Monitoring sickness insurance claimants: evidence from a social experiment

Patrik Hesselius
Per Johansson
Laura Larsson

WORKING PAPER 2005:15

Monitoring sickness insurance claimants: evidence from a social experiment[*]

by

Patrik Hesselius[♣], Per Johansson[♦], and Laura Larsson[♥]

June 19, 2005

Abstract

The paper exploits a unique social experiment carried out in 1988 in Sweden to identify the effect of monitoring on sickness absence. The treatment consists of postponing the first formal point of monitoring during a sickness absence spell, a requirement for a doctor's certificate, from day eight to day fifteen. The experiment was conducted in two geographical areas, and the treatment group was randomized by birth date. The results show strong effects on sickness absence duration from extending the waiting period in both areas. On average, the durations increased by 6.6 percent. No effect on incidence of sickness absence is found. A heterogeneity analysis reveals that monitoring affects men more than women.

Keywords: Absenteeism, Sickness insurance, Monitoring, Social experiment.
JEL-codes: J22, J28, H55, I18.

[♣] IFAU and Uppsala University; e-mail: patrik.hesselius@ifau.uu.se
[♦] IFAU and Uppsala University; e-mail: per.johansson@ifau.uu.se
[♥] IFAU; e-mail: laura.larsson@ifau.uu.se

Table of contents

1 Introduction

A cornerstone of contract theory is that incentives matter. The last decades have witnessed a great development of the theory, and it has been introduced to new fields of economics, such as labor economics and insurance markets. However, surveys show that purely theoretical studies still exceed in number empirical attempts to validate the theory (Chiappori & Salanié, 2000, and Prendergast, 1999).

One area where quite much evidence already is available is public insurance programs. In the US, the literature on labor supply effects of worker's compensation, disability insurance, and most importantly of unemployment insurance is extensive. Krueger & Meyer (2002) provide a recent survey of these studies concluding that benefit generosity plays a significant role for labor supply of the insured. European studies have mainly focused on unemployment insurance and public sickness insurance programs, common in most European countries. The results support the conclusion by Kruger & Meyer.[1]

The size of the benefit is however not the only policy instrument of interest. Theoretically, duration and incidence of benefit take-up can be influenced by measures providing more direct incentives to escape the benefit state. Monitoring and sanctions, as well as cash bonuses paid to those who stop claiming for benefits quickly are examples.

Empirical validation of these effects lags behind the literature on benefit size, but there are some nice exemptions that are based on experimental data. Meyer (1995) and Ashenfelter et al (1999) discuss results from US unemployment insurance (UI) experiments. The experiments were designed to determine the effect on unemployment of cash bonuses or job search programs that often consist of both intensified job search assistance and more extensive job search verification. The authors conclude that very little seems to be worth the cost. Cash bonus experiments that consist of payments to UI recipients who found a job quickly (and stayed with that job long enough) are found to be cost-ineffective. The relative small effects on unemployment duration are easily

---

[1] See Atkinson & Micklewright (1991) and Holmlund (1998) for surveys of the UI literature, and Ercolani et al (2002) for a survey on sickness compensation. Swedish studies by Carling et al (2001) on UI benefits and unemployment, and Johansson & Palme (1996, 2002, and 2005); Henreksson & Persson (2004); Larsson (2005) and Hesselius (2004) on sickness benefits all suggest that the benefit size plays a significant role.

overweighed by the increase in the claimant population that a permanent adoption of reemployment bonuses is likely to cause. An enforcement of sanctions seems to make no difference to total UI benefit payments or the duration of benefit claims either. The only policy tool that according to these experiments is cost-effective is job search assistance: it leads to small but significant benefits to both workers and society. European evidence from the UI system however seems to suggest that intensified monitoring and sanctions do affect duration of unemployment.[2]

Our study addresses the role of incentives through monitoring by using experimental data on public sickness insurance in Sweden. Our contribution to the literature is thus twofold. First, randomized social experiments are extremely rare in Europe which makes our data unique. The design was very clean implying that our results provide credible estimates of the 'true' effect.

Second, our results shed some new light on incentives in the 'monitoring contract' between the insurance provider and the insured. Quite contrary to the conclusion from the UI experiments in the US, we find that eligibility checks do play an important role in explaining the duration of compensation claims. Our results concern claims from the public sickness insurance, but we argue that the conclusion holds even in a broader context.

Section 2 outlines the experiment and the Swedish sickness insurance. The experiment was conducted in the late 1980s in two economically quite different areas in Sweden. Half of the insured in both areas were randomized into a treatment group according to their date of birth. Treatment consisted of looser verification of eligibility: The deadline for providing the sickness insurance authorities a doctor's certificate that proofs reduced working capacity was postponed by a week.

In Section 3 we present the results. The general randomization method based on birth date allows us to reconstruct the treatment and control samples using administrative data from the National Social Insurance Board. The results turn out pronounced: relaxing the monitoring significantly increases absence due to sickness. Among the treated, the length of sick spells increased from an average of 11.86 days to 12.64 days. We use survival analysis to show that the probability of returning back to work is largest just one day before the deadline for eligibility verification.

---

[2] See Dolton & O'Neill (1996) and van den Berg & van der Klaauw (2001) for examples of studies using experimental data.

Section 4 introduces some measures of the economic importance of the results. The experiment increased sickness benefits expenses but reduced the number of visits to a doctor. The cost however was six times higher than the benefit. Furthermore, our estimated effect of monitoring on sickness absence corresponds to the effect of a six percent increase in the benefit size. Our results also suggest that sickness absence could be reduced by nearly ten percent if the deadline for the doctor's certificate was brought forward to the third day of absence. That is the rule in many other European countries.

These three measures lead us to our conclusion in Section 5 that, quite contrary to the results from UI experiments in the US, monitoring does significantly shorten the compensated period.

2   The Swedish sickness insurance

Sweden has 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. Basically all employed workers are automatically covered by the insurance.[3] Benefits are related to the lost income during the sick spell.

Sickness benefits are and have been rather generous: in 1988, a vast majority of workers received 90 percent of their lost income from the public insurance. A benefit cap excluded workers at the very top of the income distribution from receiving a full 90 percent. However, besides the public 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 replace about 10 percent of forgone earnings, but there is considerable variation. The total compensation could thus be full 100 percent.

The public insurance has no limit for how often or how long benefits are paid. Many sick spells continue for more than a year but there are examples of

---

[3] The insurance covers even some non-employed persons such as student and job-seekers who are registered with the Public Employment Service. Students usually receive the minimum amount of benefits whereas the sick benefits of the unemployed in general are determined by their income prior to unemployment.

even much longer durations. These long spells are more likely to lead to disability pensions than to a return back to work.

Since compensation levels are so high, one would expect monitoring of the benefit claimants to be strict in order to reduce moral hazard. This is not the Swedish case however. On the contrary, public insurance offices have even in that sense been quite generous paying out sickness benefits for a week before checking the claimants' eligibility. A sick spell starts when the worker calls the public social insurance office (and his or her employer) to report sick. Within a week, at latest on the 8th day of sickness, the claimant should verify eligibility by showing a doctor's certificate that proofs reduced working capacity due to sickness. The public insurance office then judges the certificate and decides about further sick-leave. It is very rare that the certificate is *not* approved.

Of course, some exceptive rules make it possible for the public insurance offices to monitor more (or less) strictly. In case they suspect abuse, they can visit the claimant at home. Claimants who have been on sickness benefits too many times during the past may be asked to show a doctors certificate from day one. Moreover, a new sick spell starting within five working days from the first is counted as a continuation of the first making it impossible to report sick every Monday (and returning 'back to work' for the weekends) without ever visiting a doctor. Persons with chronic illnesses, on the other hand, do not necessarily have to verify their eligibility each time the illness forces them to stay at home from work.

Since 1988, some features of the system have changed. The compensation level has been reduced to 80 percent. The benefit cap has not followed the inflation rate implying that today, approximately a quarter of the workforce receives less than 80 percent of their forgone earnings from the public insurance. Among actual benefit claimants the share is however lower, slightly more than 10 percent.[4] In 1993 the generosity was further reduced by an uncompensated qualifying day in the beginning of each sick spell. Since 1991, some of the financial (and administrative) burden has been laid on employers in that they pay the sickness benefits during the first weeks of sickness.[5]

---

[4] The figures are for 2003 and come from Larsson (2003).

[5] Their responsibility for the sickness benefits, or 'sick pay', has varied from 2 to 4 weeks. In June 2005, they pay the full benefits for the two weeks of sickness (except for the qualifying day) and after the two weeks a small share of the total cost for the rest of the sick spell.

3   The experiment

The experiment we use to identify the effect of monitoring was carried out in the second half of 1988 in *Gothenburg*, the second largest city in Sweden, and in *Jämtland*, a small county in the sparsely populated Northern part of Sweden. It was initiated by the local social insurance offices.[6]

The purpose of the experiment was to see whether and how sickness absence is altered when monitoring of the insurance claimants is reduced. A randomly assigned treatment group was allowed to receive sickness benefits for two weeks without showing a doctor's certificate, instead of one week as usual. The randomization was performed by using the date of birth. All insured born on an even date were asked to show a doctor's certificate after two weeks, whereas insured born on an uneven date had to show one already after one week.

The insurance authorities had several arguments for running the experiment. In short, all of them were based on an idea that it would imply saving and less sickness absence. First, unnecessary visits to a doctor would decrease implying less cost for individuals, the medical care system and thereby for the state budget. The implementing authorities also believed that doctors, in a routine way, prescribe longer absence from work than necessary. With a two week's time limit, many individuals would have time to get back to work before receiving any such prescription. Finally, and perhaps somewhat contradictory to the above arguments, some sick spells were indeed expected to get longer, but for a good reason as sick individuals no longer were pushed back to work. This in turn would decrease the risk of recurrence of those individuals.

The background to the experiment, and thus the starting point, differed some between the two areas. The idea to test a two weeks' rule for the doctor's certificate was born in Jämtland and had been in use there for *all* insured since January 1987. In Gothenburg, the usual rule of one week's unmonitored sickness absence was in use until the experiment. Thus, in Gothenburg, the experiment implied looser rules for half of the insured, whereas in Jämtland, it implied harder rules for half of the insured. In spite of this difference, and to make

---

[6] Until recently, the public insurance was administered by 21 independent local social insurance offices that were quite free to design exceptions from the general rules (as long as they were towards more generosity). Today, the administration is centralized.

the presentation clear, we label the group with a two weeks waiting period as the treatment group and the group with a one week waiting period as the control group.

The experiment was a non-blind experiment in that all parts were informed about it in advance or at latest during the experiment. In fact, it was preceded by quite massive local information campaigns. Besides the personnel at the local social insurance offices, all employers and medical centers were informed in advance about the set-up of the experiment. Written material such as brochures and posters were spread out as well as verbal information through meetings and consultants. Also the mass media were an important channel to inform the insured. Furthermore, short information about the experiment was written on the form that every insured reporting sick must fill in to receive sickness benefits.

We use data from the National Social Insurance Board to reconstruct the treatment and control samples and to evaluate the effect of the experiment. The data include detailed information about individual sick spells, the birth date, and some other characteristics. *Table 1* shows the distribution of the insured persons in treatment and control groups in the two experiment areas. The control group is larger than the treatment group as there are more uneven than even dates.[7]

Table 2 shows descriptive statistics for both Jämtland and Gothenburg subdivided into the control group and treatment group, respectively. There are no significant differences between the treatment and control groups with respect to any of the characteristics: Gender and age distributions as well as average age and average sickness absence prior to the experiment are all but equal between the treatment and the control group. Thus, the randomization seems to be valid.

---

[7] To be exact, the samples of treated include some individuals who were excluded from the experiment, such as Government employees and some other groups that were comprised by special rules. According to the internal report of the experiment, there were approximately 30,000 (Gothenburg) plus 10,000 (Jämtland) such individuals among the insured. Half of them are randomized into the treatment groups but we cannot identify them in the data. However, this should only bias the results downward and thus reinforce our conclusion.

Table 1 Number of insured 1 July – 31 December, 1988.

	Control group	Treatment group	Total
Jämtland	33,135	31,861	64,996
Gothenburg	121,276	116,115	237,391

Table 2 Descriptive statistics.

	Gothenburg		Jämtland	
	Control group	**Treatment group**	**Control group**	**Treatment group**
Fraction women	48.1 %	48.2 %	47.6 %	47.5 %
Mean age	38.15	38.10	38.76	38.85
Income:				
Mean	1190.35	1195.38	1040.22	1042.61
25th percentile	903	910	830	830
50th percentile	1164	1170	1050	1050
75th percentile	1430	1437	1250	1250
Benefit cap:				
No above cap	8545	8389	785	783
Percent of total	7.05	7.22	2.37	2.46
Sickness absence 1/1/88 – 6/30/88:				
Incidence	1.23	1.22	1.07	1.05
Prevalence	15.62	15.52	13.95	13.72

A somewhat striking result, however, might be that before the experiment sick spells were longer and more common in Gothenburg than in Jämtland. Recall that in Jämtland, all insurance claimants were comprised by a two weeks' rule already since January 1987. The geographical variation in Swedish sickness absence has changed since 1988. Today, sickness absence is in general larger at the country side than in larger cities, and in 2004, sickness absence was significantly higher in Jämtland than in Gothenburg.[8]

To further explore on the validity of the randomization device we take a closer look at the sickness absence of the treatment and the control group before the experiment. *Figure 1* and *2* show the estimated survival functions[9] in the pre-experimental period, i.e. the fraction of sick that remains after a certain number of days in a sickness spell. As expected, there is hardly any difference between the treated and the controls.

However, all three figures indicate a monitoring effect. In Gothenburg *before the experiment*, we observe a drop in the survival rate around the 8th day of a sick spell. The pre-experimental survival functions in Jämtland look as expected somewhat different: The drop in survival function is displayed at the 15th day instead of the 8th day.

---

[8] According to figures from the National Social Insurance Board, the average number of sick days was 19.4 in Gothenburg and 27.8 in the county of Jämtland. Note that these figures do not include sick days during the employers' period that during 2004 was three weeks.
[9] We use the Kaplan-Meier estimator throughout the survival analysis (Kaplan and Meier, 1958).
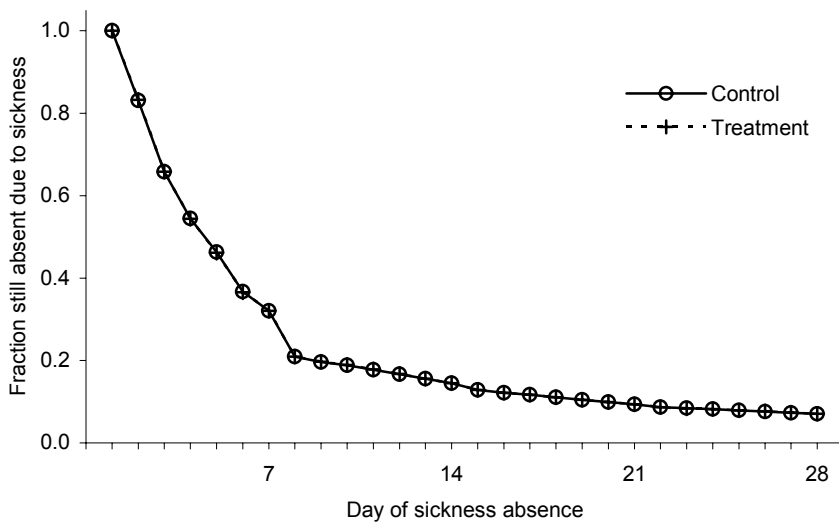
Figure 1 Fraction still absent due to sickness in Gothenburg during the half year before the experiment period (1/1/88 – 6/30/88).
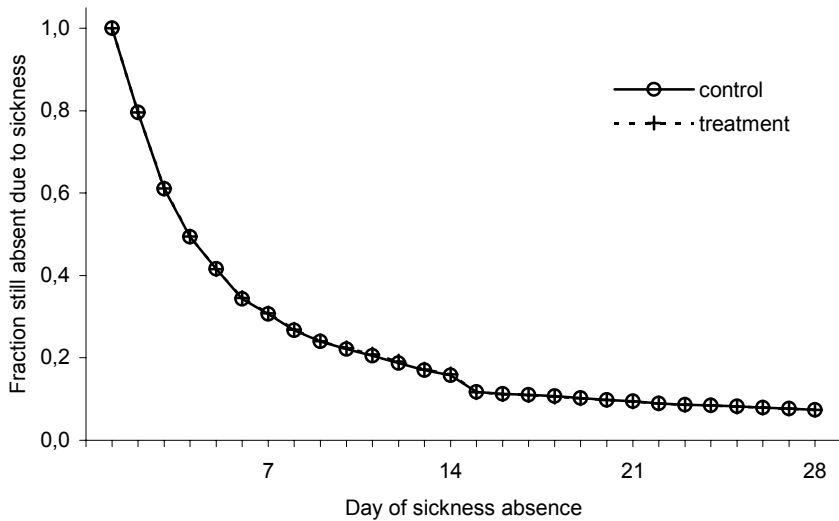
Figure 2 Fraction still absent due to sickness in Jämtland during the half year before the experiment period (1/1/88 – 6/30/88).

4 The results

We start by looking at the how monitoring affects absence duration. Given the shapes of the survival functions in Gothenburg and Jämtland prior to the experiment we have a strong prior: the looser the checks on eligibility the longer the absence spells. Besides duration, we are interested in the incidence of sick spells. We address heterogeneity by estimating the effects for various types of individuals.

4.1 Duration

Figure 3 and *4* illustrate how monitoring affects the length of the sick spells by survival functions. The dotted line shows the fraction of ongoing sick spells among the treated and the continuous line shows the corresponding fraction among the controls. The effect is distinct: the survival rate is significantly higher for the treatment group during the entire second week.[10]

A closely related way to analyze the effect is to estimate hazards rates. It gives us a more detailed picture of when the probability of ending a sick spell is largest. *Figure 5* and *6* show the results for Gothenburg and Jämtland.[11] It is evident that the control group – monitored at day 8 – tends to exit after 7 days of absence, whereas the treatment group waits another week before returning back to work. The hazard rates also show a weakly cyclical pattern: there is a peak in the hazard rate every 7th day. We believe that it is a common practice among the medical doctors to put the patients on a sick-list for full weeks.

Furthermore, the figures indicate that the monitoring effect starts already some days before the actual date of monitoring. The hazard rate of the control group is above the rate of the treatment group already at day 5 and 6. The opposite holds for day 12 and 13: the probability of returning back to work is higher among the treated. We call it a "pre-monitoring effect" and believe that it is due to weekends. Most employees work 5 days Monday-Friday and have the weekend off. Thus, in the register, a sickness spell that starts on a Monday often ends on a Friday, even though the total absent from work is 7 days.

---

[10] The survival estimates and adherent 95% confidence intervals are presented in appendix A.
[11] Estimated using the product-limit method.
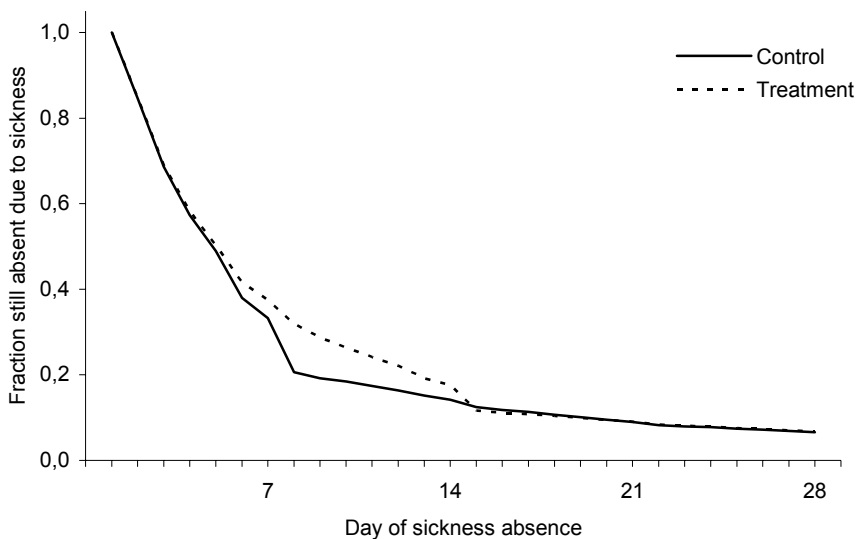
Figure 3 Fraction still absent due to sickness in Gothenburg during the experiment period.
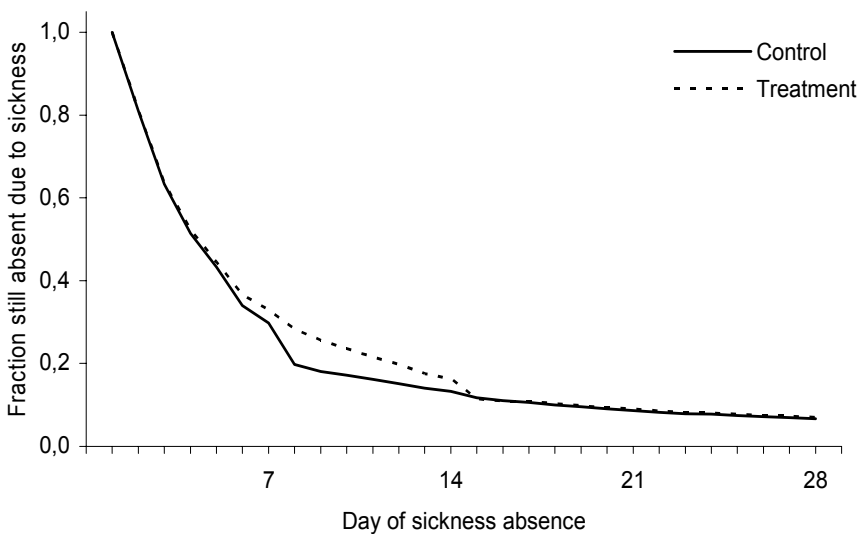
Figure 4 Fraction still absent due to sickness in the county of Jämtland during the experiment period.
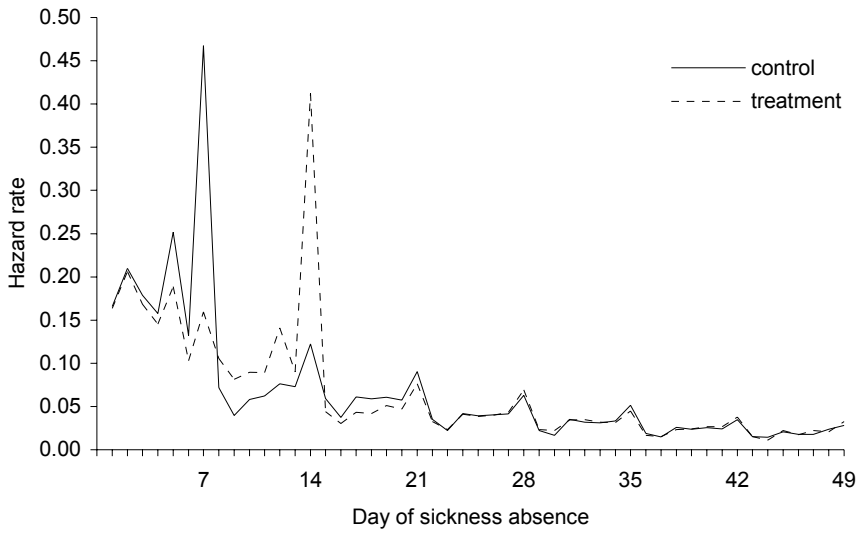
Figure 5 Hazard rate during the period of the experiment, Gothenburg.
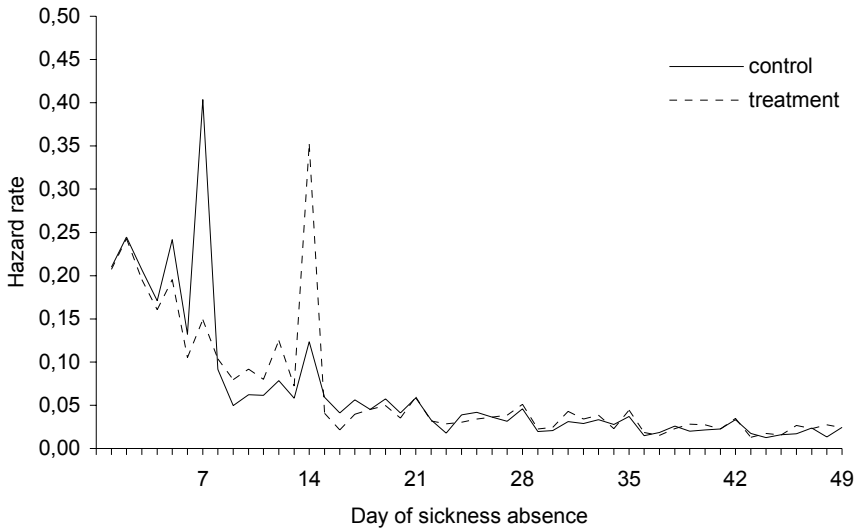
Figure 6 Hazard rate during the period of the experiment, county of Jämtland.

IFAU – Monitoring sickness insurance claimants

Finally, there seems to be a "post-monitoring effect", as well. The hazard rates of those just monitored are below the hazard rate of the comparison population for some time after the date of monitoring. This is most likely due to sample selection, or a *harvesting effect* according to epidemiology literature (see e g Schwartz, 2000).[12] Let us be more specific. It is reasonable to assume that the date of monitoring "pushes" some people back to work that without monitoring would have stayed on sick-leave for some more days. But when forced to show a doctor's certificate in order to continue on sick-leave, they perceive the cost of staying on sick-leave higher than the cost of returning to work sick. If so, the average health of the population just monitored is worse than the average health of the comparison population right after the date of monitoring. Consequently, their hazard rate is lower those days than it would have been without monitoring.

4.2   Incidence

One of the arguments for running the experiment was to reduce the risk of recurrence as the insured were not pushed back to work 'too early' while still ill. Such an effect would imply fewer sick spells during and after the experiment period. *Figure 7* and *8* display the incidence in Gothenburg and in Jämtland, before, during and after the experiment. In short, no significant effect at all is found on sickness incidence.

4.3   Are the effects the same for everybody?

We have shown that the effect of monitoring as the same in Gothenburg and in Jämtland. This is a useful result for several purposes. First, it convinces us that our experiment is not impaired by substitution bias.[13] Recall that the starting points were different in these two areas: In Jämtland, the two weeks' rule had been in use there for *all* insured since January 1987 whereas in Gothenburg, the usual rule of one week's unmonitored sickness absence was in use until the experiment. We would thus expect the risk of the control group being treated to

---

[12] An alternative explanation would be that doctors prescribe longer sick-leaves than necessary. In that case the post-monitoring effect could be taken as a 'causal' effect of the monitoring device itself.
[13] See Moffitt (2004) for a nice discussion of shortcomings with external validity in randomized trials.
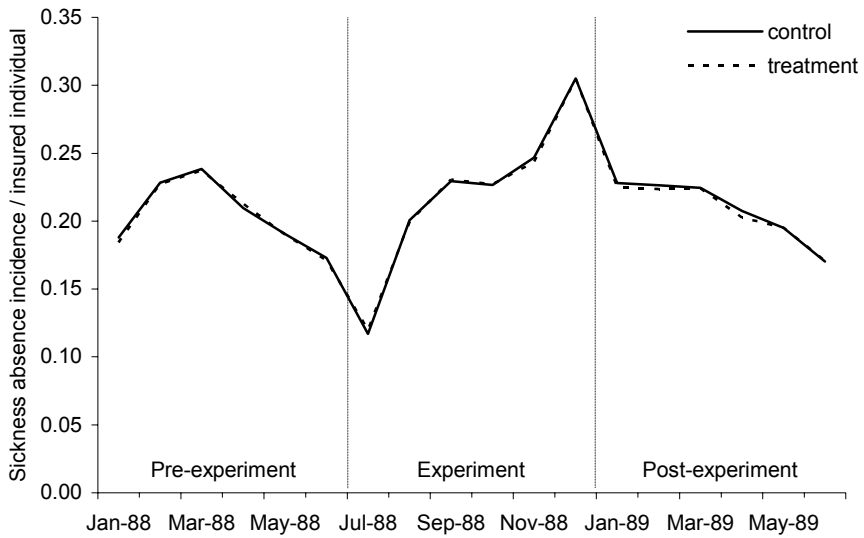
Figure 7 Sickness absence incidence per insured individual in Gothenburg, before, during and after the experiment.
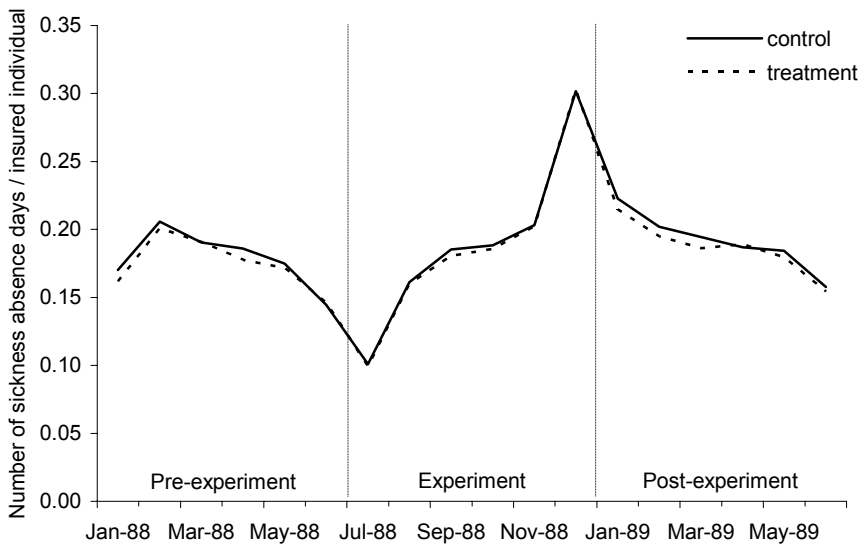
Figure 8 Sickness absence incidence per insured individual in Jämtland, before, during and after the experiment.

IFAU – Monitoring sickness insurance claimants

be higher in Jämtland than in Gothenburg, simply because it must have been more difficult to implement harder than looser monitoring. This in turn would imply that the estimated effect in Jämtland would be more biased than the effect in Gothenburg. The effects are however the same, suggesting that substitution bias is not an issue here.

Second, we are not worried about the so called Hawthorne effect either. It means that individuals respond to the experiment because they know that their behavior is being measured. Again, we believe that the asymmetry in the starting point would imply different Hawthorne effects in Gothenburg and in Jämtland. We can also compare the survival functions during the experiment (*Figure 3* and *4*) with the pre-experimental survival functions (*Figure 1* and *2*) to look for such effects. On the contrary, the treated seem to behave like all insured in Jämtland did before the experiment, and the controls behave like all insured in Gothenburg did before the experiment.

Finally, and perhaps most importantly, the fact that the effects were the same makes it possible to generalize our results. The labor market in Jämtland is very different from the labor market in Gothenburg. Jämtland is a rural area in the North of Sweden and Gothenburg is the second largest city in Sweden. The education level is higher in Gothenburg than in Jämtland and, thus, the white/blue collar worker ratio is much higher in Gothenburg than in Jämtland. This difference in labor market is also mirrored in the difference in income distributions displayed in *Table 2*. Since the effect estimate in the two regions are basically the same we believe that this average effect can be generalized to Sweden in large.

However, we are still interested in other potential dimensions of heterogeneity. Anecdotic evidence for example suggests that the moral among the young is worse than the moral among the elderly. Besides differences according to (1) age, we also test for if the effect of monitoring differs with respects to (2) gender and (3) income.

We estimate proportional hazard models (cf Lancaster, 1990) for each category and report the relative risk of the treatment group as compared to the control group at day seven and fourteen. Note that estimating proportional hazard instead of empirical hazards does not imply any more restrictive assumptions in our case since the experiment implied random assignment. Thus the standard critique of erroneously non-proportionality is not relevant, however we restrict the treatment effects to occur at day 7 and 14 in a spell, thus we here neglect the pre- and post-treatment effects.

Table 3 shows the results. A relative risk ratio below (above) one means that the treatment group has a smaller (larger) probability of exit than the control group has. The first row reports the hazard estimates for all observations, already shown in *Figure 5* and *6*. The following rows show that the only statistically significant difference in the monitoring effect is between men and women: The relative risk of the treatment group is lower among men than among women at day 7, whereas the opposite is the case at day 14. Men seem to react much stronger on monitoring than women do. But neither age nor income level seems to play any role; the estimated relative risks are basically identical for all age and income groups. In other words monitoring seems to have the same effect for all these groups.

How should we explain the different behavior between the sexes? Do men simply have (more) lax morals? Maybe, but we cannot exclude that at least part of the difference is due to selection. Let us be more specific. In general, the group of individuals who are absent due to sickness has different characteristics than the overall labor force. Women are more absent due to sickness than men, implying that the female 'sick' population *on average* differs less from the labor force than the male 'sick population' *on average* does. This, in turn, may imply that the 'sick' women to behave differently than the 'sick' men. One explanation for why women are more absent due to sickness is that they due to the fertility have more and better contacts with the medical service than men. This makes the indirect cost of visiting a doctor lower for women than for men.

Another thing to note is that, in all subgroups, the effect of monitoring is somewhat larger at day 7 than at day 14. We believe that this is due to the *harvesting effect* discussed earlier: At day 14, the sample of controls consists of persons with worse health than the treated. Thus the assignment to treatment is not really random any more at that point of time, and the monitoring effect is ambiguously biased. However, *Figure 5* and *6* indicate that this bias is very small and hardly statistically significant. In fact, the monitoring effects at day 7 and 14 are statistically significant different (at the 5 percent level) only when estimated for all observations in Gothenburg. When estimated for any of the sub-groups, or for all observations in Jämtland, the difference is not statistically significant.

Table 3 Estimated relative risks between the treatment and the control groups at sickness spell day 7 and 14. Standard errors in parentheses.

	Gothenburg				Jämtland county			
	Relative risk[a]		Parameter		Relative risk[a]		Parameter	
	Day 7	Day 14	Day 7	Day 14	Day 7	Day 14	Day 7	Day 14
All observations	0.338	3.393	-1.086 (0.014)	1.222 (0.023)	0.370	3.124	-0.995 (0.031)	1.139 (0.052)
Gender:								
Men	0.285	4.067	-1.257 (0.019)	1.403 (0.034)	0.327	3.808	-1.117 (0.043)	1.337 (0.075)
Women	0.398	2.787	-0.922 (0.019)	1.025 (0.032)	0.414	2.488	-0.881 (0.043)	0.911 (0.072)
Age:								
16-25	0.304	3.526	-1.191 (0.028)	1.260 (0.049)	0.390	3.563	-0.941 (0.064)	1.271 (0.119)
26-35	0.332	3.709	-1.103 (0.025)	1.311 (0.043)	0.339	2.911	-1.082 (0.060)	1.068 (0.101)
36-45	0.365	3.229	-1.009 (0.029)	1.172 (0.048)	0.385	3.067	-0.954 (0.062)	1.121 (0.102)
46-55	0.350	2.984	-1.050 (0.035)	1.093 (0.058)	0.391	2.887	-0.938 (0.075)	1.060 (0.120)
56-65	0.357	3.025	-1.031 (0.047)	1.107 (0.076)	0.332	3.537	-1.104 (0.098)	1.263 (0.157)
Income:								
1st Quartile	0.348	3.404	-1.055 (0.031)	1.225 (0.052)	0.332	2.976	-1.103 (0.060)	1.091 (0.091)
2nd Quartile	0.352	3.395	-1.045 (0.025)	1.222 (0.042)	0.348	3.412	-1.055 (0.050)	1.227 (0.086)
3rd Quartile	0.327	3.327	-1.116 (0.025)	1.202 (0.043)	0.432	2.933	-0.839 (0.062)	1.076 (0.110)
4th Quartile, below cap	0.320	3.371	-1.138 (0.032)	1.215 (0.054)	0.419	3.546	-0.870 (0.098)	1.266 (0.177)
Over cap	0.337	4.314	-1.089 (0.073)	1.462 (0.152)	0.442	0.893	-0.816 (0.282)	-0.114 (0.474)

[a] Relative risk = exp(parameter).

5   Economic significance of the results

We have shown that monitoring plays a statistically significant role in explaining the duration of compensation claims. To say something about the *economic significance* of the effect we introduce three measures. First, we compare the cost and the benefit of the experiment. The second measure relates the estimated effect of monitoring to the effect of an alternative reform, namely an increase of the compensation level. Third, we use our results to calculate how much sickness absence in Sweden could be reduced if the deadline for the doctor's certificate was brought forward to the 3rd day of absence. That is the rule in many other European countries.

5.1   Cost and benefits from relaxing monitoring

Our simple cost benefit analysis compares the cost of increased sickness duration to the benefit of fewer visits to a doctor. We obtain an estimate of the cost by taking the average increase in a sick spell due to the reform *times* the number of sick spells *times* the average cost of one sick day. For simplicity, let us base our analysis on figures for Gothenburg. In 1988, the average cost of one day's absence (in terms of sickness benefits) was SEK 296.[14] The number of sick spells during the experiment period was 122,123. The average increase in a sick spell due to the reform can be calculated using the estimated survival functions. The expected duration is 11.84 under a 7 days regime and 12.64 under a 14 days regime.[15] Thus, the estimated increase is 0.8 days, corresponding to 6.6 percent. Taken these figures together, the estimated cost of extending the waiting period for all insured (in Gothenburg) would have been equal to approximately SEK 29 million.

The benefit is determined by the average cost of a doctor's visit *times* the decrease in the number of visits. According to estimates from the Federation of Swedish County Councils, the average cost of a doctor's visit in 1988 was SEK 445.[16] The reduction in the number of visits is again obtained from our estimated survival functions. In the control group, 20.6 percent of all spells

---

[14] SEK 100 corresponds to USD 12.8 in June 2005.

[15] These estimates neglect that some spells are censored at day 365.

[16] The cost of a visit to a GP in 1988 is calculated using the cost in 1991 deflated by the average visit cost increase within internal medicine and ear, nose, throat care between 1988 and 1991 (LF, 1988 and LF, 1991)

were longer than one week and thus involved a visit to a doctor. In the treatment group, 11.6 percent of the spells were longer than two weeks. Thus extending the waiting period to fourteen days would potentially reduce the number of visits to a doctor by 9 percentage units[17]. The total number of sick spells being 122,123, and assuming that these spells are equally distributed among the control and treatment groups, an extension of the waiting period by a week for all insured would reduces the number of visit by 10,991. Thus, the benefit equals to SEK 4,9 million, which is considerably lower than the estimated cost of SEK 29 million.

Another way of presenting the increased insurance cost is to divide the total cost of SEK 29 million by the decreased number of forced physician visits due to the certification requirement (0.09*122,123). This yields that the reform is cost neutral if the average cost of a physician visit is SEK 2,631. In comparison to the actual cost of SEK 445, one can easily see that the cost side of the reform is greater than the benefit side.

5.2 Monitoring and compensation level

From the insurance theory we know that moral hazard can be reduced either by increasing the monitoring or by reducing the compensation. An interesting policy question is then how much the compensation level must be increased to obtain the same increase in work absence as is obtained by relaxing the monitoring by one week?

Using the estimated hazard rates for the two groups, we can simulate the proportional increase of the hazard rate necessary to obtain equivalent expected duration as monitoring at day 15th. The calculations are presented in the Appendix. Using this way of calculating the expected duration, monitoring at the 15th day yields an expected duration of 16.18 days. Setting the proportionally altered expected duration of monitoring at day 8 equal to 16.18 yields a proportional decrease of the hazard rate by 1.6 percent. Based on and elasticity estimate from Johansson & Palme (2004)[18], an increase in the compensation level by 6 percent would lead to an equally large effect as the effects from monitor-

---

[17] It is likely that a not so small fraction of the sickness absent has seen a doctor to receive medical treatment before the certificate requirement has to be enforced. Thus, the actual number of reduced doctors visits is probably smaller than 9 percentage units.
[18] The elasticity with respect to the compensation level is estimated to 0.25.

ing. In terms of the compensation level in June 2005 this would imply increasing the compensation level from 0.80 to 0.85.

5.3 The effects of changing waiting time to three days

Many other European with less work absence demand a doctor certificate already at day three in a work absence spell. It is, thus, of interest to see how much of the relatively high work absence in Sweden can be attributed to the relatively loose monitoring in Sweden. To make such an extrapolation we need some assumption about how the monitoring effect depends on duration. *Table 3* indicates that the effect is larger at day 15 than at day 8. However, we argued that the effect at day 15 is upward biased due to sample selection.

To test for whether the 'true' monitoring effect is constant over the duration, we model both the pre-monitoring effects (1-3 days before the monitoring) and the post-monitoring effects (1-7 days after the monitoring) together with the monitoring effects within a proportional hazard framework. By explicitly introducing the pre- and post-effects in the model we 'wash off' the selection effect. The estimated monitoring effects turn out to be very close to each other at day 8 and 15, as shown in *Table 4*.

The predicted baseline hazards for the treatment group and the control group are displayed in *Figure 9*. The predicted baselines are very similar for the treatment and control group. This is compelling evidence that the monitoring effect is proportionally constant on the baseline hazard. In *Figure 10*, the observed hazard rates are plotted for the control group (monitoring at day 8) together with the simulated (predicted) hazard rate when monitoring at day 8. The modeling of the direct monitoring effect together with the pre- and post-effects seem to yield a nearly precise prediction of the true hazard rate as the curves almost coincide. This allows us to calculate the effect of instead *reducing* the monitoring date from day 7, again using the estimated survival functions. The result from this exercise is presented in *Table 5*. Restricting the monitoring rule so that a doctor's certificate must be shown on the 3rd day of a sickness spell would decrease the work absence by nearly 10 percent.

Table 4 Estimates of the monitoring effect for the Gothenburg experiment using Cox proportional hazard estimation method with time-varying covariates.

	Parameter	(Standard error)	Relative risk
3 days prior	0.070	(0.013)	1.073
2 days prior	0.254	(0.012)	1.289
1 day prior	0.199	(0.017)	1.220
Monitoring effect at day 7	1.099	(0.019)	3.000
Monitoring effect at day 14	1.164	(0.053)	3.202
1 day after	-0.350	(0.023)	0.705
2 days after	-0.582	(0.030)	0.559
3 days after	-0.392	(0.027)	0.676
4 days after	-0.296	(0.029)	0.744
5 days after	-0.282	(0.026)	0.755
6 days after	-0.049	(0.031)	0.952
7 days after	-0.152	(0.044)	0.859

Table 5 The effect from changing the waiting period on the durations based on the Gothenburg experiment.

	Expected spell length (max 365 days)	Difference from monitoring at t = 8, in days	Difference from monitoring at t = 8, in percent
Monitoring at t = 3	10.75	-1.12	-9.4 %
Monitoring at t = 8	11.86	-	-
Monitoring at t = 15	12.64	0.78	6.6 %
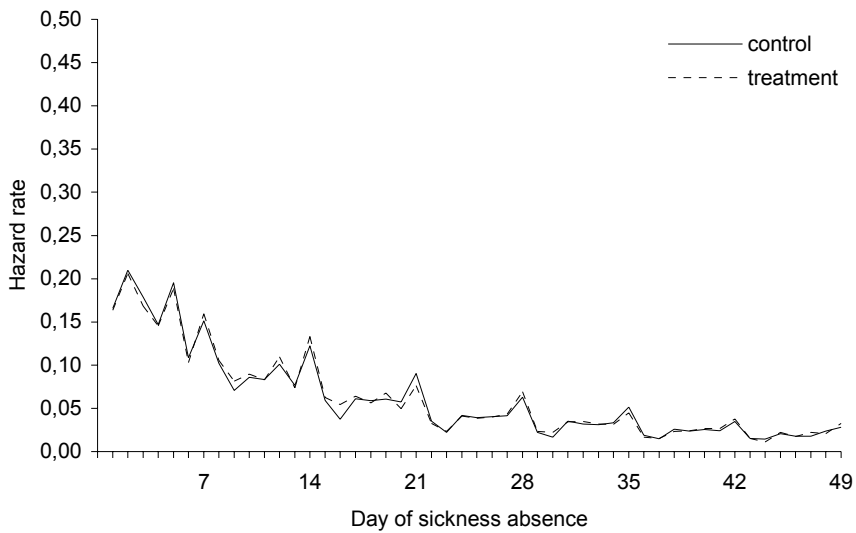
Figure 9 The calculated hazard rate in the absence of monitoring based on Gothenburg data.
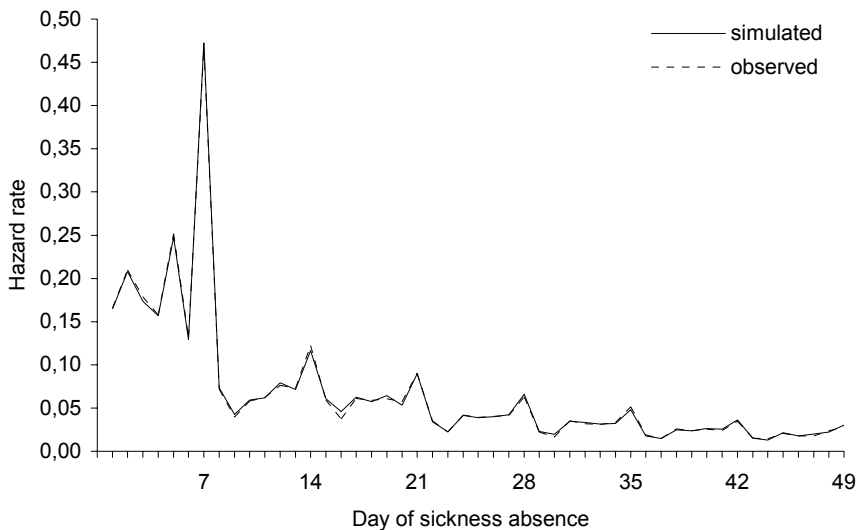
Figure 10 The observed and simulated hazard rate when monitoring at day 8th, based on the control group using the Gothenburg data.

6   Conclusions

We have shown that the degree of monitoring plays an important role in reducing moral hazard in the Swedish sickness insurance system. This evidence is based on a extremely well-performed randomized controlled experiment. Besides being highly internal valid the experiment seems to have high external validity. Thus, it is possible to simulate policy changes both outside our sample and for various potential treatments.

The results from this experiment together with results from previous studies on excess in sickness insurance, suggest that postponing the point of monitoring by one week corresponds to an increase of the compensation level by 6 percent. From a policy perspective, this is an important trade off: the distributional (and thus equity) effects of increasing monitoring are considerably different from the distributional effects of an overall reduction of the compensation level.

An interesting aspect in this context is the heterogeneity between men and women. Monitoring seems to have a stronger effect on men than on women. On the other hand, women are absent due to sickness more often. Thus, *men as a group* are hit harder by increased monitoring, whereas *women as a group* are hit harder by an increase in excess. Of course, in order to judge the equity aspects one should know why sickness absence is higher among women. To our knowledge, this is still a question for future research.

References

Ashenfelter O, D Ashmore,, and O Deschêns (1999), Do Unemployment Insurance Recipients Actively Seek work? Randomized trials in Four U.S. States, NBER working paper 6892.

Atkinson A and J Micklewright (1991), Unemployment compensation and labour market transitions: A critical review, *Journal of economic literature*, 29, 1679-1727

Carling K, B Holmlund, and A Vejsiu (2001), Do benefit cuts boost job findings? Swedish evidence from the 1990s, *Economic Journal* 111, 766-790.

Chiappori P and B Salanié (2000), Testing contract theory: a survey of some recent work, invited lecture, World Congress of the Econometric Society in Seattle, 2000.

Dolton, P and D O'Neill D (1996), Unemployment duration and the restart effect: Some experimental evidence, *Economic Journal*, 106, 387-400.

Ercolani M, T Barmby, and J Treble (2002), Sickness Absence: An International Comparison*, Economic Journal* , 12, F315-F331

Henreksson M and M Persson (2004), The Effects on Sick Leave of Changes in the Sickness Insurance System, *Journal of Labor Economics,* Vol. 22(1), 87-113.

Hesselius (2004), Sickness Absence and Labor Market Outcome, *Economic Studies 82*, Department of Economics, Uppsala University.

Holmlund B (1998), Unemployment insurance in theory and practice, *Scandinavian Journal of Economics*, 100, 113-41.

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

– (2002), Assessing the Effects of a Compulsory Sickness Insurance on Worker Absenteeism, *Journal of Human Resources*, 37:2, 381-409.

– (2005), Moral hazard and sickness insurance, forthcoming in *Journal of Public Economics*.

Kaplan E and P Meier (1958), Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, 457-481.

Krueger A and B 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.

Larsson L (2005), Sick of being unemployed? Interactions between unemployment and sickness insurance, forthcoming in *Scandinavian Journal of Economics*.

LF (1988), Kostnader per intagen patient, vårddag, läkarbesök m m 1988 (The cost per patient, day of care, doctor's visit etc 1988)", Landstingsförbundet (Federation of Swedish County Councils).

LF (1991), "Kostnader per intagen patient, vårdag, läkarbesök m m 1991 (The cost per patient, day of care, doctor's visit etc 1991)", Landstingsförbundet (Federation of Swedish County Councils).

Meyer, B (1995), Lessons from the U.S. Unemployment Insurance Experiments, *Journal of Economic Literature,* 33:1, 91-131

Moffitt Robert A. (2004), The role of randomized field trials in social science research, *American Behavioral Scientist*, 47, 506-540.

Schwartz J (2000), Harvesting and long term exposure effects in the relation between air pollution and mortality, *American Journal of Epidemiology*, 151, 440–448.

van den Berg G and B van der Klaauw (2001), Counceling and monitoring of unemployed workers: Theory and Evidence from a social experiment, working paper 2001:12, IFAU, Uppsala.

Appendix A

Table A1. Survival rates

	Gothenburg					Jämtland				
	Control		Treatment			Control		Treatment		
Time	Fraction remaining	95% CI	Fraction remaining	95% CI	Difference	Fraction remaining	95% CI	Fraction remaining	95% CI	Difference
1	100.0	100.0 - 100.0	100.0	100.0 - 100.0	0.0	100.0	100.0 - 100.0	100.0	100.0 - 100.0	0.0
2	84.7	84.5 - 84.9	84.8	84.7 - 85.0	0.2	81.0	80.6 - 81.4	81.2	80.8 - 81.6	0.2
3	68.6	68.3 - 68.8	68.9	68.7 - 69.2	0.4	63.4	62.8 - 63.9	63.7	63.1 - 64.2	0.3
4	57.3	57.1 - 57.6	58.2	57.9 - 58.5	0.9	51.5	51.0 - 52.0	52.3	51.8 - 52.9	0.9
5	48.9	48.7 - 49.2	50.3	50.0 - 50.5	1.3	43.4	42.8 - 43.9	44.6	44.0 - 45.1	1.2
6	38.0	37.7 - 38.2	41.6	41.3 - 41.8	3.6	34.0	33.5 - 34.5	36.6	36.1 - 37.1	2.6
7	33.3	33.0 - 33.5	37.5	37.2 - 37.7	4.2	29.8	29.3 - 30.3	33.0	32.5 - 33.5	3.2
8	20.6	20.4 - 20.8	31.9	31.6 - 32.1	11.3	19.8	19.4 - 20.2	28.4	27.9 - 28.9	8.6
9	19.2	19.0 - 19.4	28.7	28.4 - 28.9	9.5	18.1	17.7 - 18.5	25.6	25.1 - 26.1	7.5
10	18.4	18.2 - 18.6	26.4	26.2 - 26.6	8.0	17.2	16.8 - 17.6	23.6	23.2 - 24.1	6.4
11	17.4	17.2 - 17.6	24.2	23.9 - 24.4	6.8	16.2	15.8 - 16.5	21.6	21.1 - 22.0	5.4
12	16.4	16.2 - 16.6	22.1	21.9 - 22.3	5.7	15.2	14.8 - 15.6	19.9	19.5 - 20.3	4.7
13	15.2	15.0 - 15.4	19.2	19.0 - 19.4	4.0	14.0	13.7 - 14.4	17.6	17.1 - 18.0	3.5
14	14.2	14.0 - 14.3	17.6	17.4 - 17.8	3.5	13.3	12.9 - 13.6	16.3	15.9 - 16.7	3.1
15	12.5	12.3 - 12.7	11.6	11.5 - 11.8	-0.9	11.7	11.4 - 12.0	11.4	11.1 - 11.8	-0.3
16	11.8	11.6 - 12.0	11.1	11.0 - 11.3	-0.7	11.0	10.7 - 11.4	11.0	10.6 - 11.3	-0.1
17	11.4	11.2 - 11.5	10.8	10.7 - 11.0	-0.5	10.6	10.3 - 10.9	10.8	10.4 - 11.1	0.2
18	10.7	10.5 - 10.9	10.4	10.2 - 10.5	-0.3	10.0	9.7 - 10.3	10.3	10.0 - 10.7	0.3
19	10.1	9.9 - 10.2	9.9	9.8 - 10.1	-0.2	9.6	9.3 - 9.9	9.9	9.6 - 10.2	0.3
20	9.5	9.4 - 9.7	9.4	9.3 - 9.6	-0.1	9.0	8.7 - 9.3	9.4	9.1 - 9.7	0.4
21	9.0	8.8 - 9.1	9.0	8.9 - 9.2	0.0	8.7	8.4 - 9.0	9.1	8.8 - 9.4	0.4
22	8.2	8.1 - 8.3	8.4	8.2 - 8.5	0.2	8.2	7.9 - 8.5	8.6	8.2 - 8.9	0.4
23	7.9	7.8 - 8.1	8.1	7.9 - 8.2	0.2	7.9	7.6 - 8.2	8.3	8.0 - 8.6	0.4
24	7.7	7.6 - 7.9	7.9	7.8 - 8.0	0.1	7.8	7.5 - 8.1	8.0	7.8 - 8.3	0.3
25	7.4	7.3 - 7.6	7.6	7.4 - 7.7	0.1	7.5	7.2 - 7.8	7.8	7.5 - 8.1	0.3
26	7.1	7.0 - 7.3	7.3	7.2 - 7.4	0.1	7.2	6.9 - 7.5	7.5	7.3 - 7.8	0.4
27	6.9	6.7 - 7.0	7.0	6.9 - 7.1	0.1	6.9	6.7 - 7.2	7.3	7.0 - 7.6	0.4
28	6.6	6.5 - 6.7	6.7	6.6 - 6.9	0.1	6.7	6.4 - 7.0	7.0	6.7 - 7.3	0.3

Appendix B

Equal sickness absence rate

How much do we need to proportionally affect the hazard rate to get the same effect on expected duration of monitoring at day 15 instead of at day 8? The survival can be expressed as:

$$S(t) = e^{-\sum_{i=0}^{t-1} h(i)} \tag{1}$$

Furthermore, the expected duration is given by:

$$E[D] = \sum_{t=1}^{\infty} S(t) \tag{2}$$

Given maximum spell length of 365 days in the data, the expected duration for the control group (monitoring at day 8) is 11.86 and for the treatment group (monitoring at day 15) 12.64. This gives:

$$E[D_C] = \sum_{t=1}^{365} S_C(t) + \sum_{t=366}^{\infty} S_C(t) = 11.86 + \sum_{t=366}^{\infty} S_C(t) \tag{3}$$

$$E[D_T] = \sum_{t=1}^{365} S_T(t) + \sum_{t=366}^{\infty} S_T(t) = 12.64 + \sum_{t=366}^{\infty} S_T(t) \tag{4}$$

The hazard rates for durations longer than 365 days are assumed to be equal for the control and the treatment group. As the data does not allow us to have a greater follow-up period than 365 days, the hazard rate from the 365th day of duration and onwards is set to the mean hazard rate of the last 30 days of the 365 day follow-up period. More precise, the hazard rate from day 365 and onwards (h_{365}) is set to 0.001890.

The last term in (3) can be rewritten as:

$$\sum_{t=366}^{\infty} S_C(t) = \sum_{t=366}^{\infty} e^{-\sum_{i=0}^{t-1} h_C(i)} =$$

$$= \sum_{t=366}^{\infty} e^{-\sum_{i=0}^{364} h_C(i)} e^{-\sum_{365}^{t-1} h_{365}} =$$

$$= S_C(365) \sum_{t=366}^{\infty} e^{-\sum_{365}^{t-1} h_{365}} =$$

$$= S_C(365) \sum_{t=366}^{\infty} e^{-(t-365)h_{365}}$$

(5)

It is easy to show that:

$$\sum_{t=366}^{\infty} e^{-(t-365)h_{365}} = \frac{e^{-h_{365}}}{1 - e^{-h_{365}}}$$

(6)

Putting this into (5) yields:

$$\sum_{t=366}^{\infty} S_C(t) = S_C(365) \frac{e^{-h_{365}}}{1 - e^{-h_{365}}}$$

(7)

The survival rates for the control and treatment groups at day 365 are not significantly different from each other. Due to this, the survival rate at day 365 is set to the mean of the two groups' survival rate at this time of duration. Using this together with (7) in (3) yields:

$$E[D_C] = 11.86 + S_C(365) \frac{e^{-h_{365}}}{1 - e^{-h_{365}}} =$$

$$= 11.86 + 0.0067 \frac{e^{-0.00189}}{1 - e^{-0.00189}} = 15.40$$

(8)

This can be calculated in the same way for the treatment group, which gives an expected duration for the treatment group of 16.18.

The expected duration for the control group with a proportionally affected hazard rate can be expressed as:

$$\mathrm{E}\left[D_X\right] = \sum_{t=1}^{\infty} S_X\left(t\right) = \sum_{t=1}^{\infty} e^{-\sum_{i=0}^{t-1} h_X(i)} = \sum_{t=1}^{\infty} e^{-\sum_{i=0}^{t-1}(1+p)h_C(i)} =$$

$$= \sum_{t=1}^{\infty} e^{-(1+p)\sum_{i=0}^{t-1} h_C(i)} = \sum_{t=1}^{365} e^{-(1+p)\sum_{i=0}^{t-1} h_C(i)} + \sum_{t=366}^{\infty} e^{-(1+p)\sum_{i=0}^{t-1} h_C(i)} =$$

$$= \sum_{t=1}^{365} \left[e^{-\sum_{i=0}^{t-1} h_C(i)} \right]^{(1+p)} + e^{-(1+p)\sum_{i=0}^{364} h_C(i)} \sum_{t=366}^{\infty} e^{-(1+p)(t-365)h_{365}} =$$

$$= \sum_{t=1}^{365} \left[S_C\left(t\right) \right]^{(1+p)} + S_C\left(365\right)^{(1+p)} \frac{e^{-(1+p)h_{365}}}{1-e^{-(1+p)h_{365}}}$$

(9)

Setting the expected duration of the proportionally affected control group hazard rate equal to the expected duration of the treatment group solves the puzzle. As an analytical solution is not possible to obtain, the following equation is numerically solved:

$$\sum_{t=1}^{365} \left[S_C\left(t\right) \right]^{(1+p)} + S_C\left(365\right)^{(1+p)} \frac{e^{-(1+p)h_{365}}}{1-e^{-(1+p)h_{365}}} = 16.18 \,,$$

which gives:

$$p = -\,0.016$$

A proportional decrease of the hazard rate of 1.6 percent yields the same aggregated sickness absence rate in the society as an extension of the no-monitoring period from 7 to 14 days.

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

Rapporter/Reports

2005:1 Ahlin Åsa & Eva Mörk "Vad hände med resurserna när den svenska skolan decentraliserades?"

2005:2 Söderström Martin & Roope Uusitalo "Vad innebar införandet av fritt skolval i Stockholm för segregeringen i skolan?"

2005:3 Fredriksson Peter & Olof Åslund "Påverkas socialbidragsberoende av omgivningen?"

2005:4 Ulander-Wänman Carin "Varslad, uppsagd, återanställd. Företrädesrätt till återanställning enligt 25 § LAS i praktisk tillämpning"

2005:5 Isacsson Gunnar "Finns det en skillnad mellan samhällets och individens avkastning på utbildning?"

2005:6 Andersson Christian & Iida Häkkinen "En utvärdering av personalförstärkningar i grundskolan"

2005:7 Hesselius Patrik, Per Johansson & Laura Larsson "Hur påverkar kravet på läkarintyg sjukfrånvaron? Erfarenheter från ett socialt experiment"

Working Papers

2005:1 Ericson Thomas "Personnel training: a theoretical and empirical review"

2005:2 Lundin Martin "Does cooperation improve implementation? Central-local government relations in active labour market policy in Sweden"

2005:3 Carneiro Pedro, James J Heckman & Dimitriy V Masterov "Labor market discrimination and racial differences in premarket factors"

2005:4 de Luna Xavier & Ingeborg Waernbaum "Covariate selection for non-parametric estimation of treatment effects"

2005:5 Ahlin Åsa & Eva Mörk "Effects of decentralization on school resources"

2005:6 Cunha Flavio, James J Heckman & Salvador Navarro "Separating uncertainty from heterogeneity in life cycle earnings"

2005:7 Söderström Martin & Roope Uusitalo "School choice and segregation: evidence from an admission reform"

2005:8 Åslund Olof & Peter Fredriksson "Ethnic enclaves and welfare cultures – quasiexperimental evidence"

2005:9 van der Klaauw Bas, Aico van Vuuren & Peter Berkhout "Labor market prospects, search intensity and the transition from college to work"

2005:10 Isacsson Gunnar "External effects of education on earnings: Swedish evidence using matched employee-establishment data"

2005:11 Abbring Jaap H & Gerard J van den Berg "Social experiments and instrumental variables with duration outcomes"

2005:12 Åslund Olof & Oskar Nordström Skans "Measuring conditional segregation: methods and empirical examples"

2005:13 Fredriksson Peter & Bertil Holmlund "Optimal unemployment insurance design: time limits, monitoring, or workfare?"

2005:14 Johansson Per & Per Skedinger "Are objective measures of disability reliable?"

2005:15 Hesselius Patrik, Per Johansson & Laura Larsson "Monitoring sickness insurance claimants: evidence from a social experiment"

Dissertation Series

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

2003:2 Ekström Erika "Essays on inequality and education"