

Hesselius, Patrik

**Working Paper**

## Does Sick Absence Increase the Risk of Unemployment?

Working Paper, No. 2003:15

**Provided in Cooperation with:**

Department of Economics, Uppsala University

*Suggested Citation:* Hesselius, Patrik (2003) : Does Sick Absence Increase the Risk of Unemployment?, Working Paper, No. 2003:15, Uppsala University, Department of Economics, Uppsala, <https://nbn-resolving.de/urn:nbn:se:uu:diva-4483>

This Version is available at:

<https://hdl.handle.net/10419/82733>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Does Sick Absence Increase the Risk of Unemployment?\*

Patrik Hesselius<sup>†</sup>

March 21<sup>st</sup>, 2003

## Abstract

The paper examines whether a worker's sick absence behaviour influences the risk of becoming unemployed. Swedish panel data are used to estimate the relationship between the incidence and duration of sick leave and subsequent unemployment. The results indicate that an increase in the number of sick leaves as well as an increase in the duration of sick spells is associated with higher risk of unemployment. Women have a significantly higher risk than men of unemployment associated with sick leave spells longer than 28 days. An implication of the results is that less absence-prone workers are more likely to remain employed in a recession. This in turn may in part explain the pro-cyclical pattern of aggregate Swedish sick absence rates.

**Keywords:** Sick leave, sickness insurance, health, unemployment, duration analysis.

**JEL-codes:** C41, H55, I18, J64.

---

\* I am grateful to Meredith Beechey, Per-Anders Edin, Magnus Gustavsson, Bertil Holmlund, Per Johansson, Roope Uusitalo and seminar participants at the Department of Economics, Uppsala University for helpful comments and suggestions. A special thanks to Edward Palmer and Ola Rylander at the National Social Security Board for providing some of the data. Financial support from Jan Wallander and Tom Hedelius' Research Foundation and Swedish Model in Transition (SMIF, ALI) are gratefully acknowledged.

<sup>†</sup> Correspondence: Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden, e-mail: patrik.hesselius@nek.uu.se.

## 1 Introduction

The purpose of this paper is to test if sick absence is associated with an increased risk of future unemployment. Swedish panel data are used to estimate the effect on the risk of unemployment of previous sick-leave behaviour. The transition into unemployment is studied from January 1<sup>st</sup> 1992 to December 31<sup>st</sup> 1998 and the absence record is observed January 1<sup>st</sup> 1989 to February 28<sup>th</sup> 1991.

The everyday choice for the worker, to be or not to be present at work, will be affected by present and future repercussions of the choice. For the majority of workers in some European countries, the compensation level for sickness absence has been close to the normal wage during the last decades (Nyman et al., 2002). A smaller income loss during a sick leave spell may imply a low incentive to attend work. However, some countries that have relatively generous compensation schemes have a relatively low absence rate; Germany is a case in point.<sup>1</sup> This suggests that other factors also affect the everyday attendance choice.

The sick leave utilisation may yield a productivity signalling effect; a less absent worker is more attractive from the employer's perspective, especially if the employer pays directly parts of the costs associated with a sick spell. A worker with an extensive absence record may be tagged as expensive and less productive, which in turn may lead to dismissal (Yaniv, 1991). The penalty may exist, even though the individual has to use a sick spell to recover from illness to restore productivity, due to that the employer may have problems monitoring the health among employees. If one disregards the disutility of being sick, the cost for the individual can be divided into a direct cost of the sick leave, the possible income loss, and an indirect cost which can be seen as the possible effect on future labour market outcomes. In the analysis of the economic incentives of calling in sick, one should have in mind that some sort of non-economic work moral effect is likely.

Some attention during the last decade, among economists, has been given to absence due to sickness. This especially in Sweden where the absence rate, in 1990, was one of the highest among the industrialized countries and thereafter had a

---

<sup>1</sup> An international overview of the sickness absence rate is presented in Barmby et al. (2000).

rapid decline followed by an increase in the end of the 1990s. The changes in the Swedish sick leave insurance system during the same time period were many and resulted in large differences in the sick leave compensation rate over time. Palme and Johansson (1996) and Cassel et al. (1996) have shown that economic incentives seem to matter for the incidence of sick leave. They also find that the variation in the sick leave rate can not solely be explained by higher costs of absence; the findings shows on a significant inverse relationship between the unemployment rate and the incidence of sick absence. Johansson and Palme (1998) show that the unemployment rate affect the incidence of sick leave negatively and the incidence of recovery positively. These findings suggest a pro-cyclical pattern of the sick leave rate, which later also has been shown by e.g. Arai and Skogman-Thoursie, (2001), Askildsen et al., (2000), Askildsen et al., (2002) and Henrekson and Persson, (2001).

Leigh (1985) discusses the theoretical pro-cyclical relationship between the sick leave rate and the unemployment rate. He claims that besides an eventual disciplining effect the unemployment rate has on the sick-leave rate, a labour force composition effect is likely if absence prone workers have a higher incidence of unemployment in an economic down-turn. A fear of being laid-off that reduces the number of annual absence days is credible, if one finds a positive relationship between sick leave utilisation and the risk of getting unemployed.

The main result of the estimated models in this essay is that the both the incidence and the duration of a sick leave is associated with an increased risk of unemployment. Some evidence shows that women have a higher unemployment risk associated with long sick leave spells. Older male workers seem to have higher risk of unemployment than young male workers, given equal work experience, which may partially be explained by health differences over age.

The next section discusses the relationship between sick absence and unemployment. Section 3 presents the data, how the sample was selected and defines some key variables. Section 4 discusses sample issues, describes the estimation method and the model to be estimated. Section 5 presents the results and finally section 6 concludes.

## 2 Theoretical issues

The everyday decision for the worker, if she is going to attend work or not, cannot exclusively be seen as a choice between earning the normal wage rate or being compensated through the sickness benefit system. Health status is deemed to be important when valuing the two choices; a cold will raise the effort needed to remain as productive as in a healthier state. According to efficiency wage models<sup>2</sup>, the non-work opportunity will be more attractive for the worker if the required effort level is higher. The labour market outcome, such as the subsequent risk of unemployment, may also be an important factor in the valuation of an absence spell.

Health is one kind of human capital that is a little bit different than education since workers often start with a large health endowment that must be continuously replenished (Currie and Madrian, 1999). A sick leave can therefore be seen as an investment in health, which can prevent future sick leaves (Gilleskie, 1998). Therefore it may be necessary and valuable for the individual to call in sick once in a while to keep up productivity. The stochastic behaviour of illnesses and accidents over time affects the transition rate from work to sick absence, i.e. the incidence intensity. Return from a sick leave, the intensity of cure, can also be seen as a stochastic process, where individual recovery is a result of both personal characteristics and the level of illness. The flows into and out of sick leave are also likely to be affected by the economic incentives the individual faces to attend work.

Previous literature has various explanations of factors other than health that may affect attendance behaviour. The worker can be using sick leave, compensated by the employer or the public sick-pay system, as a way to compensate for bad job conditions or low wages. This can be seen as on the job shirking as described by Shapiro and Stiglitz (1984). Barmby et al. (1994) has formulated a theoretical relationship between wages, absence and shirking. The probability of being tagged as a less productive worker or a shirker, due to sick absence, is likely to be positively correlated to the number of days absent. Yaniv (1991) points out that the worker is taking a risk by calling in sick, this due to the fact that misuse of sick

---

<sup>2</sup> For an overview see Layard et al. (1991).

leaves may lead to dismissal and involuntary unemployment. There is also evidence that the sick leave incidence is partially non-health related, e.g. Skogman Thoursie (2002) has evidence of an increased sick leave rate among Swedish men during the Calgary winter Olympics in 1988. The previous absence behaviour of the worker can be seen as a signal for the employer or future employers of the workers' health status and/or shirking tendency.

Leigh (1985) has studied the relationship between unemployment and absenteeism and argues that workers are more absence prone in times of low unemployment. A pro-cyclical absence rate can come from a fear of unemployment during a period of high unemployment, but may also be induced by the composition of the labour force. If workers with low absence rates are retained during economic recessions and more absence prone workers are laid-off, a pro-cyclical pattern of the aggregate absence rate will emerge.

Le Grand (2000) and Arai and Vilhelmsson (2001) have estimated the risk of unemployment in Sweden during the economic downturn in the 1990s. They find that several individual and employer characteristics explain parts of the unemployment risk, such as gender, age, income, education, number of children, marital status, country of origin, experience, sector and public/private employer. There is also evidence that information about the local labour market, such as local unemployment rate, affects the unemployment risk.

The discussion suggests that the intensity of job separation may be seen as a function of exogenous macro economic variables ( $Y$ ), exogenous individual characteristics ( $X$ ) and the flow into ( $s$ ) and out of ( $c$ ) sick leave:

$$\lambda = \lambda(X, Y, s, c) \tag{1}$$

In Sweden an unemployed person can also be on sick leave and benefit from the government provided sick leave compensation; in this case the flows between unemployment and sickness insurance may be affected if there exists an economic incentive other than a gain in leisure time. The flows between unemployment and sickness insurance are not accounted for in this study. For an analysis of incentives of using sick leaves during an unemployment spell see Larsson (2002).

### 3 Data

#### 3.1 Data description and sampling procedure

The individual data used in the empirical analysis is obtained from LINDA<sup>3</sup>, which is a registry based data set of 3.35 percent of the Swedish population (about 300,000 individuals in the 1990's). To be able to observe the pattern of sick leave behaviour, the database was expanded with sick leave spell data for the period 1989-1991 from the sickness period register obtained from the National Social Insurance Board. Information about the transition into unemployment is obtained from the HÄNDEL<sup>4</sup> database. Demographic variables, collected from other sources and included in LINDA, are also used. The local job destruction rate, measured at municipality level, is used in the analysis to capture local differences in the job destruction.

The sickness period register contains information of all sick spells covered by the sickness insurance; the database does not consist of any additional health information as medical diagnosis or overall health status. For each individual, the number of sick leave spells per year and the mean sick leave spell duration will be used as measures of the pattern of individual sick leave behaviour. These measures are divided into three different spell length groups, namely short sick leave spells (less than 8 days), medium sick leave spells (8 to 28 days) and long sick leave spells (longer than 28 days). Short sick leave spells can be seen as “unverified” sickness since only sick leaves of 8 days or longer require a physician certificate. Long sick leave spells are likely to be recertified due to that the physician certificate often is time limited. The mean sick leave spell duration will be measured as the mean number of *extra* days in the sick leave spells and is obtained by decreasing the mean sick absence spell duration with 1, 8 and 29 for the short, medium and long spell durations, respectively. This, if the individual had at least one observed sick leave. E.g. the number of extra days will be 0 if no sick leave spell is observed, in a short sick spell of 1 day will the number of extra days be 0 and for a 2 day spell will the number of extra days be 1. The initial effect of the first day of a sick spell will be captured by the number of sick spells. This to avoid some of the collinearity

---

<sup>3</sup> Longitudinal INdividual DAtabase, for an overview see Edin and Fredriksson (2000).

<sup>4</sup> Included in LINDA and are job search data from the Swedish National Labour Market Administration.

between the number of spells and the mean spell duration, where the mean extra days in a sick spell solely will capture the mean effect of an extra day of sick leave. In the analysis, both full-time and part-time sick leaves will be counted as a sick leave spell.

The pattern of sick leave behaviour is measured during the period January 1989 to February 1991. During this period, the government provided sick leave replacement rate was constant at 90 percent of the wage up to a given ceiling.<sup>5</sup> One other reason to select this period is that there exists no register data of sick leave spell durations shorter than 15 days from 1992 onwards, due to the introduction of a 14 employer-paid day sick pay period in January 1992.<sup>6</sup>

The studied population is individuals 25 to 54 years of age, defined as fulltime employed during the entire period 1989 to 1991, enabling comparability with the sick leave measure. Individuals that receive their main income from self-employment are excluded from the sample because they are likely to have different incentives and possibilities to use (or not use) sick leave compensated by the National Social Security Board. The age restriction is imposed as young individuals may be in education and workers over 54 are more likely to retire during the period of study. Because the LINDA database contains no information about employment status, the individual will be regarded to be employed during a calendar year if all following criteria are satisfied: income above a given threshold, has not received any unemployment benefits, has not received any government provided study grants, has not received any income due to military service and not received any pension. The income threshold is set to the lowest quartile of the earning distribution for full time employed janitors within the public sector (follows the set-up in Edin et al., 2000). In 1990, this threshold is set to 118,000 SEK and the other years' thresholds are calculated by multiplying the threshold for 1990 with the earning development for public employed personnel<sup>7</sup>. The earnings consist of labour income and income from employment insurances as sick benefits and parental leave benefits.

---

<sup>5</sup> Some employers covered the non-benefited 10 percent of the wage during a sick spell and some employers also added up over the ceiling in the sickness insurance system during this time period.

<sup>6</sup> Sick pay is the salary paid by the employer during the first days of a sick leave and the information about sick spells within the sick pay period is not fully reported to the National Social Insurance Board during the 90's.

<sup>7</sup> The earnings development data were obtained from Statistics Sweden (SCB).



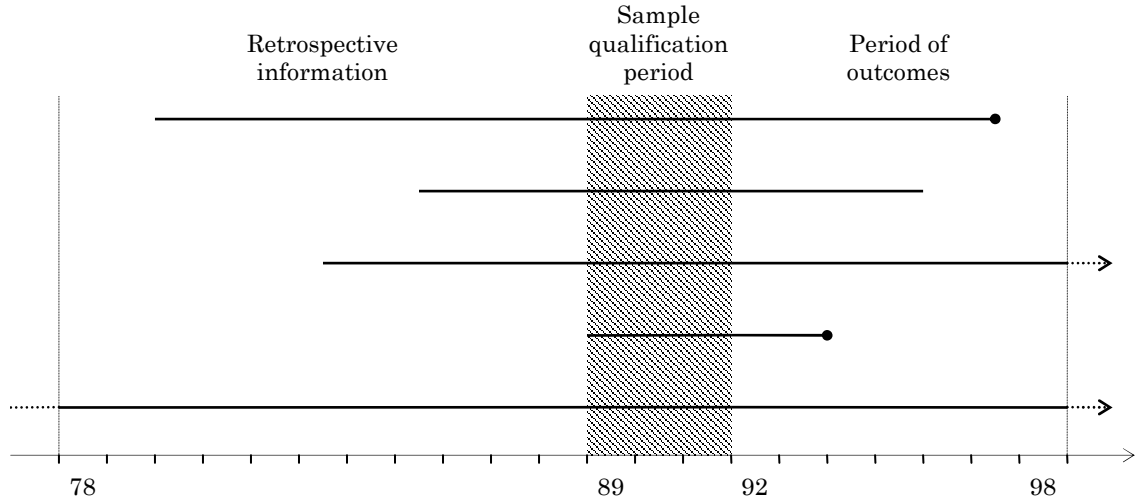
The individuals are followed from January 1<sup>st</sup> 1992 until the first unemployment event or censoring. Registering as unemployed at the local employment office is considered as an event and censoring occurs when the study period ends on the last of December 1998 or if the employment criteria above do not bind. If the employment criterion does not bind one year, the censoring date is set to December 31<sup>st</sup> of the preceding year. This means that if one observation is considered as censored in 1992, it is removed from the data set and does not add any information to the analysis.

The length of an individual's employment record is likely to affect the risk of unemployment due to the value of experience or because of labour market regulations in Sweden. In Sweden it is regulated by law that the "Last in, first out" principle<sup>8</sup> should be followed when workplaces are downsizing, suggesting that tenure might be an important factor in predicting the risk of unemployment. The length of the last continuous full time employment spell is used as a proxy for tenure and is defined as how many consecutive years before 1989 the individual had an income above the threshold and did not received unemployment benefits. Retrospective labour incomes and unemployment benefit pay-outs are studied using annual data from 1978 and individuals with unbroken employment throughout all studied years are considered having at least 14 years of unbroken experience. This yields 12 distinct experience groups. New entrances in the LINDA database (to replace deceased persons or emigrants) are treated as if they did not have any income when not observed. It is likely that the new entrances in the database do not deviate from the sick leave behaviour distribution and the unemployment risk distribution of individuals that have been followed throughout the entire period of study. Together with the small proportion these observations represent, it is unlikely that this will affect the estimates of interest.

To control for previous work absence the fraction of benefits received relative to total employment income is collected from the National Social Security Board for the period 1986 to 1988. The benefits paid out by the National Social Security Board can include: sick leave compensation, parental leave benefits, care of sick child compensation and compensation to look after a disabled person.

---

<sup>8</sup> Regulated by law and stipulates that the last worker employed will be the first one to be laid off if the workplace is downsizing. This law can be sidestepped if the employer and the union decide on another order of priority.



**Figure 1.** Visualisation of the sampling procedure, where dots represent an event.

Compensation during labour market training and military training is also included in the benefit measure, but would not affect the calculated measure as workers are defined as non-employed if they are eligible for these compensations.

The job destruction rate is used to catch the local negative shocks to employment and is defined as the sum of all negative employment changes at plants within a municipality divided by the average employment at the plants in time  $t$  and  $t - 1$ . For an extensive description of this measure and the underlying data, see Johansson (2002).

To summarise the sampling procedure (illustrated in figure 1); an individual has to be full time employed during 1989 to 1991 and not in unemployment, education, military service or receiving a pension to be eligible in the sample. Individuals are followed from January 1<sup>st</sup> 1992 to December 31<sup>st</sup> 1998 where the transition from full time employment to unemployment is observed. To capture plausible differences in risk of unemployment, due to experience, the retrospective unbroken employment spells will be considered in the analysis.

### 3.2 Descriptive statistics

Mean values for the included variables are presented in table A1 in the appendix for all observations, for those who had an unemployment spell during the observed

period and for the ones that did not register as unemployed during the observed period. The pattern of sick leave utilisation is clearly different among the workers who experience a period of unemployment between 1992 and 1998 compared with workers who were not unemployed during the same period. Workers who got unemployed had on average 2.3 short sick leave spells, compared to 1.8 for those employed continuously. This pattern persists over the other sick leave spell duration measures.

The sick benefit ratio of total labour income shows the same pattern as the sick leave measures. Workers who experienced at least one unemployment period between 1992 and 1998 had on average a 4.0 percent income fraction from sick benefits during the period 1986 to 1988, whereas workers who remained employed during the entire period had on average a previous sick benefit ratio of 3.0 percent.

Women are a smaller proportion of those experiencing unemployment, 25.0 percent against 34.6 percent among those with continuous employment record. The young are more heavily represented among the ones who experienced at least one unemployment spell. Income in 1992 is lower in the group that became unemployed (190,200 SEK against 213,200 SEK) and the workers that did not experience an unemployment period were more likely to be married, living together or have children under 16 years of age. Immigrants are somewhat overrepresented among those reported unemployed and workers with less education are more likely to face a period of unemployment. Working in the public sector also seems to be associated with a lower probability of becoming unemployed. A higher fraction of workers in the construction sector and a lower fraction of workers in the service sector became unemployed during the 1992 to 1998.

The number of workers in the sample is 51,136 and 7,079 of them experienced at least one event during 1992 to 1998. In table 1, the numbers of workers are presented for the different experience length groups, where a right censored observation is equivalent to a non-event. The percentage of right censored workers increases as experience increases, which indicates a decreased risk of unemployment as the number of years employed full-time increases. Almost 23 percent of the workers with three years of experience faces at least one period of

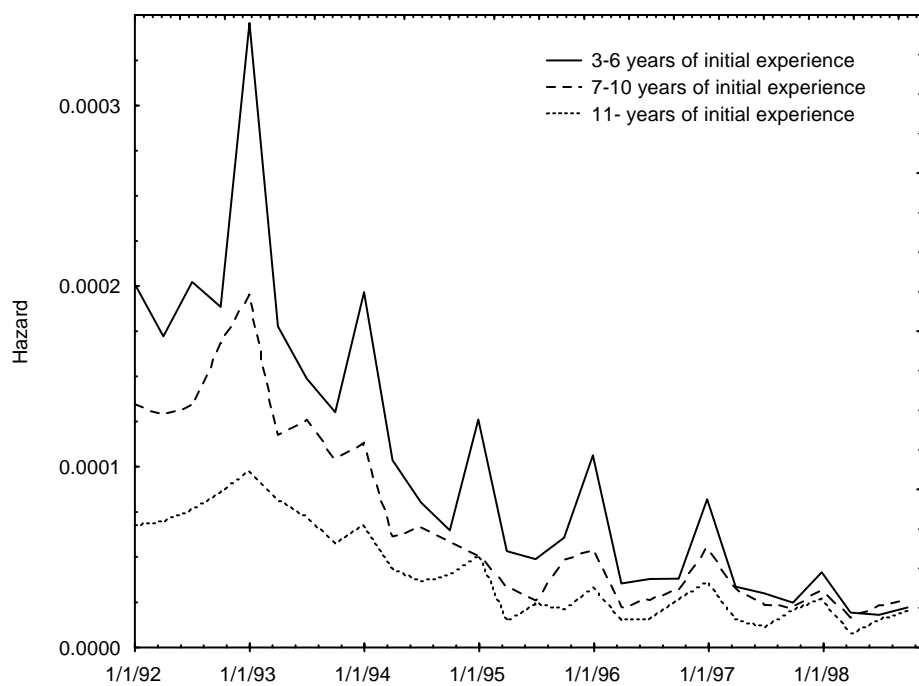
unemployment during 1992 to 1998 and for workers who had at least 14 years of experience the same figure is 9 percent.

As discussed above it is likely that workers with less seniority have a higher probability of experiencing a period of unemployment than workers with longer experience. As shown in figure 2, the risk of becoming unemployed is declining with experience, both between experience groups and over time. The unemployment hazard will also vary over time as the economy is affected by business-cycle fluctuations.

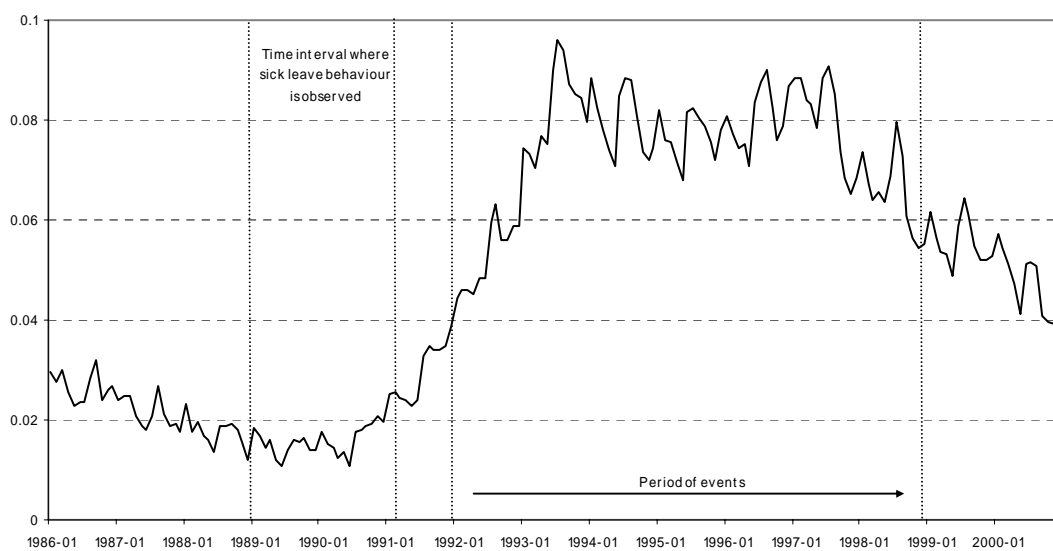
The unemployment rate in Sweden was low and relatively constant during the sample qualification period 1989 to 1991, covering a period of relative economic prosperity. Thereafter, as can be seen in figure 3, the unemployment rate increased rapidly and reached 8 percent in the mid 1990s. The period of events, in this study, coincides with the downturn in the economy, which in turn yields many observations with at least one unemployment period.

**Table 1.** Number of observations in each group of experience.

<b>Initial years of experience</b>	<b>Total</b>	<b>Event</b>	<b>Right censored</b>	<b>Percent right censored</b>
3	3,317	752	2,565	77.33
4	4,168	934	3,234	77.59
5	3,446	670	2,776	80.56
6	3,017	529	2,488	82.47
7	3,293	571	2,722	82.66
8	2,955	450	2,505	84.77
9	3,221	463	2,758	85.63
10	2,445	333	2,112	86.38
11	1,677	195	1,482	88.37
12	2,279	252	2,027	88.94
13	1,754	200	1,554	88.60
≥14	19,564	1,730	17,834	91.16
<b>Total</b>	<b>51,136</b>	<b>7,079</b>	<b>44,057</b>	<b>86.16</b>



**Figure 2.** Hazard rate estimates for different experience lengths over the studied time period.



**Figure 3.** The unemployment rate over the studied period.  
(Source: Swedish Labour Survey, AKU, SCB)

## 4 Econometric method and modelling

### 4.1 Sample issues

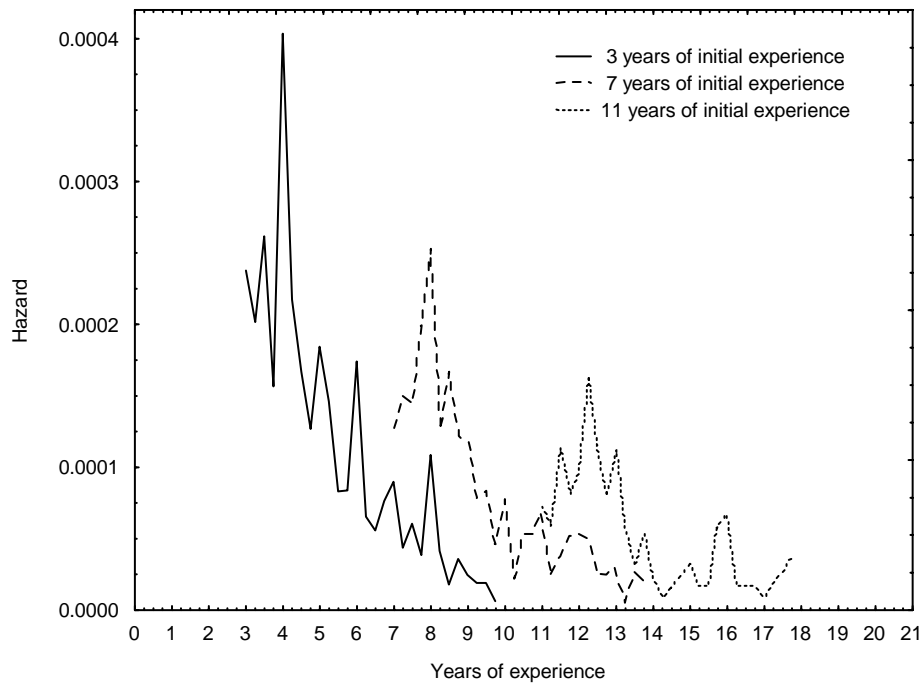
The data sample used has some properties to take into account in the econometric analysis. A sample qualification period gives the estimation problems associated with stock sampling. Using a requirement of three year of subsequent employment to be included in the sample gives that the estimates will be conditional. In data there is no information if an unemployed worker has been laid off or quitted by own means and if the worker finds a new job before the previous job ends. The estimated risk of unemployment captures both the job separation risk and the risk of entering unemployment, conditional on a separation.

The intensity of the transition from employment into unemployment can be captured in several dimensions. First, calendar time, which reflects the macro economic conditions; second, duration of employment (elapsed duration), to capture the human capital investment of tenure and/or experience.<sup>9</sup> In figure 2, the unemployment hazard rate is plotted for different experience length groups over calendar-time. This can be compared with figure 4, where the risk of getting unemployed is plotted over the duration of employment for three different initial experience lengths. If the risk is due only to seniority, one would expect that the lines in figure 4 would connect. The decreasing risk with experience in figure 2 and 4 gives that a hypothesis that the risk is independent of experience is likely to be rejected. It appears that the unemployment risk may be a function of experience but also an outcome of economic shocks in calendar-time, both local and general. Major shocks may close down entire plants and thus all employees, regardless of tenure, have to look for a new job or face unemployment. The estimation in this study will therefore be over calendar time, which follows Imbens (1994) where the calendar time dependence is left unspecified and the duration dependence parameterised. By using Cox partial likelihood, the hazard rate variations of macro economic shocks that affect all individuals in the economy will cancel out.

The sample qualification period, used to certify comparable sick leave measures, gives that the probability to be included in the sample is increasing with

---

<sup>9</sup> A more detailed description about the different ways of modelling risk sets and their implications for the estimation process can be found in Tunali and Pritchett (1997).



**Figure 4.** Hazard rate estimates for different experience lengths plotted against the duration of employment.

employment duration. The main problem with this left truncated data set is that it has an overrepresentation of long durations, which is in employment duration (time from hiring to unemployment) gives a downward bias to the hazard rate. Estimating the hazard rate in calendar-time yields an ambiguous bias, downward due to the overrepresentation of long duration and an ambiguous effect due to delayed entrance time. One way to treat the bias problem, due to overrepresentation of long durations and the delayed entry time, is to perform the analysis separate for different experience lengths and at the same time restrict the parameters to be equal over all groups. Individuals with different length of experience are likely to have separate hazard functions; the exposure to risk in calendar-time occurs at different phases in employment duration between experience groups.

The event time of unemployment is defined by a calendar date. This gives that the duration is measured in days and can be seen as continuous due to the large number of days during the period studied.

## 4.2 Econometric modelling

The econometric estimates of the duration of employment in this paper are based on the semiparametric proportional hazard model (Cox, 1972). The classic Cox model treats the covariates as constant during the time exposed to risk, whereas a refinement of the Cox model allows for covariates to vary during the time exposed to risk.

Let  $h(t | Z(t))$  define the hazard rate at time  $t$  for an individual with time dependent risk vector  $Z(t)$ . The basic model due to Cox (1972), extended with time varying covariates is as follows:

$$h(t | Z(t)) = h_0(t) \cdot e^{Z(t)\beta} \quad (2)$$

where  $h_0(t)$  is an arbitrary baseline hazard rate and  $\beta$  is a parameter vector. This is called a semi parametric form because only the covariate effect is specified in a parametric form. The Cox model belongs to the family of proportional hazard models, which comes from the fact that the ratio between two individuals with covariate values  $Z(t)$  and  $Z^*(t)$  is constant when there are no time dependent covariates.

The sample includes left truncated data. For left truncated and right censored data, only a conditional distribution of the time to event can be consistently estimated without parametric assumptions (Gross and Lai, 1996). If the event time and the truncated duration are conditionally independent, given the covariates  $Z$ , it is straightforward to show that the conditional hazard rate and the unconditional hazard rate are equal (Klein and Moeschberger, 1997). This assumption is unrealistic with the data used here since the risk of unemployment is likely to be correlated with experience as discussed in section 3.

One way to treat this problem is to stratify over experience allowing the baseline hazard functions to be different and completely unrelated between strata (Klein and Moeschberger, 1997). To incorporate the stratification, the hazard function (2) is rewritten as:



$$h_s(t | Z(t)) = h_{0s}(t) \cdot e^{Z(t)\beta}, \quad s = 1, \dots, S \quad (3)$$

where  $S$  is the total number of strata. A key assumption in using the stratified model is that the covariate effect is equal over the baseline hazard in each stratum.

Assume that the event times are distinct and ordered. The partial likelihood for each strata, based on the hazard function as specified by (3), is then given by:

$$L_s(\beta) = \prod_{i=1}^D \frac{\exp\left(\sum_{b=1}^p \beta_b Z_{ib}(t_i)\right)}{\sum_{j \in R_s(t_i)} \exp\left(\sum_{b=1}^p \beta_b Z_{jb}(t_i)\right)} \quad (4)$$

where  $R_s(t_i)$  is the stratum specific risk set in time  $t_i$  (defined by the number of individuals, in a given stratum, exposed to risk just prior to  $t_i$ ), the event times are  $t_1, t_2, \dots, t_D$  and the covariate vector for an individual is defined by  $Z_j(t) = [Z_{j1}(t), \dots, Z_{jp}(t)]$ . The partial likelihood function for the entire maximisation problem is specified as:

$$L(\beta) = L_1(\beta) \cdot L_2(\beta) \cdot \dots \cdot L_S(\beta) \quad (5)$$

The above specified likelihood function assumes that there are no ties between the event times. Due to the way of recording the event times, ties often occur in data. There exist a number of methods to handle this phenomenon, and one popular alternative specification of the partial likelihood function is Breslow's likelihood<sup>10</sup>. Let  $t_1 < t_2 < \dots < t_K$  denote the  $K$  distinct, ordered, event times and  $d_i$  be the number of events at  $t_i$ . Furthermore let  $s_i$  be the sum of the covariate vectors  $Z_j(t)$  over all individuals who has an event at time  $t_i$ . Breslow's ties-handling partial likelihood, for each stratum, is expressed as:

---

<sup>10</sup> Founded in Breslow (1974).

$$L_s(\beta) = \prod_{i=1}^K \frac{\exp\left(\sum_{b=1}^p \beta_b s_{ib}(t_i)\right)}{\left[\sum_{j \in R_s(t_i)} \exp\left(\sum_{b=1}^p \beta_b Z_{jb}(t_i)\right)\right]^{d_i}} \quad (6)$$

In the likelihood function above (6), each of the  $d_i$  events are considered as distinct and the contribution to the likelihood in time  $t_i$  of the tied events is calculated by multiplying one event's share over all ties. This method works quite well if there are few ties<sup>11</sup>. The duration in this study is measured in days, which means that the event space is relative large and therefore results in few ties. There are also other ways to correct for ties, e.g. exact likelihood. When using time-dependent covariates, the Breslows likelihood was chosen due to that some of the other methods are exploding in number of permutations in the likelihood function.

The existence of unobserved heterogeneity is hard to deny, but is not accounted for in the econometric model above. Most models who control for heterogeneity is parametric and assumes e.g. independence between the heterogeneity and the covariates. This assumption is likely to be violated in this study and is hard to relax when the covariates are time-dependent (Wooldridge, 2002). It has been shown that the effect of neglected unobserved heterogeneity on partial likelihood estimates yields a bias towards zero for the parameter estimates, whilst the estimated standard errors are approximately correct (Lancaster, 1990). Which one to choose, between a parametric model with heterogeneity control and a semi-parametric model, is not obvious. A rule of thumb (Wooldridge, 2002):

*“Introducing unobserved heterogeneity is indistinguishable from simply allowing a more flexible duration distribution.”*

The described semi-parametric estimation procedure may give a bias towards zero for the parameter estimates, which should be compared to an ambiguous bias a wrongful parameterisation of the hazard function yields.

The intensity of unemployment may, as discussed in section 2, be a function of macro economic, employer and individual parameters, where sick leave behaviour

---

<sup>11</sup> See Farewell and Prentice (1980) for a discussion about bias problems associated with ties.

is likely to affect the unemployment risk via a signalling effect of productivity and/or due to loss of experience caused by absence. Building upon equation (1) and applying the properties of a proportional hazard model yields the following model:

$$\lambda(t | X_t, Y_t, s, c) = \lambda_0(t) e^{X_t\phi + Y_t\gamma + s\alpha + c\beta} \quad (7)$$

The intensity of unemployment entry,  $\lambda(t)$ , and its baseline counterpart,  $\lambda_0(t)$ , will in the estimation procedure be referred as the hazard function,  $h(t | X, Y, s, c)$ , and the baseline hazard function,  $h_0(t)$ . Macroeconomic effects on the hazard function will to a large extent be included in the baseline hazard function (see discussion in section 4.1). Regional dummies and local job destruction rates are used to catch local labour market effects ( $M_t$ ).

The model to estimate is the stratified model presented above and the different lengths of the last continuous previously experience (measured in years) will be used as stratum divider. Applying the stratification and the changes described above to equation (7) yields:

$$h_s(t | X_t, Y_t, s, c) = h_{0s}(t | Y_t) e^{X_t\phi + M_t\gamma + s\alpha + c\beta}, \quad s = 1, \dots, S \quad (8)$$

The personal characteristic variables ( $X_t$ ) included in the model are sex, age, previous year's income, previous year's family income, number of children less than 16 years of age, married/living together, immigrant Nordic/EU/Other, educational level, type of education, employer type and sector. The variables are mainly defined as dummies; the exceptions are age, previous year's income, previous year's family income and number of children less than 16 years of age. Some of the variables are time-dependent. For more information see the descriptive statistics in table A1 in appendix A. The intensities into ( $s$ ) and out of ( $c$ ) sick leave are proxied by the yearly average number of sick leaves and the average length of sick leave respectively. The sick leaves are divided into three spell length groups to take into account the different signalling effects sick spells with various length may have. To test if an eventual sick absence effect on the unemployment risk is driven by earlier workplace absence, the sick benefit payment relative to total employment income during 1986 to 1988 is included.

## 5 Results

Table 2 consists of the parameter estimates for four model specifications, estimated with stratified Cox partial Likelihood allowing for time-varying covariates using all observations. These models assess if and how the observed sick leave behaviour pattern during 1989 to 1991 affects the risk of unemployment, controlling for an increasing number of individual/employer characteristics. The simplest model, # 1, only consists of the sick leave behaviour measures 1989 to 1991, in model 2 some individual characteristics are added and model 3 consists of all individual characteristics together with employer characteristics. The sick leave utilisation in 1986 to 1988 is introduced in model 4 to capture persistence in the individual sick leave rate induced by e.g. chronic health problems. The relative risk is calculated using  $\exp(\beta)$  and the percentage effect is obtained using  $100[\exp(\beta) - 1]$ , where  $\beta$  is the parameter estimate.

First to be commented is model 3. An increase by one in the number of short sick spells (less than 8 days, with no requirement of sickness certificate) increases the risk of unemployment by 6.3 percent and an extra medium sick spell (8-28 days, with certificate requirement) raises the risk of unemployment by 17.0 percent. The effect of one extra long sick spell (longer than 28 days, with certificate requirement) is relatively high with a 30.0 percent increase of the intensity of unemployment. All of these results are significant at a 1 percent significance level. These effects together can be seen as the effect the intensity of transition into sick leave has on the risk of unemployment. Thus, based on the results above, there is a significant positive relationship between the intensity of transition into sick leave and the risk of unemployment.

The mean duration of sick spells is also associated with an increasing risk of unemployment. An extra day in a short sick spell is penalised with a 7.1 percent higher risk of unemployment. The effects for medium and long sick spell durations are 1.1 percent and 0.08 percent respectively. This implies that there is a negative relationship between the returning-to-work rate and the risk of unemployment.

**Table 2.** Estimates of the duration models (no. individuals = 51,136, events = 7,079).

	Model 1	Model 2	Model 3	Model 4
Number short sick leave spells (1-7 days)	0.04904*** (0.00599)	0.04395*** (0.00609)	0.06077*** (0.00599)	0.05926*** (0.00608)
Number medium sick leave spells (8-28 days)	0.18570*** (0.02964)	0.17354*** (0.02947)	0.15732*** (0.02982)	0.15030*** (0.03016)
Number long sick leave spells (>28 days)	0.22402*** (0.04632)	0.27495*** (0.04649)	0.26198*** (0.04710)	0.25353*** (0.04741)
Mean extra days in short sick leave spell	0.14510*** (0.00825)	0.10157*** (0.00845)	0.06851*** (0.00856)	0.06786*** (0.00857)
Mean extra days in medium sick leave spell	0.01242*** (0.00321)	0.01285*** (0.00322)	0.01064*** (0.00324)	0.01064*** (0.00324)
Mean extra days in long sick leave spell	0.00091*** (0.00011)	0.00082*** (0.00010)	0.00077*** (0.00010)	0.00067*** (0.00012)
Fraction sick benefits of total income, 86-88				0.29783 (0.18973)
Woman		-0.67012*** (0.03067)	-0.34092*** (0.03476)	-0.34061*** (0.03474)
Age		-0.00871 (0.01629)	0.02252 (0.01650)	0.02230 (0.01650)
Age <sup>2</sup>		0.00008 (0.00021)	-0.00032 (0.00021)	-0.00031 (0.00021)
Income (t) (prev. year, 100 000 SEK)		-0.14634*** (0.02695)	-0.07997*** (0.02758)	-0.07883*** (0.02756)
Income <sup>2</sup> (t) (prev. year, 100 000 SEK)		0.00060*** (0.00014)	0.00033** (0.00014)	0.00033** (0.00014)
Household income (t) (prev. year, 100 000 SEK)		-0.11692*** (0.01717)	-0.08562*** (0.01721)	-0.08565*** (0.01721)
Household income <sup>2</sup> (t) (prev. year, 100 000 SEK)		0.00017*** (0.00003)	0.00013*** (0.00004)	0.00013*** (0.00004)
Number of kids (t)		-0.01868 (0.01567)	0.00770 (0.01574)	0.00728 (0.01574)
Married/living together (t)		-0.08842** (0.03741)	-0.16007*** (0.03754)	-0.15927*** (0.03753)
Immigrant: Nordic		0.05748 (0.05795)	0.07220 (0.05870)	0.07180 (0.05870)
Immigrant: Other EU		-0.20001 (0.23627)	-0.16104 (0.23670)	-0.16628 (0.23672)
Immigrant: Non-EU		0.01568 (0.05745)	0.12518** (0.05886)	0.12379** (0.05886)
Local job destruction rate (t)			1.25359*** (0.38297)	1.24697*** (0.38307)
Geographical dummies <sup>a</sup>			Yes	Yes
Education dummies <sup>a</sup>			Yes	Yes
Employer characteristics <sup>a</sup>			Yes	Yes
Log-Likelihood	-58,621	-58,110	-57,372	-57,370

Standard errors in parenthesis. Variables with (t) are time varying covariates.

\*, \*\*, \*\*\* - significant at 10, 5, 1 percent level.

<sup>a</sup> See appendix for a list of the variables.

Some of the personal characteristic variables yield interesting significant parameter estimates. Women have a much lower risk of unemployment compared to men (71.1 percent of the risk for males), given all other explanatory variables. This may be explained by the fact that women are underrepresented in parts of the business-cycle sensitive sectors such as manufacturing and construction, which are not captured by the sector-specific dummies in the model. One other explanation can be a higher proportion of women than men work part-time. The income is lower for a part-time worker compared to a full-time worker, given equal full-time wage. The gender dummy will in this case capture a difference in the income intercept between men and females. This goes hand in hand with the parameter estimates of the income variables, which show a significantly lower risk associated with higher incomes. The household income shows the same pattern and individuals married or living together have a significantly lower risk of unemployment.

Nordic and EU immigrants do not have a significantly different risks of getting unemployed in comparison to natives, whereas non-EU immigrants are significantly more likely to become unemployed than natives. This result goes hand in hand with the overrepresentation of non-EU immigrants among the ones who were registered as unemployed during the period of events. As one can expect, an increase in the local job destruction rate is associated with a significant increase in the risk of unemployment.

There is a possibility that the pattern of sick leave behaviour during 1989 to 1991 is a continuation of previous sick leave behaviour, which in turn can be induced by chronic health problems. The sickness benefits fraction of total individual employment income during 1986 to 1988 is included in the model to capture the previous absence behaviour. The results of the model extended with the absence measure is presented in table II, model 4 (for all covariates see table A1 in appendix).

The significant positive effects of the sick leave behaviour pattern 1989-1991 on the risk of unemployment persist in model 4, compared with model 3. The effect on unemployment risk of both the number of and the mean extra days in sick leaves is

**Table 3.** Estimates for different subgroups.

	Men	Women	Age 25-40	Age 41-54
Number short sick leave spells (1-7 days)	0.06072*** (0.00726)	0.06385*** (0.01141)	0.06040*** (0.00785)	0.05651*** (0.00978)
Number medium sick leave spells (8-28 days)	0.15795*** (0.03554)	0.10961* (0.05757)	0.18937*** (0.03880)	0.08200* (0.04791)
Number long sick leave spells (>28 days)	0.20363*** (0.05871)	0.32768*** (0.08226)	0.26855*** (0.06686)	0.22035*** (0.06817)
Mean extra days in short sick leave spell	0.05892*** (0.00961)	0.08076*** (0.01921)	0.08268*** (0.01244)	0.05408*** (0.01187)
Mean extra days in medium sick leave spell	0.01202*** (0.00374)	0.00574 (0.00656)	0.01014** (0.00445)	0.01204** (0.00475)
Mean extra days in long sick leave spell	0.00070*** (0.00014)	0.00057** (0.00028)	0.00090*** (0.00018)	0.00040** (0.00020)
Fraction sick benefits of total income, 86-88	0.13488 (0.22291)	0.68017* (0.37428)	0.06324 (0.26499)	0.49305* (0.27854)
Woman			-0.08395* (0.04692)	-0.62018*** (0.05208)
Age	0.04339** (0.01932)	-0.03385 (0.03226)	-0.00696 (0.06603)	-0.09170 (0.11716)
Age <sup>2</sup>	-0.00047* (0.00025)	0.00016 (0.00042)	0.00024 (0.00103)	0.00097 (0.00126)
Income (t) (prev. year, 100 000 SEK)	-0.04588 (0.03264)	-0.16839** (0.07649)	-0.05837 (0.04619)	-0.08435** (0.03663)
Income <sup>2</sup> (t) (prev. year, 100 000 SEK)	0.00018 (0.00019)	0.00456 (0.00369)	-0.00213* (0.00114)	0.00034* (0.00019)
Household income (t) (prev. year, 100 000 SEK)	-0.11840*** (0.02499)	-0.04535* (0.02536)	-0.13818*** (0.02969)	-0.07837*** (0.02394)
Household income <sup>2</sup> (t) (prev. year, 100 000 SEK)	0.00191*** (0.00045)	0.00007 (0.00005)	0.00440*** (0.00079)	0.00011* (0.00006)
Number of kids (t)	-0.00032 (0.01850)	0.01529 (0.03332)	-0.01272 (0.02175)	0.00481 (0.02698)
Married/living together (t)	-0.12463*** (0.04528)	-0.22925*** (0.07059)	-0.06597 (0.05559)	-0.19845*** (0.05221)
Immigrant: Nordic	0.05610 (0.07192)	0.11222 (0.10232)	0.04901 (0.08710)	0.08127 (0.08003)
Immigrant: Other EU	-0.09956 (0.25136)	-0.78759 (0.71286)	-0.30801 (0.38020)	-0.09507 (0.30345)
Immigrant: Non-EU	0.08462 (0.06943)	0.15536 (0.11319)	0.07426 (0.08488)	0.09989 (0.08309)
Local job destruction rate (t)	1.31163*** (0.43711)	0.87580 (0.80001)	0.48835 (0.53231)	2.11832*** (0.55075)
Geographical dummies <sup>a</sup>	Yes	Yes	Yes	Yes
Education dummies <sup>a</sup>	Yes	Yes	Yes	Yes
Employer characteristics <sup>a</sup>	Yes	Yes	Yes	Yes
No. observations	34,126	17,010	22,035	29,101
Events	5,308	1,769	3,858	3,219

Standard errors in parenthesis. Variables with (t) are time varying covariates.

\*, \*\*, \*\*\* - significant at 10, 5, 1 percent level.

<sup>a</sup> See appendix for a list of the variables.

not significantly altered when the previous absence proxies are included in the model. No significant effect of the model extension in model 4 can be seen for the other personal characteristic variables.

Four model specifications are presented in table 2, where the fully specified model is the last among the four. The first two models together with the two previous commented models are estimated to perform parameter stability checks, where the first only consists of the covariates of main interest and additional covariates are introduced in the second and third model towards the fully specified model. It is striking that all the sick leave behaviour parameters remain significant at a 1 percent rejection level with preserved signs. The estimated parameters are stable throughout the introduction of new covariates describing individual, employer and geographical characteristics.

Table 3 presents the estimation results of dividing the material, first into two subgroups representing the two genders and second into two age groups (25-40 and 40-54 years of age at start of the study January 1<sup>st</sup> 1992). A comparison between the two genders shows mainly equal effects of the sick leave behaviour on the risk of unemployment, except for two parameter estimates. One interesting result is that the intensity of becoming unemployed due to one extra long sick spell is increased almost twice as much for women as for men. The percentage effect on the hazard rate of one extra day of absence in a medium sick leave spell is not significant at a 5 percent rejection level for women, while the effect for men is positively significant at the 1 percent level. According to this result, the relative punishment for long periods of sick absence in terms of higher risk of getting unemployed is more than twice as big for women as for men, whereas a penalty of an extra day in a medium sick leave spell is significant for men but not for women.

The parameter estimates for men show that increased age is significantly associated with an enhanced risk of unemployment even though the increase is lowered with age. Experience is controlled for in the estimation procedure and the result suggests that an older worker is more likely to register as unemployed. One interpretation might be that an older worker may have an extensive previous unemployment record in comparison to a young worker, holding experience constant. As discussed in section 4.1, the risk of unemployment is a combination of



the risk of getting laid off and the risk of getting hired before the previous employment period ends. It is also likely that a younger individual finds it easier to get hired than an older worker. These two explanations together with a likely depreciation of human capital during unemployment or an out of labour force period support the result (Edin and Gustavsson, 2002).

Estimation of the duration model for young and older workers separately yields no significant parameter difference between the two age groups (column 3 and 4 in table 3). Nevertheless a pattern of higher penalty for young relatively older workers can be seen due to that almost all sick leave behaviour parameters are relatively higher in the young age group.

Comparing the sick leave behaviour parameter estimates of all the subgroups in table 3 with the parameters in the fully specified model (table 2) strengthens the above discussed parameter stability check. All the results presented above show a clear increased risk of getting unemployed of increased use of sick leave.

As discussed in the data section, it is likely that the impact of the pattern of sick leave utilisation 1989 to 1991 weakens as time passes. To be able to capture a change in the effect of the 1989 to 1991 sick leave behaviour parameters over time, they are divided into two to cover the 1992 to 1995 period and the 1996 to 1998 periods separately. This is made by multiplying the sick leave parameters with time dependent dummy variables. The 1992 to 1995 variables take the sick leave

**Table 4.** Estimates for different time periods.

	1992 - 1995	1996 – 1998
Number short sick leave spells (1-7 days)	0.06156*** (0.00643)	0.04491*** (0.01705)
Number medium sick leave spells (8-28 days)	0.14984*** (0.03182)	0.13347 (0.09242)
Number long sick leave spells (>28 days)	0.27775*** (0.04954)	-0.00039 (0.16134)
Mean extra days in short sick leave spell	0.07827*** (0.00917)	0.00346 (0.02307)
Mean extra days in medium sick leave spell	0.01198*** (0.00345)	0.00271 (0.00941)
Mean extra days in long sick leave spell	0.00070*** (0.00013)	0.00022 (0.00057)

Standard errors in parenthesis. \*, \*\*, \*\*\* - significant at 10, 5, 1 percent level. The estimated model is, except for the sick leave behaviour variables above, the same as the model 4 in table 2.

incidence and duration values during 1992 to 1995 and zero during 1996 to 1998; for the period 1996 to 1998, vice versa. Parts of the estimation result can be found in table 4. The parameter estimates for the first period, 1992 to 1995, do not deviate from corresponding results for the entire period (model 4, table 2), whereas the estimates from the later time period show an overall lowered and mainly insignificant effect on the risk of getting unemployed. According to this, the effect of the sick leave behaviour on the risk of unemployment is not persistent over time. This might be explained by a weakened effect over time and together with a decreased individual sick leave utilisation during the period of high unemployment a short persistence of the effect is likely.

## 6 Concluding remarks

The focus of this study has been to test if sick absence increases the risk of subsequent unemployment. Sick leave data from the end of a period of low unemployment is used to measure the pattern of sick leave behaviour. The individuals are followed throughout a period of economic recession where the event of interest is if the workers get unemployed. The empirical analysis presents evidence of an association between sick absence and the risk of unemployment. An increase in both the number of various sick spell lengths and the number of days in the sick spells yields a highly significant increase in the unemployment risk.

An individual's extended use of sick leaves may come from e.g. chronic health problems or from a workplace dissatisfaction that results in usage of the sickness benefit system. Whether or not the individual uses an extended amount of sick leaves, these personal characteristics might be visible for an employer and the individual is classified as unproductive. It is likely that a worker who is regarded as unproductive has an increased risk of getting laid-off. If this explanation holds, the result will be affected by an unobserved productivity and the observed sick leave will work as a proxy for this unobservable.

To control for earlier use of sick leaves, the 1986 to 1988 sickness benefit fraction of total income is included in the analysis. The hypothesis is that if the sick leave behaviour is due to observable personal characteristics then the earlier absence

measure will coincide with the observed sick leave behaviour in 1989 to 1991. The results from the empirical analysis show no significant evidence of this, which in turn may suggest that a penalty associated with sick leaves is a perishable. The weakened effect over time is also evident by estimating the sick leave behaviour effect for the first 4 years and the last 3 years separately in the 1992 to 1998 event period.

Several recent papers (e.g. Arai and Skogman-Thoursie, 2001, Askildsen et al., 2000, Askildsen et al., 2002 and Henrekson and Persson, 2001) have shown that the unemployment rate varies inversely with the sick absence rate and one explanation to this is that a rising unemployment rate has a disciplinary effect on the sick leave rate. The pattern of sick leave behaviour was measured during the end of a period of low unemployment in Sweden (1989-1991), which may give that the absence prone workers were confident and revealed their true sick leave behaviour. The outcome space is during a period of economic recession and even if workers adjust their sick absence behaviour, due to the increased unemployment rate, it is likely that employers remember the sick leave behavioural pattern some years ago. If the employer, during a downturn in the economy, fires workers who showed on extended sick leave behaviour during a time of low unemployment, a disciplining effect on the individual absence rate of high unemployment is likely. Due to the findings that absence prone workers have a higher risk of unemployment during a downturn in the economy, the aggregated absence rate is likely to show a pro-cyclical pattern whether or not a disciplining effect exists.

One limitation of the study is that we cannot observe if a worker is being laid off or quits and if the worker finds a new job before the previous job ends. The estimated risk measures include both an estimate of the job separation risk and of the risk of not being hired before the previous job ends.

The empirical analysis yielded some other interesting results beside the pattern of sick leave behaviour. For example, women have a much lower risk of getting unemployed. This result is not unlikely if one compares the proportion women among the ones who experience at least one unemployment spell with those who remained employed until censored. Some of the difference in risk between women and men may be explained by the interaction with the other explanatory variables

used. Parts of the gender difference may be explained by the income effect on the risk of unemployment, as women have lower income than men due to e.g. part-time work. Non-EU and non-Nordic immigrants have a significantly higher risk of getting unemployed than natives, which follows the result of Arai and Vilhelmsson (2001). The result of this study can not show if the risk differential is due to productivity differentials or discrimination, but Arai and Vilhelmsson (2001) claim that Non-EU and non-Nordic immigrants are discriminated against.

The result also suggests that women have a higher unemployment risk penalty associated with long sick leave spells. One result, for men, is that the risk of unemployment is increasing in age, given all other explanatory variables. This result is surprising at first sight. However, given equal work experience, an older worker may have more unemployment experience. It is also likely that it is easier for a younger individual to get hired, as the risk of unemployment both includes the risk of getting laid-off and the risk of not being hired before the last employment ends.

In sum, this study has shown that higher sick absenteeism is associated with a higher risk of becoming unemployed. More analysis is still needed to study if this relationship is driven by the risk of getting laid off or if it is an outcome of the risk of not getting rehired before the last employment ends.

## References

- Arai M. and Skogman-Thoursie P. 2001, "Incentives and Selection in Cyclical Absenteeism", FIEF Working Paper Series, No. 167.
- Arai M. and Vilhelmsson R. 2001, "Immigrants' and Natives' Unemployment-risk: Productivity Differentials or Discrimination?", FIEF Working Paper Series, No. 169.
- Askildsen J. E., Bratberg E. and Nilsen Ø. A. 2000, "Sickness Absence over the Business Cycle", Department of Economics, University of Bergen, Working papers in Economics: No. 04-2000.
- Askildsen J. E., Bratberg E. and Nilsen Ø. A. 2002, "Unemployment, Labour Force Composition and Sickness Absence: A Panel Data Study", IZA Discussion Papers, No. 466.

- Barmby T. A., Ercolani M. G. and Treble J. G. 2000, "Sickness Absence: An International Comparison", *The Economic Journal*, 112, F315-F331.
- Barmby T. A., Sessions and Treble J. G. 1994, "Absenteeism, Efficiency Wages and Shirking", *Scandinavian Journal of Economics*, 96(4), 561-566.
- Breslow N. 1974, "Covariance Analysis of Censored Survival Data", *Biometrika* 30, March, 89-99.
- Cox D. R. 1972, "Regression Models and Life-Tables (with discussion)", *Journal of the Royal Statistical Society, B* 34, p187-220.
- Cassel C-M., Johansson P. and Palme M. 1996, "A Dynamic Discrete Choice Model of Blue Collar Worker Absenteeism in Sweden 1991", SSE/EFI Working Paper Series in Economics and Finance No 129, Stockholm School of Economics.
- Currie J. and Madrian B. 1999 "Health Insurance and the Labor Market," in *Handbook of Labor Economics*, Ashenfelter O. & Card D. eds, Volume 3C, ch. 50.
- Edin P-A., Fredriksson P. 2000, "LINDA: Longitudinal Individual Data for Sweden", Working Paper 2000:19, Department of Economics, Uppsala University.
- Edin P-A., Fredriksson P., Lundborg P. 2000, "Trade, Earnings, and Mobility – Swedish Evidence", Working Paper 2000:24, Department of Economics, Uppsala University.
- Edin P-A. and Gustavsson M. 2002, "Time out of work and skill depreciation", mimeo, Department of Economics, Uppsala University.
- Farewell V. T., Prentice R. L. 1980, "The approximation of partial likelihood with emphasis on case-control studies", *Biometrika*, 67, 2, 273-278.
- Gilleskie D. B. 1998, "A dynamic stochastic model of medical care use and work absence", *Econometrica*, Vol.66, No. 1, 1-45.
- Gross S. T., Lai T. L. 1996, "Nonparametric Estimation and Regression Analysis With Left-Truncated and Right Censored Data", *Journal of the American Statistical Association*, September, vol. 91, No. 435, 1166-1180.
- Henrekson M. and Persson M. 2001, "The Effects on Sick Leave of Changes in the Sickness Insurance System", SSE/EFI Working Paper Series in Economics and Finance No 444, Stockholm School of Economics.
- Lancaster T. 1990, *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.
- Larsson L. 2002, "Sick of being unemployed? Interactions between unemployment and sickness insurance in Sweden", Working Paper 2002:6, Institute for Labour Market Policy Evaluation (IFAU).
- Layard R., Nickell S. and Jackman R. 1991, "Unemployment: Macroeconomic Performance and the Labour Market", Chapter 3, Oxford, Oxford University Press, 1991.

- Le Grand C. 2000, "On-the-job training, firm resources and unemployment risks: an analysis of the Swedish recession 1991-1993", Working Paper 2000:3, Institute for Labour Market Policy Evaluation (IFAU).
- Leigh J. P. 1985, "The Effects of Unemployment and the Business Cycle on Absenteeism", *Journal of Economics and Business*, 37, 159-170.
- Henrekson M. and Persson M. 2001, "The Effects on Sick Leave of Changes in the Sickness Insurance System", SSE/EFI Working Paper Series in Economics and Finance No 444, Stockholm School of Economics.
- Imbens G.W. 1994, "Transition models in a non-stationary environment", *Review of Economics and Statistics*, Vol. LXXVI, No. 4, 703-720.
- Johansson K. 2002, "Labor market programs, the discouraged-worker effect, and labor force participation", Working Paper 2002:9, Institute for Labour Market Policy Evaluation (IFAU).
- Johansson P. and Palme M. 1996, "Do economic incentives affect work absence? Empirical evidence using Swedish data", *Journal of Public Economics* 59, 195-218.
- Kaivanto K. 1997, "An alternative model of pro-cyclical absenteeism", *Economic Letters* 54, 29-34.
- Klein J. P., Moeschberger M. L. 1997, *Survival analysis: techniques for censored and truncated data*, New York : Springer, 1997.
- Nyman K., Bergendorff S. and Palmer E. 2002, "Den svenska sjukan - sjukfrånvaron i åtta länder", Rapport till ESO (Ds 2002:49), Expertgruppen för studier i offentlig ekonomi, Stockholm.
- Palme M. and Johansson P. 1998, "Assessing the effect of a compulsory sickness insurance on worker absenteeism", *Journal of Human Resources*, forthcoming.
- Skogman Thoursie P. 2002, "Reporting Sick: Are Sporting Events Contagious?", *Research Papers in Economics*, No 2002:4, Department of Economics, Stockholm University.
- Shapiro C. and Stiglitz J. 1984, "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review* 74, 433-444.
- Tunali I. and Pritchett J. B. 1997, "Cox Regression with Alternative Concepts of Waiting Time: The New Orleans Yellow Fever Epidemic of 1853", *Journal of Applied Econometrics*, Jan.-Feb. 1997, v. 12, iss. 1, 1-25
- Wooldridge, J. M. 2002, *Econometric analysis of cross section and panel data*, The MIT Press, Cambridge, MA.
- Yaniv G. 1991, "Absenteeism and the Risk of Involuntary Unemployment: A Dynamic Analysis", *Journal of Socio-Economics* 1991, Vol. 20, Issue 4, 359-372.

**Appendix, Table A1.** Estimation results and descriptive statistics of the data.

	Estimation results		Descriptive statistics (means)		
	Coefficients	Relative risk	All obs.	Obs. with an unempl. period	Obs. empl. all years of study
Number short sick leave spells (1-7 days)	0.05926*** (0.00608)	1.061	1.928	2.353	1.860
Number medium sick leave spells (8-28 days)	0.15030*** (0.03016)	1.162	0.217	0.312	0.201
Number long sick leave spells (>28 days)	0.25353*** (0.04741)	1.289	0.080	0.111	0.075
Mean extra days in short sick leave spell	0.06786*** (0.00857)	1.070	1.677	2.033	1.620
Mean extra days in medium sick leave spell	0.01064*** (0.00324)	1.011	1.996	2.694	1.884
Mean extra days in long sick leave spell	0.00067*** (0.00012)	1.001	10.674	17.077	9.646
Fraction sick benefits of total income, 86-88	0.29783 (0.18973)	1.347	0.030	0.040	0.028
Woman	-0.34061*** (0.03474)	0.711	0.333	0.250	0.346
Age	0.02230 (0.01650)	1.023	40.346	38.170	40.695
Age <sup>2</sup>	-0.00031 (0.00021)	1.000	-	-	-
Income (t) (prev. year, 100 000 SEK)	-0.07883*** (0.02756)	0.924	2.151	1.968	2.181
Income <sup>2</sup> (t) (prev. year, 100 000 SEK)	0.00033** (0.00014)	1.000	-	-	-
Household income (t) (prev. year, 100 000 SEK)	-0.08565*** (0.01721)	0.918	3.206	2.741	3.280
Household income <sup>2</sup> (t) (prev. year, 100 000 SEK)	0.00013*** (0.00004)	1.000	-	-	-
Number of kids (t)	0.00728 (0.01574)	1.007	0.672	0.618	0.680
Married/living together (t)	-0.15927*** (0.03753)	0.853	0.586	0.477	0.603
Immigrant: Nordic	0.07180 (0.05870)	1.074	0.172	0.172	0.172
Immigrant: Other EU	-0.16628 (0.23672)	0.847	0.040	0.045	0.039
Immigrant: Non-EU	0.12379** (0.05886)	1.132	0.003	0.003	0.003
Local job destruction rate (t)	1.24697*** (0.38307)	3.480	0.038	0.047	0.037
<b>Educational level:</b>					
Pre high-school education shorter than 9 years.	0.30199*** (0.06975)	1.353	0.108	0.129	0.105
Pre high-school education, 9 years.	0.24580*** (0.06533)	1.279	0.114	0.152	0.108
High-school, 2 years.	0.14320*** (0.04043)	1.154	0.318	0.416	0.302
High-school, 3-4 years.	Reference		0.136	0.126	0.137
Undergraduate education, shorter than 3 years.	-0.20407*** (0.05466)	0.815	0.146	0.100	0.154
Undergraduate education, at least 3 years.	-0.46037*** (0.06313)	0.631	0.167	0.070	0.182
Post-graduate education.	-0.86166*** (0.28162)	0.422	0.009	0.002	0.010
Unspecified level of education.	0.16288 (0.22445)	1.177	0.002	0.004	0.002

Table A1 continued	Estimation results		Descriptive statistics (means)		
	Coefficients	Relative risk	All obs.	Obs. with an unempl. period	Obs. empl. all years of study
<b>Type of education:</b>					
Common education, non-vocational.	Reference		0.269	0.331	0.259
Aesthetic, arts and religious (non-teaching).	0.27562** (0.12368)	1.317	0.015	0.013	0.016
Pedagogic.	-0.07222 (0.09676)	0.930	0.074	0.032	0.081
Administrative, behavioural and accountant.	0.02664 (0.06393)	1.027	0.182	0.148	0.188
Industrial, handcraft and natural science.	0.15364** (0.06044)	1.166	0.283	0.355	0.271
Transportation and communication.	-0.10567 (0.09660)	0.900	0.023	0.024	0.023
Health care.	-0.38806*** (0.08533)	0.678	0.090	0.043	0.098
Farming, forestry and fishing.	0.00931 (0.11058)	1.009	0.014	0.016	0.014
Service, security and military.	-0.19610** (0.09564)	0.822	0.038	0.024	0.040
Other	0.18695 (0.13380)	1.206	0.010	0.014	0.010
<b>Type of employer (t):</b>					
Individual.	0.67362*** (0.13348)	1.961	0.003	0.008	0.002
Privately owned company.	-0.62198** (0.29047)	0.537	0.003	0.001	0.003
Trading company.	0.50760*** (0.12275)	1.661	0.005	0.009	0.004
Joint-stock company.	Reference		0.570	0.713	0.547
Profit association (organization).	-0.24000** (0.09535)	0.787	0.019	0.020	0.019
Non-profit association (organization).	0.57278*** (0.08657)	1.773	0.016	0.023	0.015
Foundation.	0.21682 (0.13700)	1.242	0.007	0.007	0.007
Public/Government.	-0.32337*** (0.04320)	0.724	0.363	0.201	0.389
Other.	0.51065*** (0.08394)	1.666	0.013	0.017	0.012
<b>Branch employed in (t):</b>					
Farming, forestry and fishing	-0.11933 (0.09751)	0.888	0.009	0.013	0.008
Mining	-0.22521*** (0.05668)	0.798	0.004	0.006	0.004
Manufacturing	Reference		0.238	0.233	0.239
Energy	0.79521*** (0.05232)	2.215	0.009	0.003	0.010
Construction	0.51739*** (0.04585)	1.678	0.069	0.184	0.050
Retail trade, restaurant and hotel.	0.16464*** (0.04873)	1.179	0.105	0.136	0.101
Communication and postal services.	0.38535*** (0.04922)	1.470	0.091	0.085	0.092
Banking, insurance and distribution.	0.04832 (0.05652)	1.050	0.119	0.113	0.120
Public and private service.	0.15331** (0.06083)	1.166	0.349	0.214	0.371



Table A1 continued	Estimation results		Descriptive statistics (means)		
	Coefficients	Relative risk	All obs.	Obs. with an unempl. period	Obs. empl. all years of study
<b>Regional dummies (t):</b>					
County 1	Reference		0.230	0.204	0.234
County 3	0.29807*** (0.06752)	1.347	0.034	0.039	0.034
County 4	0.13313* (0.07376)	1.142	0.030	0.030	0.030
County 5	0.13717** (0.06282)	1.147	0.047	0.047	0.048
County 6	0.16463** (0.07281)	1.179	0.033	0.033	0.033
County 7	0.05592 (0.09171)	1.058	0.020	0.019	0.020
County 8	0.32988*** (0.07686)	1.391	0.023	0.028	0.023
County 9	0.03154 (0.17780)	1.032	0.005	0.005	0.005
County 10	0.15985* (0.09561)	1.173	0.017	0.017	0.017
County 12	0.17418*** (0.04521)	1.190	0.116	0.119	0.116
County 13	0.12812* (0.07757)	1.137	0.029	0.028	0.029
County 14	0.08647** (0.04187)	1.090	0.164	0.156	0.165
County 17	0.20944*** (0.07276)	1.233	0.030	0.032	0.029
County 18	0.28501*** (0.07141)	1.330	0.031	0.034	0.031
County 19	0.25834*** (0.07379)	1.295	0.029	0.031	0.029
County 20	0.29426*** (0.07042)	1.342	0.029	0.034	0.028
County 21	0.27587*** (0.06843)	1.318	0.032	0.037	0.031
County 22	0.16169** (0.07736)	1.176	0.029	0.028	0.029
County 23	0.27873*** (0.10234)	1.321	0.013	0.015	0.013
County 24	0.24002*** (0.07506)	1.271	0.028	0.031	0.028
County 25	0.32647*** (0.07042)	1.386	0.030	0.035	0.029

Standard errors in parenthesis.

\*, \*\*, \*\*\* - significant at 10, 5, 1 percent level.

(t) – Time varying covariates, the value from 1992 is presented in the descriptive statistics.