A Dynamic Discrete Choice Model of Blue Collar Worker Absenteeism in Sweden 1991[¤]

Claes- M. Cassel[y], Per Johansson[z]and Marten Palme[x]

October 11, 1996

Abstract

The e®ect of economic variables on the probability of being absent is studied using panel data for a sample of 1,056 blue collar workers covering day-to-day data for the time period of one year (1991), in all 365,565 observations. Also, the e®ect of a reform in the sickness insurance on worker absenteeism is studied in some detail. Duration, state dependence and individual heterogeneity are considered. The Slutsky Condition can not be rejected for any individual in the sample and the results show strong economic incentives on worker absenteeism, although the study also shows that the sharp decrease in work absence observed after the reform can not be solely attributed to higher costs of work absence.

[¤]Comments from participants in seminars at Econometrics Conference in Gothenburg, Panel Data Conference in Amsterdam and at the Economics Department, University of Lund are gratefully acknowledged. Financial support for this research is provided by the Swedish Council for Social Research.

[y]Department of Economic Statistics, Stockholm School of Economics, Box 6501, S-113 83 Stockholm, Sweden and Statistics Sweden, S-115 81 Stockholm, Sweden.

[z]Department of Economics, University of Umea, S-901 87 Umea, Sweden.

[x]Department of Economic Statistics, Stockholm School of Economics, Box 6501, S-113 83 Stockholm, Sweden.

1. Introduction

Most industrialized economies have some form of compulsory sickness insurance.[1] Obviously, the aim of such insurance is to replace the foregone earnings if the insured worker has the bad luck of having bad health and is therefore prevented from doing his regular work. However, it is always di±cult, or involvs high costs, to monitor if an insured individual really can work or not, especially for short term sickness spells. The Swedish sickness insurance, which is investigated in the empirical part of this paper, leaves the decision on whether the insured individual is entitled to receive compensation from the sickness insurance for the ¯rst seven days in a spell to the individual herself. The insured individual is entitled to compensation from the insurance if her perception of her health is such that \ it does not permit her to do her regular work\. To prevent abuse of the insurance, most sickness insurance schemes have some form of excess.[2] In the Swedish sickness insurance this excess has the form that it does not replace the full loss in earnings when the insured individual is absent from work. This means that the sickness insurance implicitly works as a incentive contract to encourage the worker not to be absent from work.

There have been few attempts to empirically examine how economic incentives implied by the compensation scheme in the sickness insurance a®ect work absence.[3] There are two institutional conditions in the Swedish sickness insurance that make it particularly useful to examine these issues. These conditions constitute the background to this study. Firstly, the fact that the sickness insurance in Sweden is compulsory means that we are able to get reliable register information on individual daily utilization of the sickness insurance (work absence). These registers can be matched with a large sample survey (The Swedish Level of Living Survey).[4] Secondly, the share of the insured individuals earnings not covered by the insurance (the excess in the insurance) has been changed recently. This gives us better opportunities to study how di®erent designs of the excess in the insurance motivates workers, i.e. to what extent economic incentives a®ect work absence.

To derive the theoretical model for the day-to-day choice of being absent from work, we apply basically the same technique as has been used in numerous labor supply studies. As the choice of work absence is binary for each day, we use a \random utility maximization\ framework. In the empirical part, we use data on work absence for each day for a sample of 1,612 Swedish blue collar workers (816 men and 796

[1] See Kangas (1991) for an overview.

[2] See Lantto (1991) for a theoretical analysis of optimal deterrence to malingering strategies.

[3] A review of most of these studies can be found in Johansson & Palme (1996). Most recently Barmby et al (1995) use data from three plants from the same ¯rm in the U.K. to estimate the e®ect of di®erent sickness insurance schemes using a random e®ects probit model.

[4] The data-set, in addition to the day-to-day information on work absence, also contains extensive individual information on, e.g., income, cost of being absent from work (which is determined by individual income from labor combined with the rules for the sickness insurance), health status, and work environment.

women) during the year 1991. This means that we potentially have 365 records for each individual. Since the interest is in the economic incentives for work absence we only analyze days absent in spells of less than eight days, i.e. for those spells where the choice of being absent is made entirely by the individuals themselves.

The advantage of this study compared to a previous study (Johansson & Palme, 1996) where we used the same theoretical model, is that the rich data set used here enables us to deal with several important issues in the estimation of the model. The panel structure of the data set makes it possible for us to deal with unobserved heterogeneity and to consider the dynamic structure of the model. State and duration dependence is likely to occur in dynamic binary choice data because people tend to persist in their behavior and unobservable variables tend to change slowly (Heckman, 1981a). This is in particular true for work absence behavior. A large part of work absence is caused by di®erent kinds of temporary illnesses, which we, of course, can not observe in the data set. As the health status of an individual on a particular day is not independent of his/her health status the day before, the probability of being absent on a particular day is likely not to be independent of whether or not the individual was absent the day before. The estimation of the parameters in a dynamic probit model (cf. Heckman, 1981a), both with and without ¯xed e®ects, are carried out with maximum likelihood (ML). Since we are using a ¯xed e®ect approach, variables that do not change over the period studied can not be included in the analysis (e.g. working condition, health status etc.). Also individuals that do not leave the states of absence or non-absence can not be included in the estimation.

The time-period of the panel contains a major reform of the Swedish sickness insurance. In March 1, 1991 the compensation levels in the sickness insurance were changed from being 90 per cent of foregone earnings from the ¯rst day to 65 per cent for the ¯rst three days, 80 per cent from day 4 to 90, and 90 per cent for day 91 and thereafter. The main advantage of having this reform within the sample period is that it creates substantial variation in the cost of being absent from work. It also enables us to test for non-linear e®ects of the cost and virtual income variables in the demand for work absence function. Furthermore, the reform enables us to analyze the dynamic structure in more detail. There are two e®ects of this reform. First, the cost of each spell is increased. Second, as the compensation level varies depending on the length of the spell, the cost of being absent one day depends on the length of the spell it is in. The reform, thus, has a direct e®ect through the change in the cost of being absent. It may also have an e®ect beyond that, as the pattern of the spells may change. This may be detected through the dynamic structure of the model.

It is well known that the design of the compensation schemes within the ¯rm may have an impact on work absence (e.g. Barmby et al., 1995). Lazear (1996) found that work absence decreased in an autoglass company when they switched the payment scheme from hourly wages to piece rates. It is also plausible that if a ¯rm applies a relatively steep wage ladder over the workers career, i.e. provides large economic incentives for the workers to stay in the ¯rm, absence may be lower. These issues can, unfortunately, not be considered in this study, as we can not get sufficient information on the compensation scheme applied for the individuals in our data-set. This is of course a limitation. As fixed effects are used to control for unobserved heterogeneity, it is, however, not likely that neglecting these issues will bias our results on the analysis of the immediate effects of the cost and virtual income on work absence. In recent decades (at least on the Swedish labor market for this category of employees) input based compensation, i.e. compensation for some time unit, has became more common. To some extent, this change can be attributed to the fact that it has become increasingly difficult for the employers to observe the output of each employee as the production technology has moved from Taylorism to be more process oriented. An important objective for an input based incentive scheme is to encourage the worker to actually be present at the workplace. This objective of the firm is likely to be highly dependent on the compensation scheme in the sickness insurance, at least for the compensation in the short run. By studying how work absence is affected by the sickness insurance, more general insights on how economic incentives and different compensation schemes affect work absence can be gained.

The study confirms that economic incentives affect work absence and that the compensation level in the sickness insurance affect the insured individuals behavior. The testable implications from the theoretical model, The Slutsky Conditions, can not be rejected for any individual in the sample and the linear functional form seems to be a good approximation of the demand relationship for work absence. However, the results also shows that the whole decrease in the absence rate following the reform in the sickness insurance can not totally be attributed to changes in the economic variables. Furthermore, for the male sub sample, the changes in the relative cost of absence spells of different length, did not account for all changes in the distribution of spells following the reform in the sickness insurance. This seems, however, to be the case for the female sub sample.

The paper is organized into six main sections. Section 2 sets up the theoretical framework. Section 3 gives a description of the institutional settings and data. The econometric models are presented in Section 4. Model evaluation and the results are presented in Section 5. Conclusions with a discussion of further research are given in Section 6.

2. Theoretical Framework

To model the every day choice of whether to be absent from work or not, we have to depart from the assumption that the individual actually prefers to be absent from work if it is entirely without costs. There are, however, several different costs associated with being absent from work. An obvious direct cost, which we consider in our model, is that the sickness insurance does not compensate for all the wage loss. One can also think of other costs. If an employee is absent frequently he may loose his job more easily, not get promoted as easily or increase his wage as fast as an employee being less frequently absent. As this model is static, this type of cost can not be considered.

We start with a utility function,

$$u = U(x; L; s); \qquad (1)$$

where x is a composite good, with the price normalized to one, L is leisure and s is a vector of socioeconomic variables. L can be broken down in to contracted leisure time, t^l; and time absent, t^a. Demand for time absent is obtained when maximizing the utility function (1) subject to the budget constraint

$$x + (1 - \delta) wt^a = wh^* + R \qquad (2)$$

where h^* is the contracted number of working hours, R is income from sources other than labor, w is net wage and δ is the share of the income the worker receives when absent. h^* can be divided into the desired number of working hours and time absent, hence . This gives the identity $T \equiv h + t^a + t^l$; where T and t^l are total available and contracted leisure time, respectively. It is important to emphasize that the contracted leisure time is exogenous in this model. Every day the individuals are assumed to choose between attending work or not, conditional to their perception of their health, the contracted number of work hours and the costs of the alternatives. We thus, assume that the contract specifying the hours of work is made between the worker and the employer in advance. Using the utility function originally proposed by Hausman (1980), the demand function for time absent is

$$t^a = h^* - \alpha w(1 - \delta) - \gamma(R + h^* w\delta) - s = h^* + \alpha c + \gamma y - s. \qquad (3)$$

This is a linear function of the cost for the individual of being absent from work (net earnings not covered by the sickness insurance, i.e. the relative cost between absence and consumption), c, and virtual income when the individual is absent, y.

If the demand for absence equation (3) is to be consistent with a well behaved utility function, The Slutsky Condition must be satisfied. For this functional form it is sufficient that $\alpha < 0$ and $\gamma \geq 0$ for every $t^a \geq 0$.

If δ is zero, the individuals would have no economic incentive whatsoever to attend their work in this model. On the other hand, it is well known that the chances of being promoted will decrease and the risk for the worker to lose his job will increase if he is frequently absent from work. Conversely, if the cost of being absent is increased very much, it is natural to think that some very ill persons would still not be able (or permitted?) to attend their workplace.[5] Almost the same reasoning holds for the virtual income. If a person has a very large non-labor income, the model will give him very small incentives to attend work. On the other hand, the social element of keeping a job probably provide incentives for limited work absence.

[5] Most people would, however, go to work if they had to pay say $1 million for being absent one day!

This discussion indicates that there can be non-linearities, at least for extreme observations on costs and income for the demand relation of work absence. The linear functional form should, thus, be seen as an approximation for a relatively short segment of this relationship. However, our data contains a major reform of the Swedish sickness insurance which created a large shift in the cost of being absent from work. To allow for possible non-linearities the cost, $c_m = D_m c$, and virtual income $y_m = D_m y$ in the period after the reform as well as the dummy variable, D_m, taking value one in the period after the reform are included in equation (3). This gives us

$$t^a = h^* + \alpha c + \gamma y + \alpha_m c_m + \gamma_m y_m + \beta_m D_m - s, \tag{4}$$

where α_m, γ_m and β_m are the parameters that corresponds to c_m, y_m and D_m.

3. Institutional Settings and Data

The data is obtained from the 1991 Swedish Level of Living Survey (SLLS, see Fritzell & Lundberg, 1994). The SLLS is a sample survey and contains detailed information on economic conditions and behavior for a random sample of about 6,000 individuals. Some limitations are made for the present study. First, the sample is restricted to individuals aged 20 to 64 years. Second, individuals not in the labor force are excluded from the sample. We also exclude self-employed, students, military personnel and white collar workers, i.e. the study is restricted to blue collar workers. The reason for excluding these other groups is to limit heterogeneity arising from differences in sickness insurance systems that can not be obtained from the available data. After these exclusions are made, the sample consists of 1,612 individuals. (Tables containing descriptive statistics for variables in the sample as well as for variables used in the estimation are given in the Appendix)

Data for the dependent variable is obtained from the National Social Insurance Board by matching with the SLLS sample. The definition for the dependent variable to indicate work absence, is that the individual is compensated from the compulsory sickness insurance that day and that the day is in a sequence of less than eight days. For days in a sequence of more than seven days, the individual has a certificate from a physician. As the data was collected from registers of actual payments to the insured individual, the quality is likely to be good. However, if we define work absence as time when the employee is absent from work which is not agreed in advance with the employer and statutory leisure time (such as statutory holiday), a small fraction of work absence is likely not to be included in the sickness insurance data.[6]

The insured individual is entitled to compensation also at weekends and holidays. We will add dummy variables for weekends and work-free days (WE), as well as the month of the holiday (HOLLY) for industrial workers (July), to find out if the utilization of the sickness insurance differ for these days compared to regular work

[6] According to one survey the amount was 2.9 per cent in 1986 (SAF, 1986).

days. Also, there are four days in the sample that are in between two work-free days which also are specified with a dummy variable (BW).

Data to construct the cost and income variables, i.e. the different components of the individual's income, are obtained from tax registers that are matched with the SLLS. As was discussed in Section 3, the cost and virtual income variables are influenced by the compensation level and by the income tax. Individual income from the social security system is taxed at the same rate as income from labor. The Swedish income tax system consists of two parts; a proportional tax imposed by the local authorities and a progressive tax imposed by central government. The local government tax is proportional and is imposed on income from labor. The local tax rate varies somewhat between Sweden's 286 communes and has a mean of 30.3 per cent. The state tax is imposed on income from capital, at a rate of 30 per cent, and on taxable income[7] from labor above 170,000 SEK at a rate of 20 per cent.

To calculate the hourly net income from labor there is a possible endogenity problem. As labor income is not fully compensated by the sickness insurance, the marginal tax rate may depend on how many days the individual is absent from work, i.e. it may not be independent of the individual decision whether or not to attend work. However, as the Swedish tax system only has two brackets and only 20 per cent of the individuals in our sample have a potential taxable income from labor above the "kink point" at 170,000 SEK, we have chosen to neglect the problem of possible endogeneity and have used the marginal tax rate each individual actually pays for the years studied.

The cost variable, c, is for labor incomes below the social security[8] ceiling calculated as $c = w(1 - \delta)$, where w is hourly net income from labor and $(1 - \delta)$ is the share of foregone earnings not covered by the sickness insurance. Before March 1 1991 this share was 0.1. After March 1, δ was decreased to 0.65 for the first 3 days absent and to 0.8 for the following 87 days in a spell. For the days after day 90, δ remained at 0.9. As described in the introduction, the reform is likely to affect absenteeism in two different ways. First, the cost of being absent from work will increase and, if the behavior of the workers is sensitive to increases in cost, average absenteeism will decrease. For the entire Swedish population the average number of days on the sickness insurance decreased from 24 days 1990 to 22.5 days 1991.[9] In our sample, the average number of normalized days absent in the period is 34.06 for the period before the reform (January 1 to February 28) and 30.16 after the reform (March 1 to December 31). Since the latter period contains the summer holiday this comparison can be somewhat misleading. For the period October 1 to November 30, a period not containing the summer holiday, the corresponding figure is 31.11, i.e. an decrease

[7] Taxable income from work, is defined as income from work minus a deduction that depends on the income. The deduction is in the range 10,304 - 18,384 SEK.

[8] The social insurance ceiling corresponded to an annual income of 241,500 SEK. All blue collar workers in our sample have income belows this ceiling.

[9] National Social Insurance Board, Facts on the Swedish Social Insurance 1992.

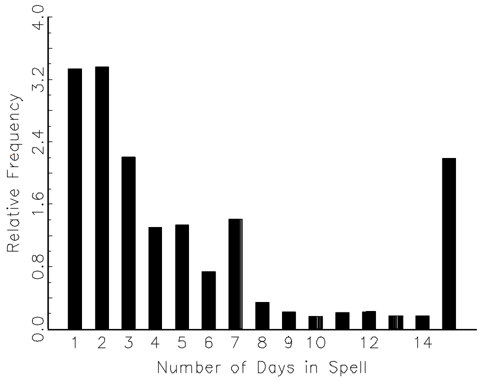

Figure 1: Relative frequencies of work absence spells of di®erent length in the sample. Calculated as the number of spells with the length 1 to 14 days as well as 15 days and more divided by the total number of days in the period for the period January 1 to Febraury 28.

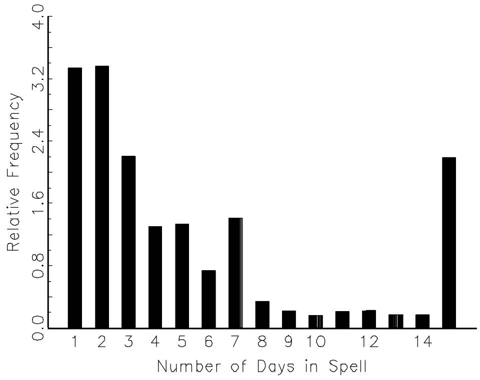

Figure 2: Relative frequencies of work absence spells of di®erent length in the sample. Calculated as the number of spells with the length 1 to 14 days as well as 15 days and more divided by the total number of days in the period for the period March 1 to December 31.

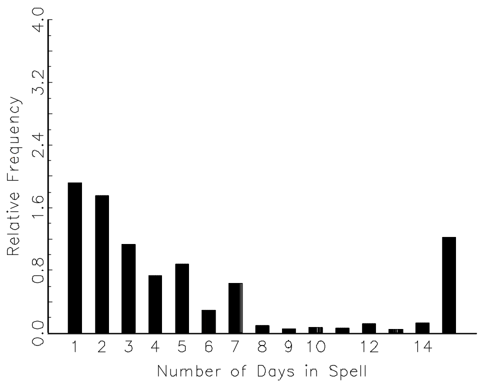

Figure 3: Relative frequencies of work absence spells of different length in the sample. Calculated as the number of spells with the length 1 to 14 days as well as 15 days and more divided by the total number of days in the period for the period October 1 to November 30.

compared to the pre-reform period. Second, the cost of shorter spells increased more compared to longer spells. Figure 1 to 3 shows the number of work absence spells with different length. As the number of days differ between the different periods, we have divided the number of spells with different length with the number of days in each period (e.g. 59 for the period January 1 to February 28) in order to make the periods more comparable. It can be seen, both when comparing 1 with 2 and 3 respectively, that the largest decrease after the reform is within short spells, i.e. in the kind of spells with the largest increase in costs.

The hourly wage rate is obtained by dividing the potential annual labor income by the number of hours of work stated in the surveys. To calculate the potential annual income from labor, we have added the share of income from labor not covered by the sickness insurance, δ for each day recorded as the individual having been compensated by the sickness insurance.

The non-labor income, y, consists of two parts. The first one, R consists of the sum of daily income from capital when not working, child and housing allowances. The second part is net daily income from sickness insurance.

4. Econometric Modeling

Under the assumption that the socio-economic variables, s, is a linear function of explanatory variables and unknown parameters the variables on the right hand side of (4) can be written as $z_{it}\beta$. Here z_{it} is a vector of the explanatory variables, $\beta = (\alpha; \psi; \alpha_m; \psi_m; \gamma')'$ is the corresponding parameter vector, γ is the parameter vector that corresponds to the socio-economic variables, $i = 1, \ldots, N$ and $t = 1, \ldots, T_i$ indexes individuals and time periods, respectively.

As been discussed above is it likely that the absence records are the result of a dynamic process. Structural[10] and duration dependence can be motivated from, e.g., the stimulus-response model developed by psychologists (Heckman, 1981a) and from the dynamics of action theory (cf., Fichman, 1989). We argued above that the change in the compensation level on March 1 may induce different dynamic behavior for the individuals before and after the reform. To model and test this statement the structural and duration dependence is incorporated as

$$\begin{aligned} t^a_{it} &= h^*_i + z_{it}\beta + \lambda d_{i(t-1)} + \lambda_m D_{mit} d_{i(t-1)} + \phi Dur_{it} + \\ &\quad \phi D_{mit} Dur_{it} + \kappa Dur^2_{it} + \kappa_m D_{mit} Dur^2_{it} + \varepsilon_{it}, \end{aligned} \tag{5}$$

where d_{it} is an indicator variable taking value one if individual i is absent day t, ϵ_{it} is an error term, $Dur_{it} = \sum_{s=1}^{t-1} \prod_{k=1}^{s} d_{i,t-k}$ and λ, λ_m, ϕ, ϕ_m, κ and κ_m are the parameters of the process.

The error term is intended to reflect errors unknown to the econometrician. Let $\epsilon_{it} = \theta_i + \omega_{it}$, where θ_i is an individual specific term[11] and ω_{it} are identically and independently distributed (iid) random variables with $E(\omega_{it}) = 0$ and $E(\omega^2_{ih}) = 1$. Assume that an individual will be absent if demand for absence is larger than some threshold value k_i. Thus if $t^a_{it} \geq k_i$ individual i will be absent day t, thus

$$\Pr(t^a_{it} \geq k_i | \theta_i, D_{it}) = \Pr(d_{it} = 1 | \theta_i, D_{it}), \tag{6}$$

where D_{it} is the information set at time period t.

Under the assumption that ω_{it} is normal we get

$$\Pr(d_{it} = 1 | \theta_i, D_{it}) = \Phi(\nu_{it}), \tag{7}$$

where $\Phi(.)$ is the standard normal distribution function and $\nu_{it} = z_{it}\beta + \lambda d_{i(t-1)} + \lambda_m D_{mit} d_{i(t-1)} + \phi Dur_{it} + \phi D_{mit} Dur_{it} + \kappa Dur^2_{it} + \kappa_m D_{mit} Dur^2_{it} + h^*_i + \theta_i - k_i$. Under the assumption that $\alpha_i = h^*_i + \theta_i - k_i = \alpha$ for all individuals the ordinary probit ML estimator can be employed. The log-likelihood function to be maximized is

[10] With structural dependence means that the disposition to be absent today depends in some way on yesterdays decision. Habit persistence states that the disposition to be absent today depends on yesterdays disposition to be absent. These two forms can, however, not be separated empirically (Heckman 1981a).

[11] The fixed effect approach is used for two reasons. First; it is quite likely, from the discussion in section 2, that unobservable attributes to being absent are correlated with the cost, c. Second; a random effect model gives consistent estimates only if the initial state of the process is observed (Heckman 1981b) see however Barmby et al. (1995). See also Guilky & Murphy (1993) for the numerical problems with large T in a non-dynamic random effect models. The incidental parameter problem is considered as a smaller problem since we have quite large T (see Heckman 1981b).

$$\ell = \sum_{i=1}^{N}\sum_{t=2}^{T} d_{it}\Phi(\mu_{it}) + (1 - d_{it})(1 - \Phi(\mu_{it})). \tag{8}$$

The robust covariance estimator

$$\mathrm{Cov}(\hat{\omega}) = H^{-1} J H^{-1} \tag{9}$$

is evaluated at the ML estimates $\hat{\omega} = (\hat{\beta}', \hat{\lambda}, \hat{\lambda}_m, \hat{\phi}_m, \hat{\kappa}, \hat{\kappa}_m)'$ (cf. Davidson & MacKinnon, 1993, Ch. 8) where

$$H = \sum_{i=1}^{N}\sum_{t=2}^{T} \frac{\partial^2 \ell_{it}}{\partial \omega \partial \omega'} \text{ and } J = \sum_{i=1}^{N}\sum_{t=2}^{T} \left[\frac{\partial \ell_{it}}{\partial \omega}\right]' \left[\frac{\partial \ell_{it}}{\partial \omega}\right].$$

For the fixed effect probit estimator equation (8) is maximized for $\omega = (\beta', \lambda, \lambda_m, \phi_m, \kappa, \kappa_m, \alpha')'$ where $\alpha = (\alpha_1, ..., \alpha_N)'$ and (9) is evaluated at the corresponding estimates $\hat{\omega}$.

5. Model Evaluation and Results

The results for the ordinary probit ML estimator and for the fixed effects estimator, for the male and female sub-sample respectively, are presented in Table 1. The t-values are calculated from the consistent matrix in (9). As mentioned in Section 4, in order to identify a fixed effect for each individual, the individuals in the sample have to leave their initial state. 578 in the male sub-sample met that requirement: 223 of the 816 individuals in the sample were not absent at all and 15 were absent all days during the time-period studied. In the female sub-sample the corresponding figures are that 195 (of 796 in the sample) were not absent at all, while 13 were absent all days. Furthermore, in order to restrict the analysis to short term absence spells we have removed absence spells longer than 7 days for which a certificate from a physician is needed to get compensation from the sickness insurance. By making this restrictions, we are able to concentrate on the absence that are left entirely to the individual's decision. Using this criteria 48 males and 62 females are removed from the sample. The resulting sample consists of an unbalanced panel with 530 men and 526 women. For the male sub-sample the total number of records, i.e. day-to-day observations is 183,213 (on average 345.7 days per individual), and for the female sub-sample it is 183,352 (348.6 days on average).

Descriptive statistics of the variables for both the original data-set (OSAMP) and for those actually used in the estimation (ESAMP) are given in the Appendix. The mean number of days absent is, of course, much larger in the OSAMP than for the ESAMP. No large differences are found for the explanatory variables, however. Descriptive statistics of some socio-economic variables not used in the estimation are also given in the tables. For these variables there are some small differences between the OSAMP and the ESAMP. The mean age for both males and females are lower in the ESAMP. This is expected since it is known that longer absence spells are more common for older people. The mean number of children are larger in the ESAMP compared to the OSAMP for the male sub-sample. The opposite is true for the female sub-sample. For both the females and males there are fewer individuals living singly in the OSAMP as compared to the ESAMP. At least to some extent this may follow from the fact that lower age and living singly is positively related in general.

The last row in Table 1 reports the log-likelihood value for each model. It can be seen that the log-likelihood increases substantially with the fixed effect model both for the male and female samples, i.e. the ordinary probit model can, as expected, be forcefully rejected. The within sample prediction ability of the two models are reported in Table 2. The prediction ability is compared with the naive model with only fixed effects. As all individuals with more than fifty per cent work absence rate are removed from the sample the model with only fixed effects predicts no work absence for all individuals every day. Since the overall absence rate in the sample is only 2.6 per cent for men and 2.5 per cent for women this model predicts 97.4 and 97.5 per cent correct for the male and female samples respectively. It can be seen that the number of correct predictions increases for both the models and both samples. The best predictions are obtained with the dynamic fixed effects probit model. With this model the number of correct predictions are increased with 1.35 and 1.29 percentage points for the males and females, respectively.

It is well known (e.g. Heckman, 1981b) that if we have a true non-dynamic model with individual effects and we estimate a model with omitted individual effects but with lagged dependent variables the estimated parameters for lagged dependent variables are likely to be significant, i.e. spurious state dependence. Hence the models may be evaluated by studying the dynamic structure. The dynamic structures of the two models have six different components: the lagged dependent variable, duration dependence and the three variables interacting with D_m. As expected, the parameter estimates of the lagged dependent variable differ substantially between the two models, both for the males and the females.

The cost and income variables are estimated to be significantly below and above zero, respectively for both samples. The parameter estimates of the dummy variable for the time period after March 1 are significantly negative for both samples. The interaction with the cost and income variables are insignificantly different from zero for both samples. Thus, The Slutsky Condition can not be rejected for any man or women in any of the two time periods of the sample. These estimates indicate that economic incentives have a fairly strong and significant impact on work absence. We can also note that the linear functional form serves as a good approximation of the demand relation for being absent from work over this cost and income interval. However, as the dummy variable for the reform is significantly negative, it also shows that the shift in the cost of being absent from work which the reform implied, does not account for the entire decrease in work absence after the reform.

Table 1: Parameter estimates and asymptotic t-values (standard errors are calculated from (9)) for the probit ML estimators.

	Men				Women			
	Ordinary probit		Fixed effects probit		Ordinary probit		Fixed effects probit	
	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t
ONE	-1.887	-28.57			-1.961	-26.25		
c	-0.046	-6.29	-38.252	-5827.71	-0.032	-4.76	-10.316	-1570.96
y	0.002	5.71	2.349	5641.34	0.013	3.61	6.679	1458.39
HOLLY	-0.290	-7.21	-0.339	-6.69	-0.205	-5.45	-0.256	-5.39
WE	-0.535	-22.61	-0.554	-19.76	-0.594	-23.56	-0.607	-21.11
BW	-0.343	-3.36	-0.398	-3.43	-0.189	-2.05	-0.231	-2.25
D_m	-0.208	-2.78	-0.187	-8.28	-0.291	-3.36	-0.272	-9.85
c_m	0.001	0.132	-0.004	-0.46	-0.0002	-0.02	-0.007	-0.81
y_m	-0.000	-0.03	0.0002	0.39	0.004	0.85	0.007	1.25
DUR	-0.035	-0.37	13.753	48.68	0.079	0.84	4.370	8.62
DUR^2	-0.026	-2.03	-4.680	-97.91	-0.036	-2.83	-1.617	-13.02
$D_m \times DUR$	0.289	2.666	-7.759	-26.93	0.135	1.25	0.338	0.65
$D_m \times DUR^2$	-0.034	-2.351	2.652	52.71	-0.019	-1.29	-0.002	-0.02
d_{t-1}	3.048	21.891	-6.131	-19.21	2.916	21.37	0.236	0.53
$D_m \times d_{t-1}$	-0.283	-1.777	5.085	15.92	-0.060	-0.38	-0.311	-0.67
ℓ	-10611.02		-9147.13		-10425.47		-8974.06	

Table 2: The predicted ($\hat{d}$) and observed absence (d) for the ¯xed e®ects and ordinary probit model for the males and female samples (Percentage ¯gures are given in parenthesis).

	Men			Women		
dn$\hat{d}$	1	0	Total	1	0	Total
			Fix e®ects probit			
1	3,137	669	3,806	2,983	626	3,609
	(1.72)	(0.37)	(2.09)	(1.63)	(0.34)	(1.97)
0	1,618	177,276	178,894	1,576	178,167	179,743
	(0.88)	(97.03)	(97.91)	(0.86)	(97.17)	(98.03)
			Ordinary probit			
1	3,076	1,131	4,207	2,912	1,069	3,981
	(1.68)	(0.62)	(2.30)	(1.59)	(0.58)	(2.17)
0	1,679	176,814	178,493	1,647	177,724	179,371
	(0.92)	(96.78)	(97.70)	(0.90)	(96.93)	(97.83)
Total	4,755	177,945	182,700	4,559	178,793	183,352
	(2.60)	(97.40)	(100)	(2.49)	(97.51)	(100)

The results for the interaction between the reform dummy and the dynamic variables di®er between the two sub-samples. For the male sub-sample, both the interaction component with the lagged dependent variable and the two components of the duration dependence are signi¯cantly di®erent from zero. For the female sub-sample, there are no signi¯cant change in the dynamic structure after the reform whatsoever. If the dynamic structure of the model remains the same after the reform, as is the case for the female sub-sample, the result can be interpreted as an inability to reject the hypothesis that the change in the spell structure, noted in Figures 1 to 3, can be attributed to changes in relative costs of di®erent spells after the reform. However, for the male sub-sample, the dynamic structure of the model did change after the reform. The spell structure change more than can be explained by changes in relative costs. In order to facilitate the interpretation of the change in the dynamic structure of the male sub-sample, we have calculated the dynamic dependence for the two periods for the latent variable, ta, for di®erent consecutive lag lengths. These results are shown in Table 3. It can be seen that the positive dependence for the ¯rst two days is larger under the pre-reform regime and is almost zero for day three for both regimes. For day four to day seven there is negative dependence that, in absolute value, is larger before the reform. Thus, some of the change in the spell structure seen in Figures 1 to 3, could not be explained by increases in relative costs of short spells, but are captured in changes in the dynamic structure.

Apart from the coefficients of various parts of the dynamic structure, the signs of the coefficient estimates are the same for the male and female sub-samples. Under the assumption of equal variance, i.e., σ_{ω}^2 , within the two populations the most striking difference between the estimates of the sub-samples is that the coefficient of the cost variable in the male sub-sample is larger in absolute value. Hence, men are more sensitive to costs of being absent in their work absence behavior. This is also found in Johansson & Palme (1996) and Johansson & Brännäs (1996).[12] At first sight this result seems somewhat surprising, given that the standard result in labor supply studies is that women have higher wage elasticity than men (cf. Killingsworth & Heckman, 1986). An explanation to the higher wage elasticity is that home-production is a closer substitute to market work for women. It is hard to see that this explanation can apply also for work absence. One plausible explanation is given in VandenHeuvel & Wooden (1995). They show that women are more sensitive to factors outside the work environment in their absence behavior than men, e.g. commuting time, stressful life events and family responsibilities and as a consequence women are likely to be less sensitive to the cost of being absent.

We have included dummy-variables for three different kind of days when it can be expected that individuals utilize the sickness insurance differently compared to ordinary weekdays. The first variable, WE, is a dummy-variable for weekends, i.e. Saturdays and Sundays. Most people in Sweden do not work on Saturdays and Sundays. During the pre-reform period the insured individual had economic incentives to not receive compensation from the sickness insurance during weekends, as it did not compensate for all foregone earnings, for both short and long sickness spells. Under the post-reform regime there are economic incentives working in different directions, as utilization of the sickness insurance during weekends can qualify to increased compensation level. However, as we in this study only include spells shorter than eight days in the analysis, this latter effect is not expected to be very strong. As can be seen from Table 1, the estimates of the parameters for the weekend dummies are as expected significantly negative.

HOLLY is a dummy variable for the four weeks in the summer which are the most common summer holiday for industrial workers in Sweden. The economic incentives to utilize the sickness insurance during the summer holiday are somewhat different from during weekends. Although they do not get full compensation from the sickness insurance if they utilize this insurance during the holiday they are entitled to have holiday in exchange for the days they utilize the sickness insurance during their regular holiday. Thus, in this case, the individual behavior depends on the individual valuation of leisure time. However, one should keep in mind that there are administrative costs associated with reporting short term sickness during the holiday. Therefore, it is not surprising that the estimates of the parameters are significantly negative.

BW is a dummy variable for days between two work-free holidays. The outcome of

[12] Johansson & Brännäs (1996) estimates a household model for work absence, however it does not provide an explanation of this lower elasticity for women than for men.

Table 3: Dependence on previous states in the ¯xed e®ects model for the male sub-sample before and after the reform.

Lag length	Before the reform	After the reform
1	2.942	2.920
2	8.786	3.876
3	-0.861	-0.270
4	-19.868	-8.472
5	-48.235	-20.730
6	-85.962	-37.044
7	-133.049	-57.414

the estimates of the parameter for this variable are dependent on two e®ects working in opposite directions. Firstly, the valuation of leisure time is probably higher for such days. This e®ect is working in the direction of a positive parameter estimate. Secondly, there are higher administrative costs to utilize the sickness insurance during planned vacations. Also, the sickness insurance gives lower compensation than planned vacations, these e®ects will work in the other direction. As can be seen in Table 1, these latter e®ects seems to dominate in our sample, as the parameter estimates turned out signi¯cantly negative.

6. Discussion

There are several similarities between the individual choice of labor supply contingent on tax and bene¯t schemes imposed by the government and the individual choice of work absence contingent on the rules for the sickness insurance. There is, however, one important di®erence: the latter choice is also contingent on a contract between the employer and the employee, i.e. the insured individual. The excess, the replacement level, in the insurance will, implicitly, work as an economic incentive for the worker to attend his or her workplace but it is a somewhat di®erent choice compared to an individual who is free to choose the number of hours of work, although he or she may be restricted by the number of hours that can be o®ered from the "demand side". There are two fundamental questions which are raised when analyzing work absence empirically. First, can the same empirical models that have successfully been used for some decades to analyze labor supply behavior be used to analyze work absence behavior as well, even though there are some di®erences in the institutional settings of the choice. Second, has economic incentives any considerable e®ect on work absence, i.e. can changes in the replacement level of the sickness insurance a®ect work absence.

On the ¯rst issue, the reform of the sickness insurance on March 1 allowed us to analyze the functional form of the demand relation for work absence, as it was a large shift in both the cost and income variables. The results show that the linear functional form can not be rejected. The Slutsky Conditions can not be rejected from the estimates for any individual in the sample. However, the reform also revealed two problems with the model. First, the parameter estimates for the dummy-variable are significantly negative for both samples. An explanation of this result can be that we failed to consider something that had a decreasing influence on work absence and took place at the same time as the reform. It can also be that the individuals "over-reacted" initially to the reform. This is seen in other areas as a reaction to economic reforms. It can also be a result of an inadequate functional form. Secondly, for the male sub-sample the dynamic structure changed significantly after the reform. The explanation to that can be on the same lines as for the negative estimate of the reform dummy-variable. However, it can also be that a rational individual not only considers the replacement level the actual day, but also the next day and the day after, when determining the absence behavior. This indicates that it may be fruitful to depart from a dynamic theoretical model even when analyzing short term work absence, which is an important subject for further research in this area.

On the second issue, it confirms what is found in a number of previous studies (e.g. Barmby et al, 1995, Johansson & Brännäs, 1996, and partially also in Johansson & Palme, 1996) that economic incentives have a significant impact on the work absence behavior. Furthermore, the results confirm that men seem to be more sensitive to economic incentives compared to women, which also has been found in previous studies.

References

Barmby, T. A., Orme, C. D. & Treble, J. G. (1995). Worker Absence Histories. Labor Economics 2, 53-66.

Davidson, R. & MacKinnon, J. G. (1993). Estimation and Inference in Econometrics. Oxford: Oxford University Press.

Fichman, M. (1989). Attendance Makes the Heart Grow Fonder: A Hazard Rate Approach to Modeling Attendance. Journal of Applied Psychology 74, 325-335.

Fritzell, J. & Lundberg, O. (1994). Vardagens Villkor. LevnadsfÄrh llanden i Sverige under tre decennier. Stockholm: Brombergs FÄrlag.

Guilky, D. K. & Murphy, J. L. (1993). Estimation and Testing in The Random E®ect Probit Model. Journal of Econometrics 59, 301-317.

Heckman, J. J. (1981a). Statistical models for discrete panel data. In Structural Analysis of Discrete Data (C. Manski, and D. McFadden, eds.) 114-178. Cambridge, Mass.: MIT Press.

Heckman, J. J. (1981b). Estimating a Discrete Time-discrete Stochastic Process. In Structural Analysis of Discrete Data (C. Manski, and D. McFadden, eds.) 179-195. Cambridge, Mass: MIT Press.

Hausman, J. A. (1980). The E®ect on Wages, Taxes, and Fixed Costs on Women's Labor Force Participation. Journal of Public Economics 14, 161-194.

Johansson, P. & BrÄnnÄs K. (1996). A Household Model For Work Absence. Umea Economic Studies 400, Umea University.

Johansson, P. & Palme, M. (1996). Do Economic Incentives A®ect Worker Absence? Empirical Evidence Using Swedish Data. Journal of Public Economics 59, 195-218.

Kangas, O. (1991). The Politics of Social Rights. Studies in the Dimensions of Sickness Insurance in the OECD Countries, Ph.D. thesis at the Swedish Institute for Social Research.

Killingsworth, M. R. & Heckman, J. J. (1986). Female Labor Supply: A Survey. In Handbook of Labor Economics (O. Ashenfelter & R. Layard, eds.) 103-204. Amsterdam: Elsevier Science Publisher.

Lantto, K. (1991). Optimal Deterrents to Malingering, Ph.D. thesis. Department of Economics. Stockholm University.

Lazear, E. P. (1996). Performance Pay and Productivity, mimeo Hoover Institution and Graduate School of Buisiness, Stanford University.

National Social Insurance Board (1992). Facts on the Swedish Social Insurance 1992. Stockholm: RFV.

SAF (1986). TidsanvÄndningsstatistik. Svenska ArbetsgivarefÄreningen, Stockholm.

VandenHeuvel, A. & Wooden, M. (1995). Do Explanaitions of Absenteeism Di®er for Men and Women? Human Relations 48, 1309-1329.

Table 1: Description of the two data sets for the men.

	OSAMP (N = 816)				ESAMP (N = 530)			
Variable	Mean	StdDev	Min	Max	Mean	StdDev	Min	Max
d	0.084	0.278	0	1	0.026	0.159	0	1
c	7.454	2.199	0.632	22.307	7.521	2.352	0.632	22.307
y	2.729	4.196	0	55.814	2.513	4.471	0	55.814
HOLLY	0.082	0.275	0	1	0.082	0.275	0	1
WE	0.315	0.464	0	1	0.315	0.465	0	1
BW	0.011	0.104	0	1	0.011	0.104	0	1
Married	0.668	0.471	0	1	0.668	0.471	0	1
Single	0.283	0.450	0	1	0.290	0.454	0	1
Age	38.400	11.988	20	64	37.040	12.072	20	64
Number of children	0.641	1.034	0	9	0.616	1.070	0	9
D_m	0.838	0.368	0	1	0.837	0.369	0	1
c_m	6.257	3.406	0	22.307	6.301	3.515	0	22.307
y_m	2.258	3.929	0	55.814	2.079	4.167	0	55.814
DUR	6.510	34.768	0	364	0.043	0.375	0	6
$D_m DUR$	4.970	28.206	0	306	0.032	0.323	0	6

Table 2: Description of the two data sets for the women.

	OSAMP (N = 796)				ESAMP (N = 526)			
Variable	Mean	StdDev	Min	Max	Mean	StdDev	Min	Max
d	0.080	0.271	0	1	0.025	0.156	0	1
c	10.638	3.734	0.767	41.922	10.914	3.397	1.487	41.922
y	16.631	5.647	3.188	64.004	16.388	4.849	3.449	37.216
HOLLY	0.082	0.275	0	1	0.083	0.275	0	1
WE	0.315	0.464	0	1	0.315	0.465	0	1
BW	0.011	0.104	0	1	0.011	0.104	0	1
Married	0.754	0.431	0	1	0.754	0.431	0	1
Single	0.158	0.365	0	1	0.166	0.372	0	1
Age	39.904	12.179	20	64	38.474	12.172	20	64
Number of children	0.783	1.024	0	5	0.814	1.015	0	5
D_m	0.837	0.369	0	1	0.837	0.369	0	1
c_m	10.625	3.749	0.767	41.922	9.140	5.087	0	41.922
y_m	16.640	5.667	3.188	64.004	13.717	7.497	0	37.216
DUR	5.905	33.143	0	364	0.065	0.482	0	7
$D_m DUR$	5.643	33.053	0	364	0.047	0.408	0	7