Est.$	$s.e$	$Est.$	$s.e$	$Est.$	se
				Males				
March ($n = 255$)	0.223	0.105	0.058	0.124	0.132	0.070	−0.372	0.073
April ($n = 248$)	0.154	0.102	−0.041	0.142	−0.267	0.059	−0.032	0.093
				Females				
March ($n = 267$)	0.072	0.106	0.004	0.128	−0.406	0.053	−0.363	0.079
April ($n = 246$)	0.121	0.107	−0.005	0.118	−0.502	0.063	−0.019	0.096

Note: Factors for spells 1-3, 4-7 and 8-90 days included in the specification.
n is the sample size.

immediately after the reform. Hence reform-effect estimates based on work-absence entrants in March and April in either year should not be affected by detrimental effects.

To test for this second source of compositional change, we now use the month of entry into a work-absence spell as a stratification unit, and separately for the first two month immediately after the reform-estimate discrete-time Cox proportional hazard models.

The results are shown in Table 6. Again, it is apparent that the results are very similar to the original ones although the precision is inferior. This test, thus, shows that the effect of the reform on the duration is most likely to be due to moral hazard rather than detrimental effects on the health of the insured workers.

5.4 Cost elasticities

In the previous analysis, both in the Kaplan-Meier and the Cox regression, we were able to make inferences regarding the direction of the behavioral response to different cost changes implied by the reform. We could not, however, say much about the magnitude of these effects. To obtain a summary measure of these magnitudes of the behavioral response, we use the estimates presented in the previous sub-section to obtain elasticity estimates with respect to the compensation level in the sickness insurance system. We obtain separate measures for the incidence and duration of work absence.

The income of being absent is affected by three components: the wage rate, the compensation level in the sickness insurance scheme and the marginal income tax rate. However, the most relevant cost is the cost of forgone income if absent. Income from the sickness insurance is taxable and for this sample of blue collar workers, with income below the social security ceiling, the wage rate and marginal income tax rate do not matter for this relative cost. Thus, pre the reform the foregone earnings is 90 percent of the income if working and post the reform the foregone earnings is 65 percent of the income if working the first three days in a spell. This implies a 28 percent decrease in the foregone earnings from 1990 to 1991.

To measure the changes in the incidence of work absence, we use both the non-parametric estimates of the hazards and the differences obtained from the Cox regression models. The first measure is thus calculated as

$$\widehat{\varepsilon} = \left(\frac{\widehat{\lambda}(pre) - \widehat{\lambda}(post)}{\widehat{\lambda}(pre)} \right) / \Delta,$$

where $\widehat{\lambda}(pre)$ and $\widehat{\lambda}(post)$ is estimated under the assumption of constant hazard for incidence and Δ is the relative decrese in forgone earnings due to the reform. The second elasticity measure, which is based on the Cox proportional hazard regersssion, is simply $\widehat{\varepsilon} = (1 - \exp(\widehat{\delta}))/\Delta$. Standard errors (se) are calculated using Gauss approximations.

It is somewhat more complicated to calculate the elasticities for the duration. This is because the cost changes and the estimates of the behavioral change both depend on the length of the spell. Under the assumption of piecewise constant hazards we calculate three sets of elasticities for the absence spells: for the first three days in a spell, for spell lengths of between 4 and 7 days and between 8 and 90 days. These elasticities are, for the second elasticities, calculated as $\widehat{\varepsilon}_j = (1 - \exp(\widehat{\beta}_j))/\Delta$, $j = 1, 2$ and 3. For spell lengths shorter than 4 days the change in forgone earnings is the same as for the work state (28 percent). However, for spells longer than 4 days and shorter than 90 days forgone earnings decrease by 11 percent.

Table 7 shows the elasticity estimates. These estimates indicate that the incidence elasticities obtained from the Cox regression models are somewhat smaller than those obtained from the non-parametric analysis. As was also apparent in the results presented in Section 5.1, the behavioral response to the reform is somewhat larger among male workers. Comparing the precision in the estimates of the elasticities for incidence with those for the duration, it can be seen that the precision is superior for incidence elasticities.

One limitation on this manner of calculating elasticities for the incidence

Table 7: Estimates of elasticity with respect to the cost of being absent on the incidence and hazard of work absence. Estimates from the non-parametric (Nonpar) analysis and the Cox regressions, respectively. Standard errors (se) are calcuated by using the delta method.

	Males				Females			
	Nonpar		Cox		Nonpar		Cox	
	$\widehat{\varepsilon}$	*se*	$\widehat{\varepsilon}$	*se*	$\widehat{\varepsilon}$	*se*	$\widehat{\varepsilon}$	*se*
Incidence	-1.05	0.02	-0.93	0.02	-0.70	0.02	-0.72	0.03
Duration, 1-3 day spells	0.04	0.17	0.23	0.11	0.25	0.17	0.25	0.11
Duration, 4-7 day spells	0.02	0.54	0.14	2.77	0.39	0.52	0.47	0.30
Duration, 8-90 day spells	0.19	0.39	-0.20	0.19	0.17	0.38	-0.87	0.16

is that changes in the expected income of remaining absent in the post-reform regime are not considered.

6 Discussion and conclusions

An unambiguous conclusion to be drawn from this study is that there is moral hazard in Sweden's sickness insurance system. This is seen primarily in the reduction in the incidence of work absence after the reform, but also in the decline in the hazard rate from the durations longer than 90 days. The result indicating that the effect of the reform diminished in spells of more than seven days, which required a certificate from a physician, may be interpreted as supporting the notion that control counteracts moral hazard, which, indirectly, shows that moral hazard is present. This result may, however, also be due to the increasing importance, in these longer spells, of the higher cost of going back to work after the reform.

The presence of moral hazard shows that there is a policy trade-off between the advantageous income-distribution properties of generous compensation levels in the sickness insurance system and disadvantageous behavioral responses to such a policy. This is particularly true of the policy reform studied in this paper. The reason for not changing the compensation level in spells of more than 90 days was to avoid negative effects on a disadvantaged group. However, our results indicate that the reform had disincentive effects on this group of workers defined by long durations only.

The estimates of the elasticities for the incidence of the work absence with respect to the compensation level in the sickness insurance system (between -1.05 and -0.93 for males and -0.70 for females) are somewhat higher, particularly for males, than elasticities estimated in traditional labor supply studies of the effect of marginal tax rates. These elasticities are typically around 0.1 for male and 0.2 for females in Swedish data, see eg Aronsson & Walker, 1997, for an overview). The estimates imply that if compensation levels are increased by 10 percent, the incidence of work absence will increase by 9 to 10 percent for males and by 7 percent for females. The pattern of higher labor supply elasticities for income security programmes than income taxes did also emerged in a recent overview by Krueger & Meyer (2002).

References

Allen, S. G. (1981). "An Empirical Model of Work Attendance", *Review of Economics and Statistics* **63**, 77-87.

Angrist, J. D. and A. B. Krueger (1999). "Empirical Strategies in Labor Economics", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics vol 3A.* Amsterdam: North-Holland.

Arai, M. and P. Skogman Thoursie (2001). "Incentive or Selection in Cyclical Absenteeism" Working Paper 2001, Trade Union Institute for Economic Research.

Aronsson, T. and J. R. Walker (1997). "The Effect of Sweden's Welfare State on Labor Supply Incentives", in R Freeman, B. Swedenborg and R. Topel (eds.) *The Welfare State in Transition*, Chicago: Chicago University Press.

Barmby, T., C. Orme and J. Treble (1991). "Worker Absenteeism: An Analysis using Micro Data" *The Economic Journal* **101**, 214-229.

Barmby, T., C. Orme and J. Treble (1995). "Worker Absence Histories: A Panel Data Study", *Labour Economics* **2**, 53-65.

Chiappori, P.A. and B. Salanié (2000) "Testing Contract Theory: A Survey of Some Recent Work", invited lecture, World Congress of the Econometric Society in Seattle, 2000.

Fritzell, J. and O. Lundberg (1994). *Vardagens Villkor. Levnadsförhållanden i Sverige under Tre Decennier (*Everyday Situations. Living Conditions in Sweden during Three Decades*)*, Stockholm: Brombergs Förlag.

Hausman, J. and J. Poterba (1987). "Household Behavior and the Tax Reform Act of 1986", *Journal of Economic Perspectives* 1, 101-119.

Henreksson, M. and M. Persson (2003). "The Effects of Sick Leave of Changes in the Sickness Insurance System", *Jornal of Labor Economics, forthcoming.*

Johansson, P. and M. Palme (1996). "Do Economic Incentives Affect Worker Absence? Empirical Evidence Using Swedish Data", *Journal of Public Economics* **59**, 195-218.

Johansson, P. and M. Palme (2002). "Assessing the Effect of Public Policy on Worker Absenteeism", *Journal of Human Resources* **37**, 381-409.

Kalbfleich, J. D. and R. L. Prentice (1980). *The Statistical Analysis of Failure Time Data*, New York: Wiley.

Kaplan, E.L. and P. Meier (1958). "Nonparametric Estimation from Incomplete Observations", *Journal of the American Statistical Association* **53**, 457–481.

Krueger, A. B. and B. D. Meyer (2002). "Labor Supply Effects of Social Insurance", in A. Auerbach and M. Feldstein (eds.) *Handbook of Public Economics 4*. Amsterdam: North-Holland.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data,* Cambridge: Cambridge University Press.

Lantto, K. (1991). "Optimal Deterrents to Malingerence", Ph.D. Thesis, Department of Economics, Stockholm University.

Lindeboom and Kerkhofs (2000). "Multistate Models for Clustering Duration Data - An Application to Workplace Effects on Individual Sickness Absenteeism", *Review of Economic and Statistics* **82,** 668-684.

Palme, M. and I Svensson (1999). "Social Security, Occupational Pensions, and Retirement in Sweden", in J. Gruber and D. Wise (eds.) *Social Security and Retirement around the World.* Chicago: Chicago University Press.

Paringer, L. (1983). "Women and Absenteeism: Health or Economics?", AEA Papers and Proceedings **73**, 123-127.

Ridder, G. and İ. Tunelı (1999). "Stratified Partial Likelihood Estimation", *Journal of Econometrics* **92,** 193-232.

Topel, R. H. and M. P. Ward (1992). "Job Mobility and Careers of Young Men", *Quarterly Journal of Economics*, **107**, 153-179.

van den Berg, G. J. (2000). "Duration Models: Specification, Identification and Multiple Durations", in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics 5.* Amsterdam: North-Holland.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2004:1 Björklund Anders, Per-Anders Edin, Peter Fredriksson & Alan Krueger "Education, equality, and efficiency – An analysis of Swedish school reforms during the 1990s"

2004:2 Lindell Mats "Erfarenheter av utbildningsreformen Kvalificerad yrkesutbildning: ett arbetsmarknadsperspektiv"

2004:3 Eriksson Stefan & Jonas Lagerström "Väljer företag bort arbetslösa jobbsökande?

2004:4 Forslund Anders, Daniela Fröberg & Linus Lindqvist "The Swedish activity guarantee"

2004:5 Franzén Elsie C & Lennart Johansson "Föreställningar om praktik som åtgärd för invandrares integration och socialisation i arbetslivet"

2004:6 Lindqvist Linus "Deltagare och arbetsgivare i friårsförsöket"

2004:7 Larsson Laura "Samspel mellan arbetslöshets- och sjukförsäkringen"

2004:8 Ericson Thomas "Personalutbildning: en teoretisk och empirisk översikt"

2004:9 Calmfors Lars & Katarina Richardson "Marknadskrafterna och lönebildningen i landsting och regioner"

Working Papers

2004:1 Frölich Markus, Michael Lechner & Heidi Steiger "Statistically assisted programme selection – International experiences and potential benefits for Switzerland"

2004:2 Eriksson Stefan & Jonas Lagerström "Competition between employed and unemployed job applicants: Swedish evidence"

2004:3 Forslund Anders & Thomas Lindh "Decentralisation of bargaining and manufacturing employment: Sweden 1970–96"

2004:4 Kolm Ann-Sofie & Birthe Larsen "Does tax evasion affect unemployment and educational choice?

2004:5 Schröder Lena "The role of youth programmes in the transition from school to work"

2004:6 Nilsson Anna "Income inequality and crime: The case of Sweden"

2004:7 Larsson Laura & Oskar Nordström Skans "Early indication of program performance: The case of a Swedish temporary employment program"

2004:8 Larsson Laura “Harmonizing unemployment and sickness insurance: Why (not)?”

2004:9 Cantoni Eva & Xavier de Luna “Non-parametric adjustment for covariates when estimating a treatment effect”

2004:10 Johansson Per & Mårten Palme “Moral hazard and sickness insurance: Empirical evidence from a sickness insurance reform in Sweden”

Dissertation Series

2002:1 Larsson Laura “Evaluating social programs: active labor market policies and social insurance”

2002:2 Nordström Skans Oskar “Labour market effects of working time reductions and demographic changes”

2002:3 Sianesi Barbara “Essays on the evaluation of social programmes and educational qualifications”

2002:4 Eriksson Stefan “The persistence of unemployment: Does competition between employed and unemployed job applicants matter?”

2003:1 Andersson Fredrik “Causes and labor market consequences of producer heterogeneity”

2003:2 Ekström Erika “Essays on inequality and education”