LUND UNIVERSITY

An Analysis of Sick Leave in Sweden using Panel Data 1985-1997

Wells, Curt; Edgerton, David; Kruse, Agneta

2004

Document Version:
Other version

Link to publication

Citation for published version (APA):
Wells, C., Edgerton, D., & Kruse, A. (2004). *An Analysis of Sick Leave in Sweden using Panel Data 1985-1997*. (Working Papers, Department of Economics, Lund University; No. 3). Department of Economics, Lund University.

Total number of authors:
3

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Lund University, Department of Economics WP 2004:3
January 27th 2004

An Analysis of Sick Leave in Sweden using Panel Data 1985-1997*

by

Curt Wells, David L. Edgerton and Agneta Kruse

Key Words: Sick insurance, sick leave compensation, panel data, fixed effects negative binomial

JEL Code: C23; C25; J22

Abstract:
Since the beginning of the 1990's, sick leave benefits accruing from the tax-financed Swedish social security insurance have been reduced a number of times. In this paper we use data from the survey on the Household Market and Non-market activities to estimate a fixed effects negative binomial model. We find a direct relationship between the rate of compensation and the amount of sick leave taken.

* The authors wish to thank the Swedish Council for Social Research for financial support (grant F0076/1998). An earlier version of this paper was presented at The Econometric Society World Conference, Seattle, 2000. Communications can be sent to the authors by e-mail to curt.wells@nek.lu.se, david.edgerton@nek.lu.se or agneta.kruse@nek.lu.se

1. Introduction

In this paper we address the question of how individual behavior changes when the rules governing the social welfare system are altered. To this end, the Swedish sickness insurance provides an excellent study object, since the system has often been changed during the past few years. As the question of employee compensation for sick leave is one the more widely discussed aspects of Swedish social welfare legislation, it is of interest to examine how individuals respond to changes in this legislation, and a number of studies have appeared in the literature.

Johansson and Palme (2002) examine changes in the transitions between work and work absence that follow a reduction in the rate of compensation in 1991. They find that increasing the cost for taking sick leave decreases both the incidence and the duration of sick absenteeism. Voss and Diderichsen (2001) use data from a specific company to study changes in the incidence of sick absenteeism following the introduction of a qualifying day in 1993. Their findings are inconclusive since most effects are not significant.

In this paper we use data from a panel, consisting of four waves between 1985 and 1997, to investigate how individuals respond to changes in sickness benefits. In particular we want to see whether differences in socioeconomic and demographic characteristics as well as the work environment are important explanatory factors. Our enquiries lead us to conclude that there is a definite and inverse relationship between the cost of illness to an individual, in terms of foregone income, and the number of sick weeks the individual takes. Unsurprisingly, strenuous work is found to increase the demand for sick leave; however, this seems only to be the case for those living in and around urban areas.

The paper is organized in the following manner. The Swedish sickness benefit system is briefly described in the next section. The basic theoretical model is presented in Section 3, and the data is presented and described in the fourth section. The econometric models are presented in Section 5, while the final two sections present our results and our conclusions.

2. Sickness Insurance in Sweden

Sickness benefits were introduced in Sweden in 1955 as a tax financed, uniform and compulsory insurance, with compensation based of the ‘loss of income’ principle. From its inception until 1991, the benefits were steadily increased. When started, there was a qualification time of three days but these were gradually reduced and abolished in 1987.[1] From December 1987 until March 1991, almost everybody received a sick pay equal to 100% of his or her wage.

As a result of the recession that began in the early 1990's, the government successively reduced a number of welfare benefits, including the sickness insurance. After 1991, the remuneration rate was decreased a number of times. In January 1992, the compensation rate was reduced to 75% of income for the first three sick days. This rate rose to 90% on the fourth and following days. In addition, the first two weeks of compensation was to be paid by the employer rather than the Social Security. In April 1993 the first day became a qualifying day with no compensation, and the compensation for longer sickness periods was reduced. Finally, in 1997 the two weeks of compensation that were the responsibility of the employer were extended to four.[2] Together these reforms have lead to quite a dramatic change: what was once a collective insurance has now become an insurance where the burden has partially been shifted to the individual.

Obviously sickness insurance is important to individuals, and changes in the rules governing the benefits will affect their behavior. The changes in the 1990's have made absence from work more expensive, and we would therefore expect to find a decrease in the time an individual is on sick leave. The changes in 1992 should be observable in the 1992 wave; those from 1993 in the 1995 wave and from 1997 in the final wave of the survey data used in this study.

3. The Economic Model

Following, for example, Johansson and Palme (1996) we can analyze sick leave in the context of the usual neo-Classical model. This model posits a utility function which depends positively upon leisure time, consumption of the composite good, and individual characteristics such as marital status, number of children, education, working conditions, *etc*. Leisure time is "purchased" by abstaining from working time, which is priced at the going wage rate. When utility is maximized, assuming an interior solution, the marginal rate of substitution between

---

[1] To be exact we note that these rules were applied to the blue collar workers only, since white collar workers had a wage contract stipulating no qualification time at all.

[2] There have also been a number of changes in the sickness insurance after 1997. We do not report these changes, since our investigation period ends in 1997. Some of the changes have increased the 'generosity' of the system; and it can be noted that since 1997 the number of persons on sick leave has increased dramatically.

leisure time and working time will be equal to the real wage rate net of tax. This is of course the familiar solution from a basic course in microeconomics.

However, the real world differs from this in a number of ways. First of all, the individual may be unable to work because of illness. Because of the existing social security system in Sweden, the individual is insured against income loss resulting from absence due to illness. This insurance system requires that the restrictions placed on the utility function must be modified. Note that working time and sick time cannot be chosen independently. One has a certain amount of contracted time and sick time must be deducted from this given amount. Thus the individual chooses sick time to maximize his utility and working time becomes a residual. The second adjustment concerns the budget restriction. The individual is compensated for income loss with a percentage of income, δ.[3] Thus sick time becomes a poorer paid substitute for working time.

Given a specified utility function, the demand for sick leave can be obtained as a function of the individual's disposable income, the proportion of income paid as illness compensation and various individual characteristics.[4] Without knowledge of the utility function we can, of course, use a linear approximation to this demand function, see for example Johansson and Palme (2002). We expect that an increase in disposable income will reduce the demand for sick leave. Further, increasing δ, the proportion of the individual's wage paid as illness compensation should increase the demand for sick leave. How individual characteristics affect demand is difficult to say *a priori*.

Directly estimating a linear approximation to the demand function can lead to negative predictions of absences for sickness. To avoid this it is common practice to linearize the logarithm of the demand function. The count data models we use in this study are thus of the form $\ln y = \boldsymbol{X\beta}$.

---

[3]This is not quite true: one is compensated with δ x 100 percent of income up to 7.5 inflation adjusted "base amounts" set by the government each year. However, for those with income above this amount, there is an additional insurance paid by the employer. As a first approximation we therefore assume that the individual receives the same compensation regardless of income.

[4]For example, Hausman (1980) has derived a utility function that will exactly yield a linear demand function. Johansson and Palme (1996) have used such an approach.

4. HUS Data

The data used in this study have been taken from a panel survey of Swedish households. The first wave of the Swedish *Household Market and Nonmarket Activities* (HUS) was completed in 1984, following a pilot study reported in Klevmarken (1984). In all, 2131 households and a total of 3757 individuals were selected.[5] The data used here include the 1986, 1992, 1995 and 1997 waves. The original net sample consisted of 1993 households with 3552 individuals. The database used in this paper consists of 14,390 observations. Those above 64 and below 16 years of age and those who have not provided information on sick leave are dropped from the study.[6] The 1984 wave has also been excluded since questions about the respondent's working environment were not asked.

The available data includes information on sick leave - the number of weeks absent from work with paid sick leave - which is our dependent variable. From our discussion of the economic model, we expect income, socioeconomic variables and information on the work environment to be important explanatory variables. Although the entire sample that we have at our disposal is in excess of 13,600 observations, only some 10,500 observations have data on both demographic data and labor market participation (Sample A). Of these only 8638 also have provided information on income (Sample B), and only 7209 also have information on the work environment (Sample C). The sample we used for estimation contains only those individuals who provide all this information.

The dependent variable in the estimations in the next section is presented in Tables 1 - 3. While averages cannot present the whole picture, they can give us a feeling for the data. Therefore we present some statistics for the different samples in Table 2 and 3. While the average number of sick days differs between the samples, the trend toward lower sick absence is noted in all three. The exception is the slight increase for men in the 1995 wave compared to the 1992 one. We also note that the proportion of both men and women who have taken sick leave has fallen during the three samples. However, the trend is for longer sick leave periods with the exception of the 1997 wave where the number of sick weeks is reduced from

---

[5] The sample procedure is described in Klevmarken (1984). More detail is found in the first volume of Klevmarken and Olovsson (1993)

[6] There were 62 individuals who were absent an entire year: these were not included in the sample. Also, some of the questions asked were not answered by all those interviewed. As variables describing an individual's working environment proved important, the sample includes only those who provided an answer to these questions. In all we use 7209 observations on 4123 individuals.

the high level of 1995. This decrease may be coupled to the extension of the employer's responsibility for the of sick pay from the 2 week period fixed in 1992 to four weeks.

There is a problem with the dependent variable. The respondents were asked if they had taken sick leave during the previous year. If the answer was positive, then they were asked "how many weeks" were you absent from your job. They were also asked to round off their answer to the nearest number of weeks, which means that if they were absent one or two days they were to answer "zero weeks". Thus some of those in the count who are registered as not having sick leave have in reality been absent up to a couple of days. There are 504 in the entire sample cases where a 'zero' answer is actually a rounded down answer. This is about 15.7% of those in the entire sample who have answered 'zero'.

Table 4 presents the distribution of age cohorts in the entire sample of 7209 observations. When one considers those in the sample who took sick leave, we find differences between ages and sexes. Basic statistics for the variables we use are presented in Table 5.

5. The Econometric Methodology

The dependent variable in this study is discrete; thus from the beginning we chose a method that would exploit this aspect of the data. The count model used assumes that the spell is the number of sick weeks per year.[7] The large number of zeros suggests a negative binomial model; the heterogeneity of the data suggests a panel data model; and the suspected relationship between the individual effect and the independent variables suggests a fixed effects model. This is the model fitted here.

In the cross–sectional Poisson regression model, exogenous variables are entered into the model by defining the Poisson parameter λ by $\ln\lambda = \boldsymbol{X\beta}$. In a panel data setting usual specification is $\tilde{\lambda}_{it} = \exp(\boldsymbol{X}_{it}\boldsymbol{\beta} + \mu_i)$ where the individual effects are $\alpha_i = \exp(\mu_i)$. In the fixed effects model presented here, the density is made conditional on the sum of the counts for each individual. Thus the individual effects are conditioned out and the likelihood is[8]

---

[7] The probability function assumes that the count may take any integer value; the data here take on values up to 52. Thus the function should be censored at 52. However, the probability for the count to be greater than 52 is so small that this finess has been ignored.

[8] The derivation is detailed in Hausman *et al* (1984), see also Lee (2002).

$$L_i = \frac{\left(\sum_t y_{it}\right)!}{\prod_t y_{it}!} \cdot \prod_t \left(\frac{\exp(\boldsymbol{X}_{it}\boldsymbol{\beta})}{\sum_s \exp(\boldsymbol{X}_{is}\boldsymbol{\beta})} \right)^{y_{it}} \tag{1}$$

A disadvantage with the Poisson formulation is that the problem of overdispersion: the expectation and the variance of the count are seldom equal in empirical work. One way out of this dilemma is to use the heteroskedastic–consistent covariance matrix, $H^{-1}(G'G)H^{-1}$ where H is the second derivative matrix of the likelihood function and G is the score matrix, see Cameron and Trivedi (1998). A second way out is to use a fixed effects Negative binomial model. Here it is common to use the formulation presented in Hausman *et al* (1983)

$$L_i = \left(\prod_t \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)} \right) \left(\frac{\Gamma\left(\sum_t \lambda_{it}\right)\Gamma\left(\sum_t y_{it}+1\right)}{\Gamma\left(\sum_t \lambda_{it} + \sum_t y_{it}\right)} \right), \tag{2}$$

where Γ denotes the gamma function. Maximization is straightforward.[9] The basic assumption is that individual heterogeneity is modeled by a gamma function with the parameters $\left(\exp(\boldsymbol{X}_{it}\boldsymbol{\beta}, \phi_i / \exp(\mu_i)\right)$. Thus the expectation of y_{it} is $\exp(\boldsymbol{X}_{it}\boldsymbol{\beta} + \mu_i)/\phi_i$ and the variance is $E(y_{it})[1+\exp(\mu_i/\phi_i)]$, Cameron and Trivedi (1998). Here, as above, μ_i are the individual effects and ϕ_i is the individual specific overdispersion term. It should be obvious that only the ratio μ_i/ϕ_i may be estimated.

In linear models, fixed effects estimation is performed by conditioning on the mean of the variables. This conditioning obtains by using variables with their respective mean subtracted (the individual effects may be recovered after estimation). The differencing removes all individually specific variables whose value does not change during the sample. Such variables are sex, education and perhaps occupation. The fixed effects negative binomial model in (2), however, is conditioned on the individual heterogeneity rather than the individual mean. It is thus this heterogeneity that is conditioned out of the model. One implication of this is that, contrary to linear fixed effects models, the constant and individual specific terms may now be estimated; see Allison and Waterman (2002).

There is some discussion in the literature about the model in (2). For example Greene (2001) has proposed a solution where the fixed effects are estimated directly, thereby elimi-

[9] Curt Wells has programmed this function along with its first and second derivatives in Gauss. These routines are available on request.

nating the need for conditioning. This introduces a small sample bias, but Greene suggests that this is "probably not too severe" when the sample has at least 8 observations per unit. Here, however, we have at most four observations. We have also found that this estimation procedure is very sensitive to specification, and have therefore not used it in this study.

There is a further complication when using the Fixed Effects Negative Binomial model. The heterogeneity has been modeled using a gamma function with parameters (λ_{it}, ω_i). The second of these two is the individual specific heterogeneity, which is integrated out before estimation. Now, the expected value of the gamma is $\lambda_{it}\omega_i$ which implies that the marginal effects will contain the unestimated heterogeneity terms. However, as $\lambda_{it} = \exp(\boldsymbol{X}_{it}\boldsymbol{\beta})$, we can obtain 'marginal effects' as follows.

For a continuous independent variable, x_s, we can differentiate the expectation of y with respect to this variable. Taking logarithms on both sides of

$$\mathrm{E}(y_{it}) = \exp(\boldsymbol{X}_{it}\boldsymbol{\beta})\omega_i \tag{3}$$

and differentiating with respect to x_s, we obtain

$$\frac{\partial \ln \mathrm{E}(y_{it})}{\partial x_s} = \beta_s . \tag{4}$$

The coefficient can thus be interpreted as the proportional effect on the dependent variable of a marginal change in the independent variable.

Things become a bit more complicated when the independent variable is binary. The usual method of calculating the marginal effects, by taking the difference between the expected values when the binary is one respectively zero, will not work as the individual heterogeneity remains. However, the following is a useful result. Let the k^{th} independent variable be binary taking the values of one or zero. Then, letting $\tilde{\boldsymbol{X}}$ stand for the remaining $k-1$ independents and $\tilde{\boldsymbol{\beta}}$ the conformal coefficient vector, we may write, for $\theta = 0,1$

$$\mathrm{E}(y_{it} \mid x_k = \theta) = \exp(\tilde{\boldsymbol{X}}_{it}\tilde{\boldsymbol{\beta}} + \beta_k\theta)\omega_i \tag{5}$$

The ratio of the expectation given that the binary is unity to that when the binary is zero will therefore give us an idea as to how the binary variable effects the dependent variable:

$$\frac{\mathrm{E}(y_{it} \mid x_k = 1)}{\mathrm{E}(y_{it} \mid x_k = 0)} = \exp(\beta_k) \tag{6}$$

6. The Empirical Results

6.1. A Fixed Effects Negative Binomial Model

Our discussion in section three suggests that both disposable income and compensation for sick leave will have an effect on the individuals demand for sick leave. These two variables are thus included in our final model. The first of these is the net hourly wage (defined as after tax income divided by hours worked). The second is illness compensation calculated according to the rules for such compensation. We expect these two to have opposite signs, negative for income and positive for compensation. Further the annual dummy variables are expected to pick up effects of the changes in social security legislation. These variables should have a negative sign in the regressions.

Other variables are less obvious. Those representing working conditions are hard to assign *a priori* signs. While those having a *strenuous* job would be expected to be absent more than others all else being equal, it not as easy to predict the sign for one who has attended a vocational school. Micro data with its observations on thousands of individuals offers the possibility of examining interaction effects. As above, we are uncertain as to how these interactions will affect the demand for sick leave. Many background variables have been used in our preliminary analyses. In Table 6 we present those variables and two-way effects that were found to be significant in our models. In many cases the interaction proved significant while the one-way effects were not.

Selection between the fixed effects, random effects and cross-sectional models was made using Hausman tests. Testing the null of the cross–sectional against the fixed effects model yielded a test statistic of 891.9, which greatly exceeds the critical value at any reasonable significance level. There is obviously individual heterogeneity that the cross–sectional model cannot cope with. We also tested the null of a random effects model against the fixed effects model. The test statistic here is 93.3, and this null is therefore also rejected. The results for the fixed effects model are presented in Table 7.

Using a fixed effects specification has its drawbacks. First of all, those individuals with but one occurrence in the database do not contribute to the estimate. The estimation algorithm likewise excludes individuals whose dependent variable never changes. This implies that those individuals who have never taken sick leave do not contribute to the likelihood function. A random effects model would not suffer from these deficiencies, but to be consistent needs zero correlation between the regressors and the individual heterogeneity. Unfortunately the Hausman tests we performed rejected this hypothesis.

The model presented in Table 7 is fairly parsimonious; although the method used allows us to estimate individual specific characteristics, none turn out to be significant. First of all we notice that the dummy variables for the last three panel years are significant. Each of the waves collected during the 1990´s represents a different change in the rules governing sick leave so at to make sick leave more costly for the individual. From the beginning of 1992, compensation was decreased and employers were made responsible for the first 14 days of sick pay. In April of 1993, a qualifying day was introduced; the effect of this change is observable in the 1995 dummy. In January of 1997 the period of employer responsibility for the sick pay was extended to 4 weeks. This steady increase in costs is observed by the decreasing size of the coefficients on the year dummies. Using equation (5) we find, for example, that the ratio of the expected number of sick weeks in 1997 compared to that in 1985, where full compensation was rule, has been halved as exp(-0.678) = 0.51.

Finally, we can see that the hourly wage (net of taxes) does have an effect on the weeks of sick leave. As we would suspect, those with higher wages are more infrequently absent with sick leave.

However a number of cross–effects also appear in the model. We notice a direct price effect for women (x_7) and for men (x_8).[10] If compensation is decreased then absence for sick leave also decreases. This is perhaps the main result of the study. This result is illustrated in Figure 1, for men, and Figure 2 for women. Here we have plotted the expected number of sick weeks for an average individual against the rate of compensation. The latter has been varied from 0.5 to 1, which of course would represent full compensation. Note the direct relationship between the expectation and the rate of compensation. Note also that the response curve is steeper for women. Changes in the system do seem to have an effect on individual behavior.

[10] Compensation is based on income, up to 7.5 basic units, and is here calculated net of taxes. For the 1992 wave taxable income was multiplied by 0.81, reflecting that weekly compensation was 75% of gross income for the first three days and 90% of gross income for the remaining two days. For the 1995 and 1997 waves the corresponding multiple was 0.63, which reflects the introduction of a qualifying day. The basic units were 21,800 Swedish Kronor (1985), 33,700 (1992), 35,700 (1995) and 36,300 (1997).

Two other cross–effects are presented. First is the effect of having a strenuous job for those living in and around the larger cities (x_6). As would be expected, having such a job increases the number of sick weeks. The other cross–effect is that between capital income and a 2-year education beyond the compulsory elementary schooling, which we interpret as vocational training (x_5). The negative sign should be seen in the same light as the effect of the hourly wage: those with higher incomes are sick less.

6.2 Some Further Reflections

A first question that should be asked is whether our choice of model has had an influence on our results. To answer this question we created a binary variable, which is unity if a person has taken paid sick leave during the year regardless of the length of such absence. This variable was then modeled using a fixed–effects logit. As Table 8 indicates, our conclusions are robust to this change of method.

A second question of importance is whether our results are biased by the exclusion of individuals, either because they did not answer questions on their working environment, or because of our use of panel data methods. While no really satisfactory answer may be given, some information can be obtained by looking at Tables 2 and 3. The basic statistics seem about the same whether one uses samples A, B or C, *i. e.*, there does not seem to be any systematic effect between answering the questions on income or environment and the propensity to take sick leave.

In sample D the use of panel data excludes those individuals where the amount of sick leave per year is the same in all waves. In practice this means removing those who have never taken sick leave, and this is the sample used in the estimations in Tables 7 and 8. The exclusion of these individuals will, of course, raise the proportion of observations showing some sick leave during a year. However, the downward trend over time is just as apparent in sample D as in the larger sample.

Finally, in Table 9 the results are presented for a Fixed Effects Negative Binomial model using sample B, which excludes the environmental variables. The included variables are of course somewhat different, but the main results remain essentially unchanged. The Hausman tests leave no doubt as to the necessity of using a fixed effects model, and the dummy variables for 1992, 1995 and 1997 have the expected negative sign. We note that in this model net compensation for both men and women appears with the expected positive sign.

To summarize, the results of our robustness studies do not seem to give any evidence that supports the presence of serious model or sample selection bias.

7. Conclusion

The basic conclusion of our study is that there is an undeniable relationship between social security benefits and the amount of sick leave taken. This is in line with previous findings, such as those of Palme and Johansson (2002). Decreasing the compensation given to those on sick leave reduces the number of weeks individuals are absent from their job.

Our study also shows that there is a great deal of unobserved individual heterogeneity, which is to be expected since we have no data on health status. It is also clearly evident that this heterogeneity is correlated with important explanatory variables, such as income, and that a fixed effects model is therefore necessary. Studies based on cross sectional data or random effects models should therefore be treated with caution.

In general fixed effects models do not allow individually specific explanatory variables, such as gender, to be included. Although this is not theoretically the case for the negative binomial, since it is the variance that has the individual heterogeneity, we find in practice that these variables are not significant. Those explanatory variables that are included in our model have the expected signs.

References

Allison, P. D. and R. P. Waterman (2002). "Fixed–Effects Negative Binomial Regression Models", *Sociological Methodology*, 32(1), 247–265.

Cameron, A. C. and P. K. Trivide (1998). *Regression Analysis of Count data*, Cambridge University Press, New York.

Greene, W. (2001). "Estimating Econometric Models with Fixed Effects", Unpublished manuscript, Department of Economics, Stern School of Business, New York University, available at http//:www.stern.nyu.edu/_wgreene.

Hausman, J. (1980). "The Effect of Wages, Taxes and Fixed Costs on Women's Labor Force Participation", *Journal of Public Economics*, 14(1), 161–194.

Hausman, J., B. H. Hall and Z. Grilliches (1984). "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", *Econometrica*, 52(4), 909–938.

Johansson, P. and Palme, M. (1996). "Do Economic Incentives Affect Work Absence? Empirical Evidence using Swedish Data", *Journal of Public Economics*, 59(2):195–218.

Johansson, P. and Palme, M. (2002). "Assessing the Effect of Public Policy on Worker Absenteeism", *Journal of Human Resources*, 37(2), 381-409.

Klevmarken, N. A. (1984). "Household Market and Nonmarket Activities: The First Year of a Swedish Panel Study", *Vierteljahrsfefte zur Wirtschaftsforschung*, 35(4), 452–457.

Klevmarken, N. A. and P. Olovsson (1993). *Household market and nonmarket activities: Procedures and codes 1984 -- 1991*. Almqvist & Wiksell International, Stockholm.

Lee, M. (2002). *Panel Data Econometrics, Method–of–Moments and Limited Dependent Variables*. Academic Press, San Deigo, California.

Voss, M., B. Floderus and F. Diderichsen (2001). "Changes in Sick Absenteeism Following the Introduction of a Qualifying Day for Sickness Benefit – Findings from Sweden Post", *Scandinavian Journal of Public Health*, 29, 166–174.

APPENDIX: Tables and Figures

Table 1. Paid Sick Leave; N = 7209

Wave	Sample size	Mean number of weeks sick leave			Proportion who have taken sick leave		
		All	Men	Women	All	Men	Women
All	7209	1.278	1.042	1.550	0.375	0.333	0.418
1985	1673	1.628	1.349	1.918	0.501	0.483	0.520
1992	2260	1.335	1.031	1.647	0.390	0.335	0.447
1995	1534	1.253	1.082	1.435	0.317	0.268	0.371
1997	1742	0.929	0.739	1.149	0.283	0.250	0.321

Table 2. Weeks of Paid Sick Leave in Three Different Samples

Wave	Sample A; 10509 obs			Sample B; 8638 obs			Sample C; 7209 obs		
	Proportion women	Mean		Proportion women	Mean		Proportion women	Mean	
		men	women		men	women		men	women
All	0.500	1.183	1.719	0.502	1.208	1.775	0.483	1.042	1.550
1985	0.504	1.719	2.160	0.510	1.722	2.140	0.490	1.349	1.918
1992	0.512	1.217	2.007	0.512	1.201	2.041	0.493	1.031	1.647
1995	0.490	1.085	1.469	0.503	1.203	1.552	0.484	1.082	1.435
1997	0.498	0.778	1.183	0.478	0.763	1.210	0.463	0.739	1.149

Table 3. Proportion with Some Paid Sick Leave in Four Different Samples

Wave	Sample A; 10509 obs		Sample B; 8638 obs		Sample C; 7209 obs		Sample D*; 3097 obs	
	Men	Women	Men	Women	Men	Women	Men	Women
All	0.319	0.392	0.327	0.402	0.333	0.418	0.549	0.597
1985	0.478	0.495	0.486	0.498	0.483	0.520	0.736	0.680
1992	0.334	0.422	0.330	0.427	0.335	0.447	0.576	0.657
1995	0.250	0.346	0.261	0.359	0.268	0.371	0.480	0.538
1997	0.232	0.303	0.237	0.307	0.250	0.321	0.426	0.488

* Sample D consists of those observations that contribute to the likelihood of the fixed effects negative binomial model

Table 4. The Age Distribution in the Sample; N = 7209.

	1985 wave		1992 wave		1995 wave		1997 wave	
	Total number	Proportion female	Total number	Proportion female	Total number	Proportion female	Total number	Proportion female
16 ≤ AGE ≤ 24	121	0.521	192	0.594	52	0.519	85	0.511
25 ≤ AGE ≤ 34	381	0.478	456	0.467	287	0.453	335	0.499
35 ≤ AGE ≤ 44	556	0.491	619	0.501	415	0.506	441	0.454
45 ≤ AGE ≤ 54	391	0.481	656	0.492	519	0.499	591	0.462
55 ≤ AGE ≤ 64	224	0.482	337	0.460	261	0.444	290	0.417
Total	853	0.490	2260	0.493	1534	0.484	1742	0.463

Table 5. Basic Statistics for the Variables in the Model; N=7209.

	1986 wave		1992 wave		1995 wave		1997 wave	
Variable	Men	Women	Men	Women	Men	Women	Men	Women
Size of wave	853	820	1145	1115	792	742	935	807
Means for the Continuous variables								
Age	41.08	40.59	42.09	41.26	43.94	43.75	43.94	42.75
Capital income*	0.03	0.03	0.05	0.04	0.38	0.43	0.06	0.05
Disposable income*	0.66	0.53	1.29	0.94	1.40	1.02	1.36	0.97
Hours worked	41.46	32.40	41.39	34.22	41.71	35.54	41.76	35.31
Weeks of sick leave	1.35	1.92	1.03	1.65	1.08	1.44	0.74	1.15
Proportions for the Binary variables								
City dweller	0.557	0.585	0.551	0.587	0.566	0.571	0.549	0.569
Customers	0.730	0.733	0.806	0.830	0.828	0.811	0.817	0.800
Hectic job	0.798	0.860	0.869	0.909	0.891	0.902	0.882	0.921
Monotonous job	0.204	0.211	0.202	0.209	0.213	0.205	0.218	0.226
Skilled work	0.775	0.744	0.821	0.827	0.867	0.863	0.862	0.840
Strenuous job	0.467	0.535	0.479	0.543	0.431	0.539	0.464	0.570
Vocational school	0.351	0.399	0.329	0.311	0.265	0.275	0.261	0.286

* In units of 100,000 Swedish Kronor

Table 6. Variable Definitions

Variable	Definition	Interacted with	Used in Table
x_1	1992		7, 8, 9
x_2	1995		7, 8, 9
x_3	1997		7, 8, 9
x_4	Net hourly wage*		7, 8
x_5	Capital income	Vocational schooling	7, 8
x_6	Strenuous work	City dweller	7, 8
x_7	Net sick leave compensation**	Female	7, 8, 9
x_8	Net sick leave compensation**	Male	7, 8, 9

* In units of 100 Swedish Kronor per hour

** In units of 100 Swedish Kronor per week

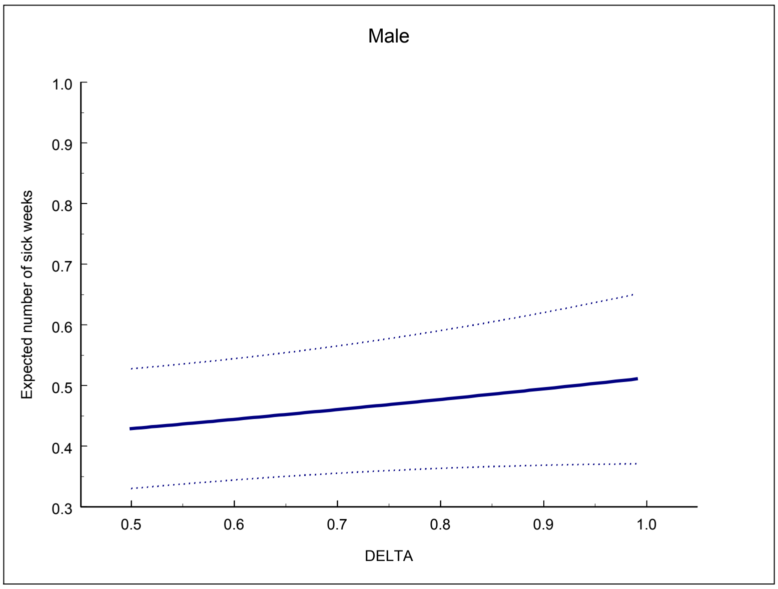
Table 7. Fixed Effects Negative Binomial Model. (Enviromental Variables Included)

Independent Variable	Number of weeks absent from work with paid sick leave.
Number of observations	3097
Number of df	3088
Number of individuals	1126

Hausman test:	Statistic	P-Value
H_0 RE and H_1 FE	93.31	0.0000
H_0 NB and H_1 FE	981.90	0.0000
0.1% critical value of χ^2 with 9 df	27.88	

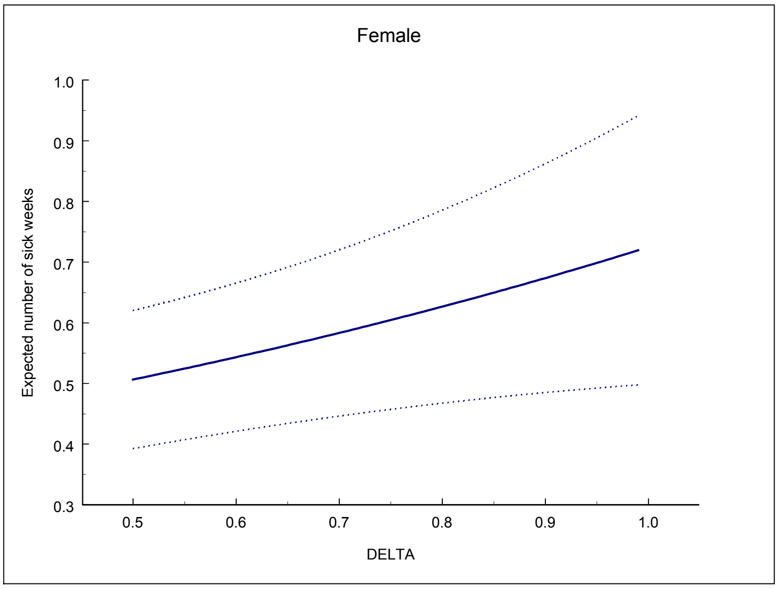
Variable	Coefficient	Standard error	t-statistic	P-Value
Constant	–0.5906	0.1227	–4.812	0.0000
x_1	–0.3710	0.1012	–3.667	0.0002
x_2	–0.4868	00972	–5.010	0.0000
x_3	–0.6778	0.0948	–7.150	0.0000
x_4	–0.8861	0.6039	–1.467	0.1424
x_5	–0.4651	0.1626	2.860	0.0042
x_6	0.2667	0.0943	2.289	0.0047
x_7	0.4420	0.1456	3.036	0.0024
x_8	0.2286	0.1308	1.748	0.0806

Figure 1. Expected Sick Absence as a Function of Compensation Rate (δ). Males



The dotted lines are the 95% confidence interval.

Figure 2. Expected Sick Absence as a Function of Compensation Rate (δ). Females



The dotted lines are the 95% confidence interval.

Table 8. Fixed Effects Logit Model.

Independent Variable	Binary: Individual has paid sick leave during the year
Number of observations	3097
Number of df	3088
Number of individuals	1126

Variable	Coefficient	Standard error	t-statistic	P-Value
x_1	–1.3281	0.2281	–5.823	0.0000
x_2	–1.4386	0.2029	–7.090	0.0000
x_3	–2.0141	0.2061	–9.772	0.0000
x_4	–3.8566	1.4550	–2.650	0.0080
x_5	–1.2455	0.3671	–3.393	0.0007
x_6	0.4441	0.2253	–2.650	0.0487
x_7	1.9896	0.3496	5.596	0.0000
x_8	0.5271	0.3270	1.612	0.1069

Table 9. Fixed Effects Negative Binomial Model. (Enviromental Variables Excluded)

Independent Variable	Number of weeks absent from work with paid sick leave.
Number of observations	3825
Number of df	3819
Number of individuals	1362

Hausman test:	Statistic	P-Value
H_0 RE and H_1 FE	102.50	0.0000
H_0 NB and H_1 FE	1284.67	0.0000
0.1% critical value of χ^2 with 6 df	22.46	

Variable	Coefficient	Standard error	t-statistic	P-Value
Constant	–0.6772	0.0282	–8.181	0.0000
x_1	–0.4064	0.0679	–5.987	0.0000
x_2	–0.6094	0.0703	–8.674	0.0000
x_3	–0.6814	0.0768	–8.872	0.0000
x_7	0.2497	0.0792	3.154	0.0016
x_8	–0.0986	0.0296	3.336	0.0008