THE EFFECT OF SICKNESS ON EARNINGS*

Daniela Andrén[α] and Edward Palmer[β]

Abstract

The question addressed in this paper is whether sickness history affects annual earnings and hourly wages in Sweden. If poor health makes people less productive, we expect to find a negative effect of previous health history on hourly wages. If, instead, poor health reduces people's working capacity, but not their productivity, this implies only a decrease in hours worked. Using a longitudinal database for individual sickness, we estimate both (annual) earnings and (hourly) wage equations, and find that people who are healthy in the current year, but have long-term sickness in the previous five years have lower earnings than persons without long-term sickness.

Key words: human capital model, age-earnings profiles, earnings, and wage equations, sickness spells, health variables.

JEL Classification: I10, I12, J24, J28.

---

* We would like to thank Thomas Andrén, Lennart Flood, Donald Storrie, Roger Wahlberg, and participants at The Third Meeting of Young Economists in Berlin for useful comments on earlier versions of this paper. The usual disclaimer applies.

[α] Göteborg University, Department of Economics, Box 640, SE 405 30 Göteborg, Sweden; E-mail: Daniela.Andren@economics.gu.se

[β] Uppsala University and National Social Insurance Board, SE 103 51 Stockholm, Sweden; E-mail: edward.palmer@rfv.sfa.se

1 Introduction

The human capital approach was built on the analysis of costs and returns to investments in human capital through the computation of earnings differentials. The main results of the human capital model show that earnings depend on investments in individual education and training, but also that the effect on earnings of a given investment in human capital may decline with age. Training and experience will have a positive effect on earnings profiles, but as people approach the end of working life, profiles typically turn downwards. The shapes of an individual's earnings profile and wage profile are not necessarily the same. For example, a decline in annual earnings close to retirement age can be explained by either a decrease in hours worked, or by less overtime with a wage bonus, or by a combination of these. In this study we are interested in the effect of health on annual earnings.

If there are short-term or long-term effects of past poor health on current earnings, we expect that they would take one of the following forms: 1) unchanged hourly wages and fewer hours worked; 2) decreased hourly wages; 3) decreased hourly wages *and* fewer hours of work (per year). This is why, *in examining the effect of health on earnings*, one should analyze *both* annual earnings *and* hourly wages. By studying hourly wages and annual earnings, we can discern whether an effect on earnings is derivable from an effect on hourly wages, or if not, attributable to a change in hours worked. People with a poor previous health history may simply have to face wage discrimination, in spite of unchanged productivity, however this possible effect is not analyzed here.

In general, the investment-earnings relation is a reduced form of two simultaneous structural equations: a demand function relating individual investments to their marginal rates of return, and a supply function relating the obtainable funds for such investments to their marginal costs. *Investments in health* (including a nutritional diet, exercising, environmental quality, etc.) help improve or maintain productivity. If poor health makes people less productive, we expect to find a negative effect of their previous health history on their hourly wages. If poor health reduces working capacity, leading to a decrease in hours worked, we expect to find a negative effect of previous health history on annual earnings. We would expect that while the investments in health keep working capacity from deteriorating and enable people to maintain a normal level of hours worked (and better annual earnings), it would not necessarily increase their hourly wages. This is another reason why, *in examining the effect of health on earnings*, one should analyze *both* annual earnings *and* hourly wages.

In studies of annual earnings and hourly wages, the most common approach is *not* to control for health status. When health status has entered studies, two approaches have been taken. It is either formulated as a binary exogenous variable or it is used as a stratification criterion for obtaining samples of "healthy" and "unhealthy" men and women, black and whites, etc. In this study, we are able to specify health status using information about days of sickness during five previous years, the year when the first spell of long-term sickness was recorded, and the diagnosis. Using these variables, we analyze the annual earnings and hourly wages for a sample of insured people, and various subsamples as well.

Another aspect of "health status" is the *quality* of this variable, which depends on the source of the information (individual interviews and/or register information). We use variables for health status that do not rely on an individual's self-evaluation, which might give a biased measure of health. Our source of data (administrative registers) enables us to use reliable information on compensated days of sickness, sickness diagnosis, and also relatively reliable information on earnings during 1985-1990, and for hours worked, and hence hourly wages for 1988.

This paper adds health related variables to the human capital model in order to analyze the effects of previous poor health on actual hourly wages and annual earnings. In Section 2 we will review literature on human capital and labor supply, considering health. Section 3 outlines a model of earnings, wages, and health, which is followed, in Section 4, by data description, and earnings and wage profiles in Section 5. In Section 6 we present the econometric specification, with the empirical results in Section 7, and the conclusions in Section 8.

2 Previous studies

The "human capital" literature has expanded rapidly since its inception in the late 1950s. Research leading toward a theory of human capital was pioneered by Theodore Schultz,[1] whose work primarily concentrated on a number of strategic questions related to conditions for efficiency in the employment of production resources, attaching crucial importance to vocational skills, schooling, research, and its application. Schultz and his students showed that, for a long time, there was a considerably higher yield on "human capital" than on physical capital in the American economy, and that this tension resulted in a much faster expansion of educational investments than in other investments.

Mincer (1958) formulated the first schooling model. With their seminal papers, Mincer (1958, 1962) and Becker (1962) stimulated a large amount of research on human investment decision. There is little economic research on the effects of health investments, compared to education and training, however.

In earlier work, Becker (1962) mentioned medical care and vitamin consumption as ways of investing in human capital. Furthermore, he referred to investment in mental and physical health that can be made within the firm (medical examinations, luncheons, protection against accidents) and outside firms by individuals. In theory, a firm would be willing to compensate employees for individual costs leading to improve human capital if it could benefit from a resulting increase in productivity. Mushkin (1962) analyzed health as an investment. Her paper dealt with capital formation through health care and with returns to investment in health. Becker (1964) and Fuchs (1966) emphasized that health capital is one component of the stock of human capital, and Grossman (1972a, 1972b) constructed the first model of the demand for health capital itself. According to Grossman (1999), if increases in the stock of health simply increased wage rates, one could apply Becker and Ben-Porath's models to study the decision to invest in health. He argued that health capital differs from other forms of

[1] From "Human Capital and Modern Labor Economics: The Early Days", Gary Becker' s talk to the First EALE/SOLE World Conference, Milan 2000.

human capital. While the personal stock of knowledge affects both the market and the non-market productivity of people, health status also determines the total amount of working time. He used the household production function model of consumer behavior to account for the gap between health as an output and medical care as one of the many inputs into its production. The model also emphasized the equally important difference between health capital and other forms of human capital. It provides a theoretical framework for making predictions about the impacts of many variables on health, and an empirical framework for testing those predictions.

Finally, it is important to recall that annual earnings and earnings profiles of men and women of different ages have always been a subject of interest for economists. Mincer's original study began with the observation that these profiles differ. Fase (1970) notes the historical interest of economists and members of other professions, especially insurance actuaries, in age-income profiles. Normally, economists focus on lifetime income and its distribution over the life cycle. Many researchers have worked with statistical models of earnings profiles to determine differences between groups, e.g., occupational groups. For Sweden, Klevmarken's (1972 and 1992) studies are still the most comprehensive.

Poor health is traditionally associated with a loss of earnings capacity mainly associated with withdrawal from the labor market. The Grossman model of the demand for health (1972 a, b) identified the complex interrelation among work-time, wages, and health. Following the 1972 studies by Grosssman, a lot of studies (mainly done on US data), focused on work, wages and health.

Variables used are the *hourly wage* [e.g., Luft (1975), and Mitchell and Burkhauser (1990)], *log of hourly wages* [e.g., Lee (1982), Johnson and Lambrinos (1985), and Baldwin and Johnson (1994)], *log real hourly wages*, computed dividing annual earnings by annual hours [e.g., Haveman et al.(1994)], *log annual wages* [e.g., Berkovec and Stern, 1991], *annual earnings* [e.g., Mitchell and Burkhauser (1990)], *log annual earnings* [e.g., Luft (1975), Bartel and Taubman (1979), Chirikos and Nestel (1985)], and *log personal annual income* [e.g., Mullahy and Sindelar (1991, 1993, 1995)]. Hours worked have also been analyzed in some of the mentioned earlier studies as hours per week [e.g., Luft (1975)], as log of hours per week [e.g., Bartel and Taubman (1979)], or as annual hours [e.g., Chirikos and Nestel (1985), Mitchell and

Burkhauser (1990), Haveman et al.(1994)].

Grossman and Benham (1974) used the household production model to examine the effect of health on wages (weekly wage) and on weeks worked, treating health as an endogenous variable. The estimated structural equations for wage determination and labor supply indicated that good health had a positive effect on these two components of earnings.

Luft (1975) investigated several aspects of the impact of health status on earnings, including the computation of the overall loss of earnings to the economy in a year. He measured the effects of health status by comparing the different components of earnings (labor force participation, hourly wage, and hours worked per week) of persons who were well with those of persons who were disabled. By analyzing out subsamples of men and women, and blacks and whites, he estimates different ways in which disability affects different groups. His results suggested that there are different ways in which poor health may affect different groups. For example black males are more likely to drop out of the labor force or work fewer weeks than white males, while the latter take larger cuts in hourly wages and annual earnings.

Bartel and Taubman (1979) estimated the effect of specific diseases (physician diagnosed) on wage rates and hours worked to try to determine which diseases have bigger effects on current earnings, and how long the effects exist. They analyzed both earnings and wage rate equations, and explained the different effects using labor supply equations. They found negative effects on both wages and annual earnings given by heart disease/hypertension, psychoses/neuroses, arthritis and bronchitis/asthma.

Chirikos and Nestel (1985) examined the effect of health histories over the preceding ten-year period on current economic welfare, using a two-equation model. First, health history effects on wage rates adjusted for sample selectivity bias were estimated, and then the influence of health history and wages on annual hours of work are evaluated using Tobit regressions. Analyzing people grouped by gender, race, and health status, they found that health problems in the past (up to 10 years) adversely affected current earnings.

Johnson and Lambrinos (1985) used a national sample from the 1972 Social Security Survey of Disabled and Non-Disabled Adults to estimate wage discrimination against handicapped workers. The samples were further partitioned by gender, and four earnings functions were estimated. Their results showed (through three different experience variables) that discrimination had occurred. Handicapped women were also subject to discrimination based on gender. Discrimination accounted for about 33% of offer wage differentials between not handicapped and handicapped men and 40% for women.

Mitchell and Burkhauser (1990) developed a procedure for examining the separate impact of arthritis on wages and hours worked, and these results are translated into earnings. They found that hours worked tended to be considerably more affected by arthritis than were wage rates, although this was more pronounced for men and younger women than for older women.

Mullahy and Sindelar (1991) determined the structure of gender differences in labor market responses to alcoholism. The results indicate that the effects may depend on the control variables, the age distribution of the sample, and the choice between examining participation and income. The effects of alcoholism were greater on household income for women, but it reduced personal income for both men and women. The effect of alcoholism was found to be stronger on labor market participation than on income.

Mullahy and Sindelar (1993) found that inferences about the effects of alcoholism on income depended critically on the age group being studied. Their results also support the proposition that alcoholism has a more significant impact on the likelihood of working than it does on how much is earned when working.

Haveman et al. (1994) studied white males using longitudinal data. They specified a 3-equation simultaneous model, designed to capture interrelationships among work-time, wages and health, which is estimated using Hansen's (1982) generalized method of moments' technique. Simpler models were then estimated with more restrictive assumptions, and substantial differences were found between these estimates and those from the simultaneous model (e.g., the positive relationship between work time and health disappeared when the relevant simultaneities were considered). The implicit demand for health function is the only available estimate that accounted for the interrelationships among health, work time, and wages.

Baldwin and Johnson (1994) analyzed the extent of labor market discrimination against men and women with disabilities in US using the 1984 panel of the Survey of

Income and Program Participation. Using a two-stage estimation of a quasireduced system of wages and health, they found large differences in employment rates and hourly wages between disabled and not disabled men.

Mullahy and Sindelar (1995) expanded the standard approach to the welfare analysis of health-related economic costs by accounting for risk aversion and the variance in income that depends on health status. Their results suggest that an evaluation of alcoholism's welfare costs in terms of productivity differentials alone may significantly understate such costs.

Muller et al. (1996), using multivariate analysis based on US pooled cross-sectional time-series data, estimated the probability that a recipient of a Supplemental Security Income disability benefit would perform work. They also estimated the annual earnings equation. They used distinct groups, based on their diagnoses, and found that patterns of work and earnings varied over time. Changes in the probability of work of disabled and in the level of earnings, seemed to mirror economic trends, as measured by the unemployment rate.

Thomas et al. (1997) investigated the impact of four indicators of health on wages of urban workers in Brazil, finding that health yielded a substantial return, at least in the market wage sector. The indicators used (height, body mass index, *per capita* calorie intake, and *per capita* protein intake) do not fully capture health, but they measure various dimensions of it.

Smith (1998, 1999) noted that there are several pathways through which health may affect wealth accumulation, and referred specifically to lowered earnings and increased medical expenditures.

Currie and Madrian (1999) presented an overview of the American literature linking health, health insurance, and labor market outcomes, such as wages, earnings, employment, hours, occupational choice, job turnover, retirement, and the structure of employment. The empirical literature surveyed by them suggests that poor health reduces the capacity to work and has significant effects on wages, labor force participation, and job choice. It is difficult to conclude anything about the magnitudes of these effects, given that they are sensitive to both the choice of health measures and to identification assumptions.

The literature on this subject in Sweden includes studies of the effect of time out of labor force on subsequent wages that give contradictory results. Gustafsson (1981) used an OLS cross-sectional regression, and estimated the effect of absenteeism on monthly salaries in 1974 for women aged 30-44 years (from a random sample of white-collar workers). She found a negative effect of time spent out of work on monthly salaries. Edin and Nynabb (1992) used a restricted sample of persons employed during the interview week in 1984, with no internal missing values, and reinterviewed in 1986. They found a positive effect on (log) hourly earnings of time out for women and no significant effect for men. Stafford and Sundstrom (1996) used OLS cross-sectional regressions and found negative effects (significant at the 10 % level only for men) of time spent out of work on wages. Sundberg (1996) using basically the same model as the one used by Haveman et al. (1994), where *ill health* is a self-assessed variable, found that poor health affected wages negatively.

Albrecht et al. (1998) examined the effects of career interruption on subsequent wages by estimating cross-sectional and fixed-effects specifications of earnings functions that included time-out of labor force variables, using month-by-month event history data for individuals over their entire work life merged with employer-reported Swesish wage data. They estimated separately the effect of total time, and (disaggregate) time out (parental leave, household time, other time out, diverse leave, unemployment and military service). In the cross section, the total time out had a significant effect on both women's and men's wages. Parental leave had no effect on women's wages, but had a significant negative effect on men's wages.

Skogman Thoursie (1999), for the first time, estimated of the extent of unexplained wage differentials between disabled and nondisabled workers in Sweden, using data from the Swedish Level of Living Survey for 1981 and for 1991. He found that the unexplained component due to differences in returns on wage determinants is insignificant in the 1981 case but is highly significant in 1991, constituting around 50-60% of the average log wage differential.

None of these studies analyzed specifically time out due to sickness. Hansen (2000) uses information about short-term absence among Swedish employees to investigate the potential wage loss attributed to absence, and finds that women's wages were significantly reduced by work absence due to their own sickness, while absence to care for a sick child had no significant wage effect. He also finds that the distribution of the gender wage gap depends to a large extent on work absence.

3 The model

Our point of departure is the human capital model. Central to human capital theory is the assumption that an individual can, by forgoing earnings, spend time on education or training and thereby, augment the quality and the value of his/her labor services. Schultz' (1960) focus on education as a key to raising productivity led to the modern emphasis on "human capital" as a factor in production. Indeed, his work paved the way for Becker's analysis of human skills as a source of productivity growth, which relates hourly earnings to the effects of schooling, on-the-job training and work experience. This is typically expressed as

(1) $$\ln y = \beta_0 + \beta_1 s + \beta_2 \exp + \beta_3 \exp^2 + u$$

where the schooling coefficient (β_1) provides an estimate of the individual return to education (s), and e*xp* is experience.

The typically observed concave profile for lifetime earnings is captured by the experience variable, measured by years of work or approximated by age, and the quadratic of experience, with positive and negative expected values of β_2 and β_3, respectively. Typically, schooling occurs prior to entering the job market and training thereafter, although it is possible for individuals to leave the workforce and acquire more schooling. Human capital is the sum of all investments made in schooling and training in all years. The simple human capital model postulates proportionality between earnings and human capital, with the factor of proportionality being the earnings of the individual. The higher one's human capital, the greater are one's earnings, according to this theory. This model, which is the point of departure in many studies of earnings formation and differences in earnings between various groups, is our point of departure too.

Ben-Porath (1967) developed a model in which, in each year of one's life, one invests in oneself (education, qualifications, experience) in accordance with the benefits and costs of the investment at the stage of the life cycle. Polachek and Siebert (1993) derived an earnings function based on the assumption that, when actual earnings deviate from potential earnings, they do so because of investments in human capital. They introduced Mincer's expression for time equivalent investments, i.e. the fraction of potential earnings foregone in order to acquire (invest in) additional human capital. Total investments encompassed years of schooling (*s*) and years of post-school experience (*t*). Post-school training investments were assumed to increase at a decreasing rate. This means that earnings, i.e., the return on the capital stock, would also increase at a decreasing rate.

Health is similar to education and training in the sense that it is a "stock" that can be enhanced and/or maintained with investment (good nutrition, exercise, etc.) through life, although it is likely that the normal process of aging can increase the likelihood of some specific diseases. We consider the investment in health to be the same as the investment in education and/or training,[2] and try to estimate the effect of poor health on hourly wages and annual earnings. We include in the earnings function[3] (2) school-dummies, age and age-squared (using age as a proxy for the work experience), variables related to personal characteristics (*X*), and health, including previous health history (*Z*):

(2) $$\ln y = \beta_0 + \beta_1 s + \beta_2 age + \beta_3 age^2 + \delta X + \sigma Z + \varepsilon \,.$$

Previous studies of the relation between wages and the characteristics of individuals have focused on the estimation of an earnings function rather than a wage function, because of the lack of data on hourly wages. Willis (1986) provides a survey and exposition of the development of earnings function as an empirical tool for the analysis of determinants of wage rates.

People acquire jobs that require different amounts of health capital. People with little education (and/or training) usually get jobs requiring a lower level of professional

[2] We consider that health investment is (much) easier for people to accumulate, since it does not involve or require too much effort and resources from the individual. Health is an investment for which people do not need skills in order to accumulate capital. Nevertheless, investments in health require financial, informational and labor resources.

[3] "The term *earnings function* has come to mean any regression of individual wage rates or earnings on a vector of personal, market, and environmental variables thought to influence the wage." (Willis, 1986)

skills, but often requiring more physical effort, poorer work environment, etc. In this case, it may be especially important for these people to have good health. In fact, many jobs (food handling, health care, day care, etc.) require good health on a daily basis.

Finally, we note that all individuals in Sweden are covered by public health care, which means that they have free access to medical care. Additionally, the state, health organizations and other institutions distribute a huge amount of information. In other words, people have access to medical and health care, and information concerning prevention and control, and therefore the *investment* is very much determined by the decision of individuals to care for themselves.

4 The data

The data employed in this study come from the Swedish National Social Insurance Board's LS-database. This is a longitudinal database covering spells of sickness during the time period January 1, 1983 to December 31, 1991 for people who are registered with the national sickness insurance scheme[4] in 1986, and are in working age, i.e. 16-64 years, during 1983-1991. We analyze a random sample, representative for insured population in working age, of 1688 individuals. We left out all persons who died or were classified as long-term sick during 1988 (i.e., 187 persons). The reason for choosing the year 1988 is that there is unique data on hours worked because of a change in the social insurance law in December 1987 that required everyone to report hours of work along with current earnings to the social insurance office. We have no information on hours of work for the preceding years, and only information with, we believe, lower quality for the years after 1988.

There are two choices for the earnings variable in our data, both of which have some problems. Since the dataset is based on the social insurance files, we have data on the level of income that was reported (earnings from work and related to work, such as

[4] During the analyzed period all Swedish residents were registered with a social insurance office upon reaching 16 years of age, and were entitled to a daily sickness allowance if they had an income from work of at least SEK 6,000 per year.

sickness cash benefit during spells of sickness). Over all six years of the study period, we have earnings data of this kind for about 72% of all year-persons. An alternative measure of earnings, taxable income that gives pension rights, which collected on a yearly basis, was used to fill in the missing values.

The level of education is another constructed variable, measured by three categories: (1) *low*, which means primary and secondary education; (2) *medium*, which includes gymnasium and post-gymnasium education, but without a university degree; and (3) high, which means at least a university degree. The information on education was collected by interview.

Table 1 presents basic descriptive statistics for the groups of men and women, not long-term sick in 1988, and stratified according to their sickness history.

Table 1 Descriptive statistics (individual characteristics) of men and women, by sickness status 1983-88

Variables	Not LT sick in 1988				Not LT sick during 1983-88				LT sick during 1983-87 & Not LT sick in 1988			
	Men (N=849)		Women (N=839)		Men (N=766)		Women (N=737)		Men (N=83)		Women (N=102)	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Age	41.09	11.24	40.80	11.28	40.62	11.08	40.45	11.09	45.43	11.86	43.34	12.38
Age-groups												
16-30 years	0.22	0.42	0.22	0.42	0.23	0.42	0.23	0.42	0.13	0.34	0.19	0.39
31-45 years	0.42	0.49	0.42	0.49	0.43	0.50	0.43	0.50	0.37	0.49	0.35	0.48
46-50 years	0.11	0.31	0.12	0.32	0.11	0.32	0.12	0.32	0.08	0.28	0.10	0.30
51-55 years	0.10	0.30	0.11	0.31	0.10	0.30	0.11	0.31	0.11	0.31	0.14	0.35
56-65 years	0.14	0.35	0.13	0.34	0.12	0.33	0.12	0.32	0.30	0.46	0.23	0.42
Citizenship												
Swedish born	0.88	0.32	0.88	0.33	0.89	0.32	0.87	0.34	0.84	0.37	0.94	0.24
Foreign born	0.06	0.24	0.09	0.28	0.06	0.25	0.09	0.29	0.05	0.22	0.03	0.17
Nationalized	0.05	0.23	0.04	0.20	0.05	0.22	0.04	0.20	0.11	0.31	0.03	0.17
Education												
Low	0.51	0.50	0.47	0.50	0.49	0.50	0.45	0.50	0.69	0.47	0.62	0.49
Medium	0.35	0.48	0.34	0.47	0.36	0.48	0.35	0.48	0.27	0.44	0.26	0.44
High	0.14	0.34	0.20	0.40	0.15	0.35	0.21	0.40	0.05	0.22	0.13	0.34
Marital status												
Unmarried	0.41	0.49	0.34	0.47	0.42	0.49	0.34	0.48	0.33	0.47	0.30	0.46
Married	0.54	0.50	0.55	0.50	0.54	0.50	0.55	0.50	0.58	0.50	0.58	0.50
Divorced	0.04	0.20	0.10	0.29	0.04	0.19	0.10	0.30	0.10	0.30	0.09	0.29
Widower	0.01	0.08	0.01	0.12	0.01	0.09	0.01	0.11	0.00	0.00	0.03	0.17
Zero earnings	0.07	0.25	0.07	0.25	0.06	0.24	0.06	0.24	0.17	0.38	0.12	0.32
Annual hours w	1591	837	1287	849	1628	812	1324	838	1247	977	1025	870
Hourly wages (kr)	68	36	54	25	70	36	55	25	55	39	51	28
Earnings*	142	80	94	51	145	78	96	51	118	88	83	48

Note: *Italics indicates dummy variables. **Earnings have been adjusted to 1997 values using the CPI, and are expressed in thousand Swedish crowns per year.

The first two columns include all men and women in the study, while the next two sets of two columns each divide the sample into those who were not, or were, long-term (LT) sick sometime during 1983-87. The descriptive statistics are for year 1988.

A first conclusion is that those who experienced at least one long-term sickness spell during 1983-87 were generally older and had a lower level of education than those who did not. However, the proportion of older men with long-term sickness history is greater than that of women in the same age group. This could reflect the statistical fact that older women with sickness history tend to leave the labor market with full disability more often than men do[5].

Given the almost perfect gender-balance of the total sample, it is also clear that a higher proportion of women than men had a history of LT sickness. In addition, among persons younger than 30 years with long-term sickness, women occur more frequently, which could be explained by problems in conjunction with pregnancy.

Another difference has to do with citizenship. The proportion of naturalized men (i.e., foreign-born men with Swedish citizenship) is much higher among those with long-term sickness (10.8% compared with 5.4% for the total sample). This could reflect poor human capital and/or health capital of men who came to Sweden during the 1970's or earlier, and/or the likelihood that their jobs were mainly in categories requiring greater physical effort or consisting of more demanding work environments. The proportion of Swedish-born women is much higher in the group of women with previous spells of long-term sickness (94.1%, compared to 87.6% of total sample). An expected difference is related to hours worked per year (*Annual hours w*, in Table 1), hourly wages and earnings. Both men and women who had not experienced a long-term sickness spell in the preceding five years worked more hours per year, with higher hourly wages and higher annual earnings than did people who had had such long-term sickness spell. The difference in hourly wages for these groups is much higher for men

[5] Andrén (2000, Paper 4 of this thesis).

than for women, which could be explained by the relatively better education of those without long-term sickness in the preceding five years.

The variable *Zero earnings* refers to people who reported zero earnings on their tax declaration for 1988. These are people (about 7%) who declared that they had received neither salary nor any compensation for earnings loss due to an event such as unemployment, sickness, disability, or parental leave.

Table 2 presents further descriptive statistics (of health variables) for the groups of men and women, classified as in Table 1.[6] Sickness history is measured on an *individual* basis first by the number of days of sickness (*Compensated days of sickness per year*), second by sickness cohort[7] dummies (*Year* 1983 - *Year* 1987), which reflect the year of the first long-term sickness, and third by diagnosis. The average compensated days of sickness increased by year "only" for people who were not long-term sick during 1983-1988. The decrease for the other two groups during 1987 and 1988 is explained by the design of the sample: we selected only those people who were not long-term sick or disabled in 1988. This implies that we "left out" people with ongoing spells of long-term sickness in 1988 (about 80% of these spells started in 1987, and about 10% in 1986 or earlier).

We use two diagnosis variables, days of sickness with a specific diagnosis, and dummies for the occurrence of long-term sickness spells with a specific diagnosis (*musculoskeletal, cardiovascular, respiratory, mental, general symptoms, injuries and poisonin*g, and *other*). Unfortunately, the database is not large enough to enable us to use finer categories, which might more clearly reveal the true effects of the more serious diagnoses. Data on diagnoses cover the observation period 1986-1991 for both short-term and long-term spells of sickness, whereas we have information on duration of the sickness spells for the period 1983-1991. We present the descriptive for the period 1986-1988. During this period, women had on average more compensated days of sickness than men, and in fact this is generally true in Sweden, regardless of the period

---

[6] A more detailed analysis of the different sickness cohorts is presented in Table A1 in Appendix.

[7] A sickness cohort *j* consists of people who had their first long-term sickness in year *j*, i.e. people selected with regard to the occurrence of a first (within the window observation) long-term sickness spell.

studied. Women also had on average more days of compensated sickness with general symptoms and other diagnoses, while men had more days with injuries or poisoning.

With respect to *long-term spells*, the musculoskelatal diagnosis (usually back pain) was the most frequent for both men and women, while cardiovascular, respiratory and general symptoms were least frequent.

Table 2 Descriptive statistics (health variables) of women and men, by sickness status 1983-1988

Variable	Not LT sick in 1988				Not LT sick during 1983-1988				LT sick during 1983-1987 & not LT sick in 1988			
	Men (N=849)		Women (N=839)		Men (N=766)		Women (N=737)		Men (N=83)		Women (N=102)	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Compensated days of sickness per year												
1983	6.65	15.00	8.84	18.48	5.13	10.49	6.75	11.73	20.65	32.87	23.96	39.58
1984	8.73	27.36	12.18	37.57	5.26	11.91	6.75	13.29	40.75	72.60	51.45	93.01
1985	11.97	47.53	12.54	38.54	5.85	12.08	7.69	14.16	68.48	135.7	47.54	97.23
1986	12.05	38.53	17.80	48.82	5.80	10.92	8.58	21.53	69.72	102.5	84.45	106.2
1987	9.89	26.57	15.93	40.90	6.40	13.01	8.71	16.23	42.11	67.52	68.11	93.96
1988	7.40	11.80	9.40	12.81	6.71	10.83	8.80	12.30	13.77	17.34	13.77	15.43
Sickness cohorts												
Year 1983*	0.02	0.13	0.03	0.16					0.17	0.38	0.23	0.42
Year 1984	0.03	0.17	0.03	0.16					0.31	0.47	0.22	0.41
Year 1985	0.02	0.15	0.03	0.16					0.23	0.42	0.23	0.42
Year 1986	0.01	0.11	0.02	0.15					0.13	0.34	0.20	0.40
Year 1987	0.02	0.12	0.02	0.13					0.16	0.37	0.14	0.35
Sickness spells by diagnosis, 1986-88												
Musculoskeletal	0.47	1.20	0.53	1.26	0.42	1.13	0.46	1.16	0.92	1.68	1.03	1.77
Cardiovascular	0.03	0.23	0.06	0.30	0.03	0.21	0.06	0.31	0.06	0.36	0.07	0.29
Respiratory	1.59	2.37	2.19	2.85	1.57	2.34	2.17	2.79	1.75	2.69	2.31	3.25
Mental	0.06	0.42	0.08	0.41	0.04	0.35	0.06	0.35	0.23	0.80	0.19	0.71
Gen. symptoms	0.35	0.90	0.58	1.21	0.33	0.88	0.57	1.17	0.61	1.05	0.68	1.48
Injuries etc.	0.26	0.72	0.19	0.49	0.23	0.68	0.16	0.45	0.58	0.95	0.36	0.73
Other	0.69	1.43	1.27	2.01	0.64	1.38	1.20	1.94	1.11	1.77	1.75	2.41
Compensated days of sickness by diagnosis, 1986-88												
Musculoskeletal	12.50	75.10	17.02	96.67	5.46	30.32	5.39	19.94	77.48	212.1	101.0	257.9
Cardiovascular	2.45	39.25	5.12	80.92	1.69	35.85	0.61	4.12	9.51	62.34	37.7	230.2
Respiratory	9.38	48.93	10.63	20.11	7.44	12.92	10.07	15.32	27.24	151.1	14.7	40.33
Mental	4.74	64.13	3.74	35.24	0.74	6.64	1.30	9.28	41.63	201.4	21.4	96.52
Gen. symptoms	2.10	11.68	4.07	25.44	1.67	5.50	2.74	7.58	6.10	33.32	13.7	69.61
Injuries	6.33	44.26	3.94	30.06	3.92	34.22	1.88	8.82	28.53	93.72	18.8	81.72
Others	6.32	29.88	12.61	39.34	4.13	12.85	8.74	27.62	26.52	85.08	40.6	79.88
LT sickness by diagnosis, 1986-87												
Musculoskeletal*	0.021	0.144	0.031	0.173					0.169	0.377	0.225	0.420
Cardiovascular	0.004	0.059	0.005	0.069					0.024	0.154	0.010	0.099
Respiratory	0.002	0.049	0.001	0.035					0.072	0.261	0.059	0.236
Mental	0.007	0.084	0.010	0.097					0.012	0.110	0.049	0.217
Gen. symptoms	0.001	0.034	0.007	0.084					0.145	0.354	0.069	0.254
Injuries etc.	0.018	0.132	0.011	0.103					0.096	0.297	0.206	0.406
Other	0.011	0.102	0.027	0.163					0.096	0.297	0.206	0.406

Note: *Italics indicates dummy variables.

5 Earnings and wage profiles

Since there is interest in comparison of earnings between points of time and among individuals, we have tried to distinguish between these two aspects, but at the same time to link them together, using cross sectional profiles and cohort profiles. Differences among individuals measured in cross sections are commonly interpreted as if the same differences had been observed over time. This sort of interpretation, with implicit agreement of results between cross sections and time series, cannot of course be expected to hold in general. Nevertheless, this is the best we can do, without much longer longitudinal data sets.

We analyze the age-earnings and age-wage profiles, taking into account sickness history (both diagnosis and duration of sickness). Figures 1 and 2 show the age-earnings profiles, for men, women, and combined, with and without considering persons with zero-earnings in the years when they did not work.

The difference due to gender is only reflected in the level, as the shapes are similar (which is clearer in Figure 2, where zero-earnings are excluded). The earnings increase was strongest until 30–35 years, after which earnings were relatively flat until 50-53 years, when they start to decline. Given the shape of the age-earnings profiles, we expect a positive effect of age on earnings, and a negative effect of age-squared. What we cannot determine from the age-earnings profiles is whether the later decline in earnings is solely the result of health, even if we compare the age-earnings profiles for the sickness cohorts (persons who were sick with a spell of at least 60 days) with the others. It could be that what we are seeing is really the effect of other factors, such as education, vocational training, experience and/or increasing age. The remainder of our paper is devoted to the task of going deeper into analysis.

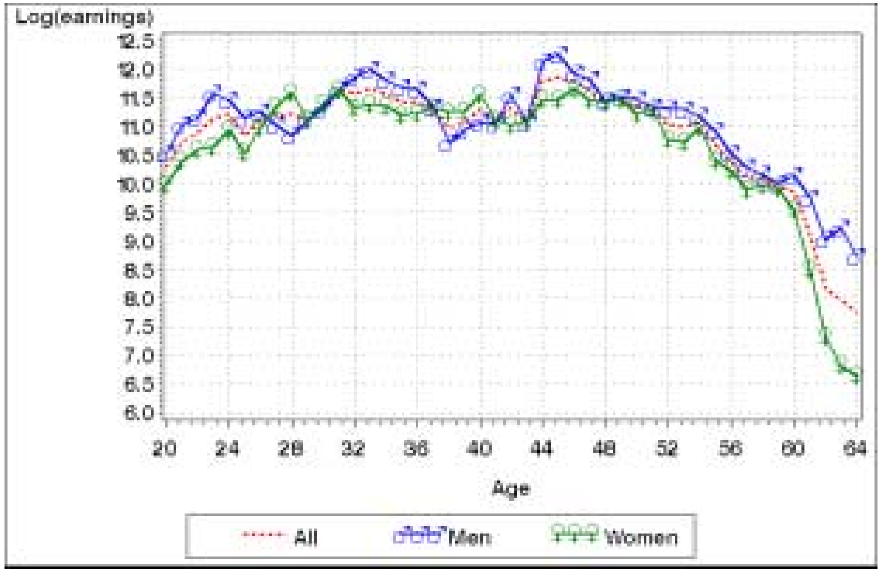

Figure 1 Age-earnings profiles, considering zero-earnings

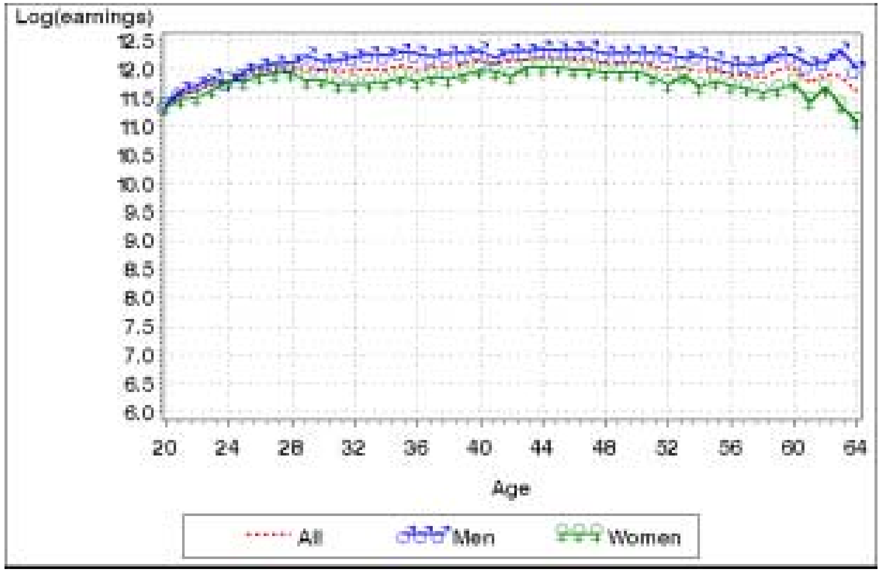

Figure 2 Age-earnings profiles, without considering zero-earnings

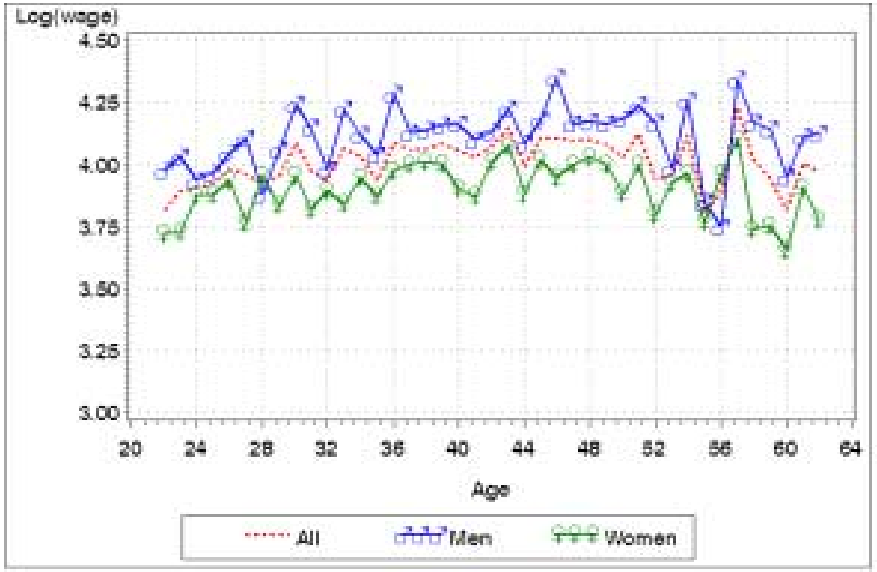

Figure 3 Age-wage profiles

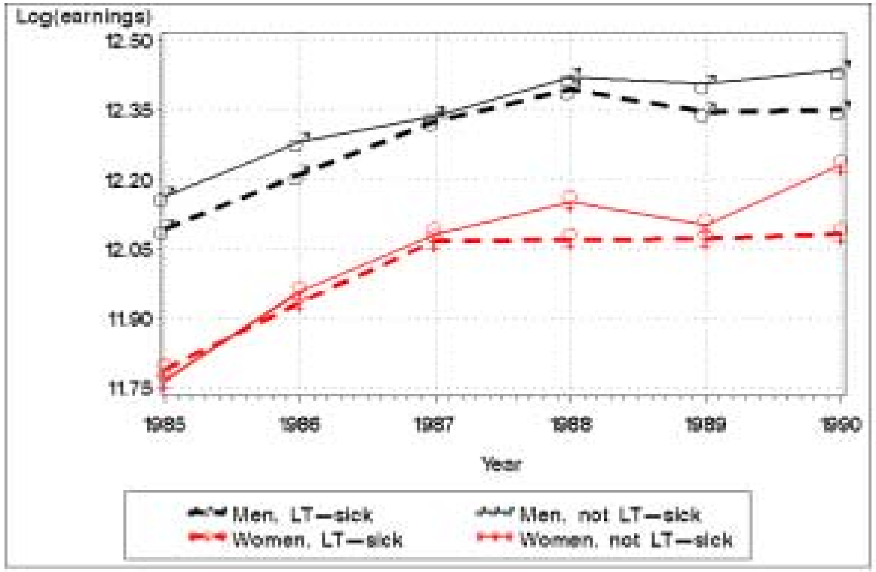

Figure 4 Earnings profiles by gender, and sickness groups

Given the lack of data, we could not get a picture of the age-wage profiles over the same period as for earnings. Figure 3 shows a picture of the age-wage profiles by gender for 1988. They are flatter than the age-earnings profiles, and as expected, women's wages were generally lower than men's.

Figure 4, which shows the earnings profiles of people who were not long-term sick in 1988, illustrates another impact of sickness on earnings. Given absenteeism from the labor market due to long-term sickness sometime during 1983-87, both women and men earned less during this period when the replacement rate was 90%, but also in 1988 and in the following years. This could imply that working capacity was not fully recovered because either people experienced new sickness spells, or returned to work for fewer hours, or both.

Figure 5 shows that earnings profiles also differ across diagnoses. The higher average earnings of people who had a musculoskeletal diagnosis could be explained by the high proportion of people aged 40-50 years with these problems, who have in average higher experience. For those with cardiovascular diagnosis the explanation could be higher education (i.e., more stressful jobs), but more likely, it reflects a tendency for persons not to be out of work until a serious cardiovascular condition occurs, and then when this occurs, they are also usually over 40.

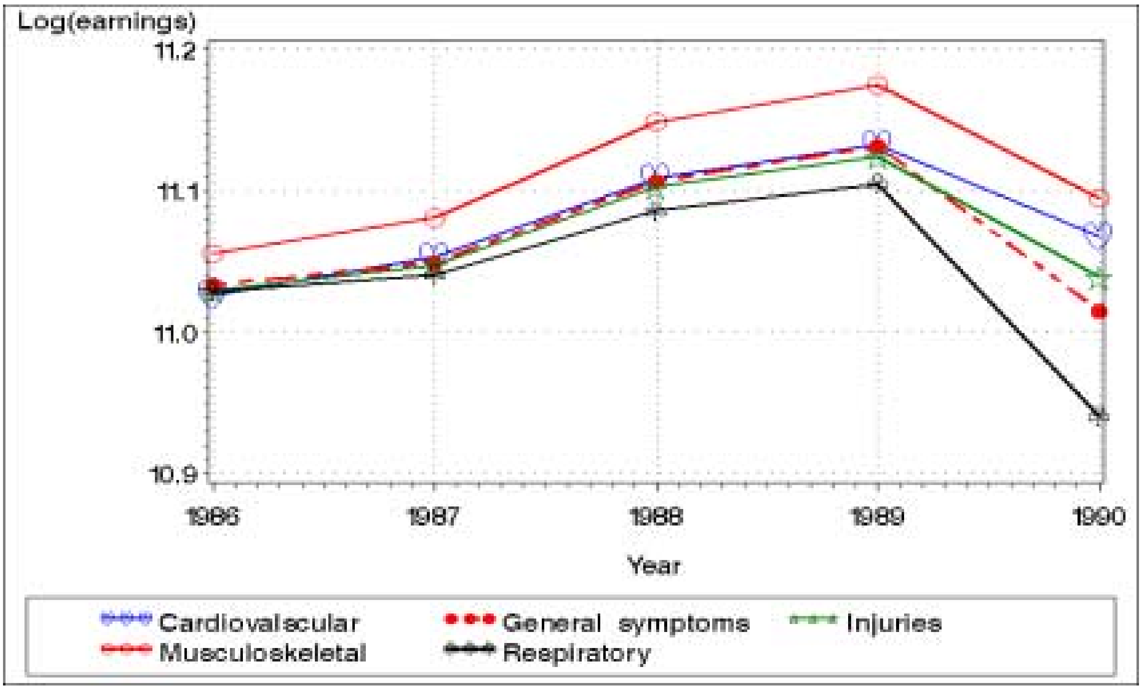

Figure 5 Earnings profiles by diagnosis

Figure 6 shows the "sickness effects" on earnings by sickness-cohort, compared to the almost constant average earnings over time for insured people who did not experience long-term sickness at all during 1983-1988. Given the short span of time and the (very low) number of observations for each cohort, it is difficult to draw conclusions about the effect of sickness on earnings in the long-run, but it seems that the average earnings of people with earlier long-term sickness spells decreased after some years, which could be explained by a full, or partial, exit from the labor market, for some period of time, or permanently.

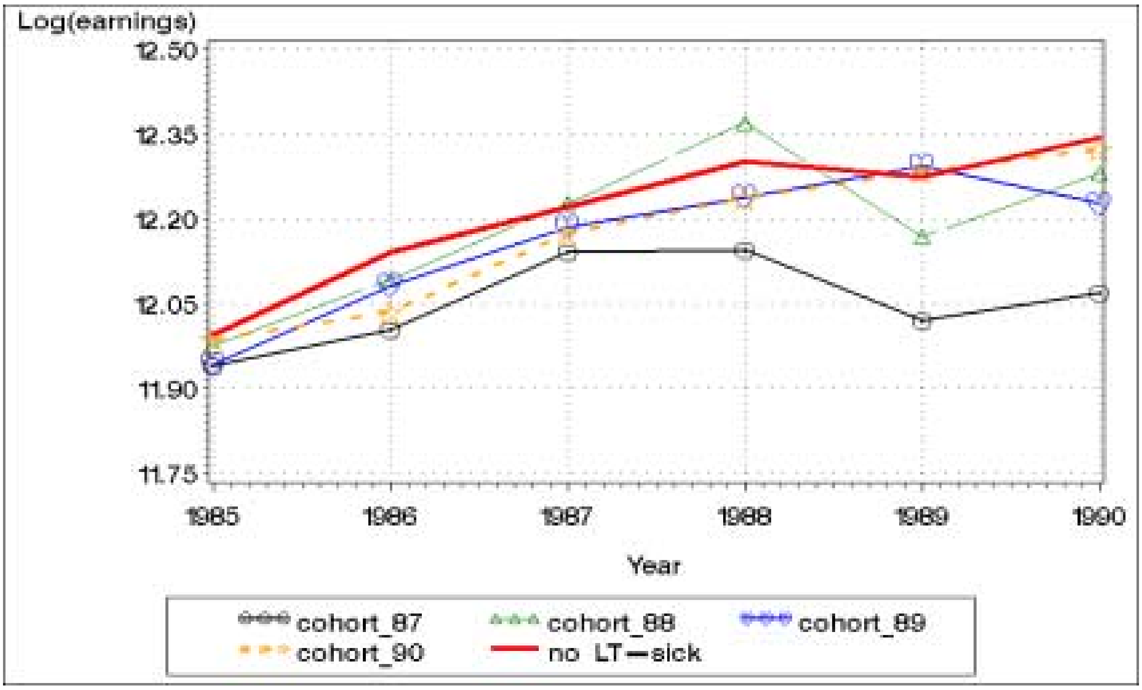

Figure 6 Age-earnings profiles for sickness cohorts

6 The econometric specification

In order to estimate the effect of sickness on earnings, we use an empirical model based on equation **(2)**, with the following form

(3) $\ln y_i = x_i^{'}\beta + u_i, \; i = 1, 2, \ldots, n$

where y_i is the earnings variable, i.e. annual earnings or hourly wage, for the individual

i, β is the parameters vector, and x_i is a vector of known constants, u_i are the residuals that are independently and normally distributed, with mean zero and a common variance σ^2.

We estimate two models that differ only by the dependent variable. In the first model we use annual earnings as a dependent variable, while in the second model we use hourly wages. An important characteristic of the data is that there are several observations where annual earnings are zero, and even more observations where the hourly wage is not observed. If the data on annual earnings have a mass-point at zero, the linearity assumption might be destroyed so that the least squares method would be inappropriate for estimating the earnings equation. If the dependent variables are limited in their range, Tobit models are the appropriate approach for estimating such regressions.

The Standard Tobit model is used for estimating the annual earnings equation. Similar to the pioneering work of Tobin (1958), who used data with several zero values for the dependent variable, we use annual earnings as the dependent variable, which has zero values when people neither worked, nor were absent from work due to sickness.[8] The Standard Tobit model is:

$$\textbf{(4)} \quad \begin{cases} y_i^* = x_i^{'}\beta + u_i \\ y_i = \begin{cases} y_i^* \text{ if } y_i^* > 0 \\ 0 \quad \text{if } y_i^* \leq 0 \end{cases} \end{cases}$$

where β is a vector of unknown parameters, x_i is a vector of known constants, and the residuals u_i are assumed to be identically and independently distributed (i.i.d.) drawings from $N(\beta, \sigma)$. It is assumed that x_i and are y_i observed for $i = 1, 2, \ldots, N$, but the y_i^* are unobserved if $y_i^* \leq 0$. If one can at least observe the exogenous variables, x_i, when $y_i^* \leq 0$ (i.e. our data used in estimating annual earnings equation), the model is known as the *censored* version of the Standard Tobit model, and has the following likelihood

[8] These people have a reservation wage greater than zero.

function:

$$\textbf{(5)} \quad L_c = \prod_{y_i^* \le 0} \left[1 - \Phi\left(x_i'\beta / \sigma\right)\right] \prod_{y_i^* > 0} \frac{1}{\sigma} \phi\left[\left(y_i - x_i'\beta\right) / \sigma\right]$$

where Φ and ϕ are the distribution and density function, respectively, of the standard normal variable. Our problem is to estimate β and σ^2 on basis of N observations on y_i and x_i.

When the dependent variable is hourly wages, we use Heckman's full information maximum likelihood method. If y_1 is the wage offer (or market wage) and y_2 is the reservation wage, we never observe y_2 but we observe y_1 for most of people who work. If $y_1 > y_2$, we observe that the individual is in the labor force. If $y_1 < y_2$, we observe that the individual is not employed, and we do not observe either y_1 or y_2. The partitioning of the sample, for example, into employed and not employed is based on the "self-selection" of individuals into the two groups based on the relationship between wage offers and reservation wages. Selectivity is an important issue, which is often considered in the estimation of labor supply models by considering only the subsample of individuals who work. We want to estimate the wage equation **(6a)** of the following model:

$$\textbf{(6)} \quad \begin{cases} y_{1i}^* = x_{1i}'\beta_1 + u_{1i} & \text{(a)} \\ y_{2i}^* = x_{2i}'\beta_2 + u_{2i} & \text{(b)} \\ y_{1i} = y_{1i}^* \ \text{ if } y_{2i}^* > 0 & \text{(c)} \\ y_{1i} = 0 \quad \text{if } y_{2i}^* \le 0. & \text{(d)} \end{cases}$$

We do not know the wages for people who do not work, and therefore we use **(6c)** and the selection equation **(6b).** Becker's wage equation **(1)** relates hourly earnings to the effect of years of school and work experience. Given the fact that people who do not work are usually those who are only able to get fairly low wages given (some of) their observed characteristics that are in both x_1- and x_2- vectors, u_1 and u_2 are expected to be positively correlated. It is assumed that u_1 and u_2 are i.i.d. drawings from a bivariate normal distribution with zero mean, variances σ_1^2 and σ_2^2, and covariance σ_{12}, that only the sign of y_{2i}^* is observed, and that y_{1i}^* are observed only when $y_{2i}^* > 0$. Given these assumptions, model **(6)** has the following likelihood function:

$$(7)\quad L=\prod_{y_{2i}^{*}\le 0}\left[1-\Phi\left(x_{2i}^{'}\beta_2/\sigma_2\right)\right]\prod_{y_{2i}^{*}>0}\Phi\left\{\frac{x_{2i}^{'}\beta_2+\frac{\sigma_{12}}{\sigma_1^2}\left(y_{1i}-x_{1i}^{'}\beta_1\right)}{\sqrt{\sigma_2^2-\frac{\sigma_{12}^2}{\sigma_1^2}}}\right\}\times\frac{1}{\sigma_1}\phi\left(\frac{y_{1i}-x_{1i}^{'}\beta_1}{\sigma_1}\right)$$

Instead of using this full information maximum likelihood method, Heckman's two-step method can be used. Puhani (2000) shows that exploratory work to check for collinearity problems is strongly recommended before deciding on which estimator to apply. In the absence of collinearity problems, the full-information maximum likelihood estimator is preferable to the limited-information two-step method of Heckman, although the latter also gives reasonable results.

Under the assumption that the regression and selection models are both correctly specified, a test whether lambda[9] is significantly non-zero checks that the disturbances of the selection and regression processes are correlated. If that test is not significant, then we cannot reject the hypothesis that the selection and regression process are not correlated. If they are not correlated, selection is random. In this case, we effectively have a regression process with some data that are missing at random and OLS will produce unbiased estimates for both parameters and standard errors.

7 Results

We estimated equations for both annual earnings and the hourly wage. All equations were estimated for 1988, considering sickness records in the preceding five years.

Table 3 presents Tobit estimated coefficients of the annual earnings equation for all individuals together, and for different groups: men, women, people with no long-term sickness spell during 1983-1988, and those with at least one spell of long term sickness during 1983-87. As noted earlier (from Table 1), about 7% of observations had

[9] Lamda is the inverse Mills ratio, $\lambda_i = \phi_i / \Phi_i$, where Φ_i and ϕ_i are the distribution and density function, respectively, of a standard normal variable.

zero annual earnings. Given this, it is not surprising that the Tobit estimates are almost the same as the OLS estimates (reported in Table A2 in Appendix) in size, sign, and significance.

Table 3 Estimated parameters of the earnings equation in 1988

Variables	All Insured		Men		Women		Not long-term sick 1983-1988		Long-term sick 1983-1987	
	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.
Female	**-0.55**	0.14					**-0.60**	0.15	-0.14	0.62
Age	**0.41**	0.05	**0.40**	0.07	**0.43**	0.07	**0.43**	0.05	0.18	0.22
Age-Squared/100	**-0.52**	0.06	**-0.51**	0.08	**-0.55**	0.08	**-0.55**	0.06	-0.27	0.25
Citizeanship (CG: Swedish born)										
Foreign born	**-2.19**	0.28	**-2.42**	0.42	**-2.06**	0.36	**-2.20**	0.27	-0.48	1.57
Nationalized	**-0.61**	0.34	**-0.91**	0.45	-0.50	0.51	**-0.84**	0.35	1.03	1.22
Education (CG: low)										
Medium	**0.93**	0.17	**0.98**	0.23	0.83	0.24	**0.86**	0.17	1.07	0.77
High	**0.95**	0.21	**0.57**	0.32	1.18	0.28	**0.90**	0.21	**2.20**	1.13
Married	0.05	0.18	0.38	0.26	-0.27	0.23	0.08	0.18	-0.24	0.73
Sickness cohorts										
Year 1983	**-2.10**	0.67	**-2.61**	1.02	-0.98	0.89			-1.64	1.47
Year 1984	**-1.14**	0.52	**-1.61**	0.71	-0.85	0.74			-0.78	1.13
Year 1985	-0.21	0.51	0.26	0.76	0.10	0.68			-0.32	1.13
Year 1986	-0.13	0.61	1.49	0.99	-0.85	0.77			0.09	1.16
Year 1987	-0.54	0.63	**-1.96**	0.98	0.98	0.85				
Sickness days (1986-88), by diagnosis										
Musculoskeletal	0.002	0.001	**0.004**	0.002	-0.002	0.001	0.003	0.003	0.002	0.002
Cardiovascular	0.001	0.001	0.001	0.003	-0.001	0.001	0.001	0.003	0.000	0.002
Respiratory	0.003	0.002	0.002	0.002	0.008	0.006	**0.013**	0.006	0.001	0.003
Mental	-0.001	0.002	0.002	0.002	**-0.008**	0.003	-0.006	0.009	0.000	0.002
Gen. symptoms	0.000	0.004	**-0.021**	0.010	0.004	0.004	0.006	0.012	-0.003	0.005
Injuries	**-0.005**	0.002	0.000	0.002	**-0.014**	0.004	0.003	0.003	**-0.013**	0.004
Other	0.000	0.002	0.004	0.004	-0.002	0.003	0.002	0.004	-0.001	0.004
Compensated days of sickness, by year										
1983	**0.020**	0.005	0.006	0.008	**0.019**	0.007	0.011	0.007	0.016	0.011
1984	**-0.005**	0.003	**0.008**	0.005	**-0.012**	0.004	0.003	0.007	**-0.010**	0.004
1985	**-0.007**	0.002	**-0.012**	0.003	-0.003	0.003	0.002	0.006	**-0.009**	0.003
1986	-0.003	0.003	**-0.018**	0.005	**0.007**	0.003	0.000	0.005	-0.002	0.005
1987	0.003	0.003	**0.017**	0.006	0.000	0.003	0.004	0.006	0.000	0.005
1988	**0.051**	0.006	**0.051**	0.009	**0.048**	0.009	**0.036**	0.008	**0.070**	0.020
Intercept	**2.959**	1.022	**2.964**	1.455	**2.279**	1.395	**2.487**	1.037	7.405	4.628
Ancillary parameter	2.902	0.053	2.885	0.075	2.790	0.073	2.747	0.053	3.838	0.226
Left-censored obs.	117		59		58		91		26	
Uncensored obs.	1571		790		781		1412		159	
LR chi-squared[a]	375.47		221.22		225.17		293.360		73.51	
Log-likelihood	-4097		-2054		-2004		-3589.2		-471.8	
Pseudo-R^2	0.044		0.051		0.053		0.039		0.072	

Notes: [a]Prob >chi-squared = 0.000 for all samples. *Param.* indicates the parameter estimate in the semi-log annual earnings equation; **Bolds** indicate parameters significant at less than 5 % -level; *Italics* indicate dummy variables. CG is the comparison group.

Both the Tobit and OLS estimates show that previous sickness history (in 1983, 1984, and 1985) had a negative impact on the amount of 1998 annual earnings. The effect was higher for men. Being woman has a negative effect on the annual earnings, which is not significant for those with long-tem sickness records before 1988, however. This indicates that persons in this situation (i.e., LT-sick in the past) are not different in this respect.

Age has a significant positive impact on annual earnings. It is higher for those without recent long-term sickness than for those with, which is an indication of the sickness on earnings. Age-square's effect is negative, annual earnings increased with age at a decreasing rate, as expected.

Medium and high education had positive effects on annual earnings. The high education effect is even higher for people with recent long-term sickness than for those without. The high education effect, which is one of the few significant effects on the earnings of people with recent long-term sickness, is much higher for women than for men.

Marital status had a negative effect on women's annual earnings, while it is positive for men, but in neither case was significant at the 10% level, or less. In general, foreigners and nationalized Swedes, both men and women earned less than did Swedish born people. However, for those with recent long-tem sickness there was a positive effect (though not significant at the 10% level) for nationalized Swedes with previous long-term sickness, compared with Swedish born people.

Table 4 shows the coefficients of the wage equation estimated using Heckman full maximum likelihood. Given that we do not know the hourly wages of people who did not work during 1988, these estimates are better than the OLS estimates (presented in Table A3, in Appendix). Women had lower wages than men regardless of their previous history of sickness, but the difference was smaller when only people with recent long-term sickness were compared. As expected, the hourly wage increased with age, at a decreasing rate. Medium and high education had positive effects on hourly wages, higher for men than for women.

Foreigners had lower hourly wages than Swedish born people had, the difference being higher for women than for men, and more than double for those with recent long-term sickness, compared to those without. In fact, being a "foreigner" had the most significant negative effect on the *wage rate* of those with long-term sickness history.

Table 4 Estimated parameters of the wage equation in 1988

Variables	All Insured		Men		Women		Not long-term Sick 1983-1988		Long-term Sick 1983-1987	
	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.
Female	**-0.21**	0.02					**-0.22**	0.02	**-0.14**	0.05
Age	**0.04**	0.01	**0.04**	0.01	**0.03**	0.01	**0.04**	0.01	**0.04**	0.02
Age-Squared/100	**-0.04**	0.01	**-0.04**	0.01	**-0.04**	0.01	**-0.04**	0.01	**-0.05**	0.02
Citizenship										
Foreign	**-0.11**	0.04	-0.06	0.06	**-0.14**	0.04	**-0.10**	0.04	**-0.24**	0.13
Nationalized	0.02	0.04	0.03	0.06	-0.01	0.06	0.05	0.05	-0.14	0.10
Education										
Medium	**0.11**	0.02	**0.20**	0.03	0.02	0.03	0.13	0.02	-0.02	0.06
High	**0.24**	0.03	**0.33**	0.04	**0.16**	0.03	0.26	0.03		
Married	0.03	0.02	**0.08**	0.03	-0.01	0.03	0.03	0.02	0.02	0.06
Sickness Cohorts										
Year 1983	0.08	0.09	0.05	0.15	0.03	0.10			-0.03	0.12
Year 1984	0.06	0.07	0.00	0.11	0.11	0.09			-0.08	0.08
Year 1985	-0.10	0.06	-0.13	0.10	-0.07	0.08			**-0.24**	0.07
Year 1986	**0.18**	0.07			**0.18**	0.08				
Year 1987	-0.04	0.08			0.00	0.10				
Sickness Days (1986-88), by diagnosis										
Musculoskeletal	-0.0002	0.0001	-0.0005	0.0003	-0.0001	0.0002	-0.0003	0.0004	-0.0002	0.0001
Cardiovascular	-0.0001	0.0002			0.0000	0.0002				
Respiratory	0.0004	0.0002	0.0003	0.0003	0.0007	0.0006	0.0009	0.0007	0.0003	0.0002
Mental	-0.0001	0.0002	-0.0001	0.0003	0.0001	0.0005	-0.0013	0.0012	0.0001	0.0002
Gen. Symptoms	0.0008	0.0005	0.0023	0.0024	0.0007	0.0004	0.0028	0.0014	0.0009	0.0004
Injuries	0.0001	0.0003	0.0002	0.0004	-0.0002	0.0006	0.0001	0.0004	**-0.0002**	0.0005
Others	-0.0002	0.0003	**-0.0011**	0.0005	0.0002	0.0003	-0.0003	0.0005	-0.0004	0.0003
Compensated days of sickness, by year										
1983	-0.0007	0.0007	-0.0017	0.0012	-0.0002	0.0008	-0.0002	0.0009	-0.0012	0.0010
1984	-0.0001	0.0004	-0.0004	0.0006	0.0001	0.0005	-0.0003	0.0008	-0.0002	0.0005
1985	-0.0002	0.0003	0.0004	0.0007	-0.0005	0.0003	-0.0003	0.0008	-0.0004	0.0003
1986	-0.0002	0.0003	0.0004	0.0006	-0.0004	0.0004	-0.0006	0.0006	0.0002	0.0003
1987	-0.0001	0.0004	0.0000	0.0007	-0.0005	0.0004	0.0006	0.0008	**-0.0007**	0.0004
1988	0.0010	0.0008	0.0001	0.0012	0.0013	0.0010			0.0015	0.0016
Intercept	**3.35**	0.13	**3.18**	0.19	**3.29**	0.16	**3.27**	0.14	**3.68**	0.37
rho	-0.39	0.10	-0.44	0.12	-0.58	0.11	-0.31	0.13	-0.42	0.28
sigma	0.34	0.01	0.37	0.01	0.31	0.01	0.35	0.01	0.29	0.02
lambda	-0.14	0.04	-0.16	0.04	-0.18	0.04	-0.11	0.05	-0.12	0.09
LR Test of independent equations (rho=0)										
Chi2(1)	11.64		11.21		12.66		4.74		1.77	
Prob >chi2	0.001		0.001		0.000		0.030		0.184	
Censored obs.	117		59		58		91		26	
Uncensored obs.	1571		790		781		1412		159	
Wald chi2	331.52		140.44		97.1		293.49		62.81	
Log likelihood	-797.64		-421.63		-274.89		-721.04		-69.10	

Notes: *Param* represents the estimates of the parameters in the semi-log hourly wage equation. The selection equation is presented in Table A4, in Appendix. **Bolds** indicate parameters significant at less than 5 % -level. *Italics* indicate dummy variables. [light green box] indicates that the variable was not included in the model due to few or no observations.

In sum, excepting the sample of people with recent long-term sickness, for which a few health status variables had significant effects on the hourly wages, very few other significant effects were found. This confirms our hypothesis that recent long-term sickness in a previous period can decrease both (current) annual earnings, and (current) hourly wages.

8 Conclusions

The data for annual earnings and the hourly wage provide evidence that both are affected by a history of sickness. The earnings profiles confirm the usual paradigm of a flat concave profile, first increasing with, then tapering off, and eventually declining with age. Age-wage profiles are much flatter than earnings profiles, which is typical for the Swedish labor market. Nevertheless, there is a premium for education, according to the results. High education enhances both earnings and the hourly wage rate, suggesting that the effect goes mainly through the wage rate.

The results answer the central question of this study, what is the effects of the sickness on earnings. Persons with a history of long-term sickness have lower earnings than those without. In the multiple regression analysis, previous history of long-term sickness has a negative effect on earnings when estimated for both genders together and for men when men and women are estimated separately. Days of sickness per year entered as a separate variable have a more ambiguous effect, however.

There are also clearly observable differences between earnings profiles for different diagnosis categories, where people with musculoskeletal sickness histories have higher profiles, and people with respiratory problems notably lower profiles. The multivariate analysis did not reveal significant effects through the diagnosis when men and women were aggregated, but did reveal that men with musculosketal problems had on average a higher wage and men with general symptom diagnoses had a lower wage, whereas there was a negative effect on earnings for women for mental diagnoses and injuries. The latter was also significant in the separate equation for persons with long-term sickness.

Multivariate analysis indicates that there were only few and weak effects of sickness history only on the wage rate of those with poorer health. The results also indicate that being “a foreigner” had the most significant negative effect on the wage rate of those with long-term sickness history, but had not significant effect on their annual earnings.

People who previously had a spell of long-term sickness had lower earnings in following years, even if they did not experience a new spell of long-term sickness. The conclusion of this study is that since the effect cannot be attributed to an effect on the wage rate, it has to have resulted from a reduction in time spent working. An implication for the policy is that the work alternative should always be more attractive than the alternative of disability for people who can still work. Therefore, it is desirable to have programs directed to improve the social and physical work environment, and individual performance (through training and/or vocational rehabilitation of those individuals).

References

Albrecht, James W., Per-Anders Edin, Marianne Sundstrom, and Susan B. Vroman (1999), Career interruptions and subsequent earnings: a reexamination using Swedish data, Journal of Human Resources 34, 294-311.

Baldwin, M., and W. Johnson (1994), Labor market discrimination against men with disabilities, Journal of Human Resources 29, 1-19.

Bartel, Ann, and Paul Taubman (1979), Paul Health and Labor Market Success: The Role of Various Diseases, Review of Economics and Statistics 61, 1-8.

Becker, Gary S. (1962), Investment in human capital: A theoretical analysis. Journal of Political Economy 70(5), Part 2: Investment in Human Beings, 9-49.

Becker, Gary S. (1964), Human Capital (Columbia University Press for the National Bureau of Economic Research, New York).

Becker, Gary S. (1965), A theory of the allocation of time, Economic Journal 75, 493-517.

Becker, Gary S. (1967), Human Capital and the Personal Distribution of Income: An Analytical Approach (University of Michigan, Ann Arbor, Michigan). Also available in: G.S. Becker (1993), Human Capital, Third Edition (University of Chicago Press), 102-158.

Ben-Porath, Yoram (1967), The production of human capital and the life cycle of earnings, Journal of Political Economy 75, 353-367.

Berkovec, J., and S. Stern (1991), Job exit behavior of older men, Econometrica 59, 189-210.

Chirikos, Thomas N., and Gilbert Nestel (1985), Further Evidence on the Economic Effects of Poor Health, Review of Economics and Statistics 67, 61-69.

Currie, Janet, and Brigitte C. Madrian (1999), Health, Health Insurance And The Labor Market, in Handbook of Labor Economics, Volume 3, Edited by O. Ashenfelter and D. Card, Elsevier Science B.V, 3309-415.

Edin, Per-Anders, and Jenny Nynabb (1992), Gender Wage Differentials and Interrupted Work Careers: Swedish Evidence, Working paper 1992:17, Department of Economics, Uppsala University.

Fase, M.M.G. (1970), An econometric model of age-income profiles: A statistical analysis of Dutch income data 1958-1967, Rotterdam.

Fuchs, V.R. (1966), The contribution of health services to the American economy, Milbank Memorial Fund Quarterly 44, 65-102.

Grossman, Michael (1972a), On the concept of health capital and the demand for health, Journal of Political Economy 80, 223-255.

Grossman, Michael (1972b), The Demand for Health: A Theoretical and Empirical Investigation (Columbia University Press for the National Bureau of Economic Research, New York).

Grossman, Michael (1999), The Human Capital Model of the Demand for Health, NBER Working Paper No. W7078.

Grossman, Michael, and Lee Benham (1973), Health, hours and wages, in M. Perlman, ed., The economics of health and medical care (Wiley, NewYork).

Gustafsson, Siv (1981), Male-Female Lifetime Earnings Differentials, in Studies in Labor Market Behavior: Sweden and the United States, ed. Gunnar Eliasson, Bertil Holmlund, and Frank Stafford, 235-68. Stockholm: Industrial Institute for Social and Economic Research.

Hansen, Jörgen, (2000), The effect of work absence on wages and wage gaps in Sweden, Journal of Population Economics 13(1), 45-55.

Haveman, Robert and Barbara Wolfe (1990), The Economic Well-Being of the Disabled: 1962-84, Journal of Human Resources 25, 32-54.

Haveman, Robert, Barbara Wolfe, Brent Kreider, and Mark Stone (1994), Market work, wages, and men's health, Journal of Health Economics 13, 163-82.

Hause, J.C. (1972), Earnings profile: Ability and schooling, Journal of Political Economy 80, Part 2: Investment in Education: The Equity-Efficiency Quandary, S108-S138.

Heckman, James J., 1979, Sample Selection Bias as a Specification Error, Econometrica 47, 153-161.

Johnson, W.G., and J. Lambrinos (1985), Wage discrimination against handicapped men and women, Journal of Human Resources 20, 264-77.

Klevmarken, Anders (1972), Statistical methods for the analysis of earnings data: With special application to salaries in Swedish industry, IUI, Almqvist & Wiksell, Stockholm.

Klevmarken, Anders (1992), On ageing and earnings, Memo 171, Department of Economics, University of Göteborg.

Lambrinos, James (1981), Health: A Source of Bias in Labor Supply Models, Review of Economics and Statistics 63, 206-12.

Lee, Lung-Fei (1982), Health and wages: a simultaneous equation model with multiple discrete indicators, International Economic Review 23, 199-221.

Luft, Harold S. (1975), The Impact of Poor Health on Earnings, Review of Economics and Statistics 57, 43-57.

Mincer, Jacob (1958), Investment in Human Capital and personal Income Distribution, Journal of Political Economy 66, 281-302.

Mincer, Jacob (1962), On-the-job training: Costs, returns, and some implications, Journal of Political Economy 70(5), Part 2: Investment in Human Beings, 50-79.

Mincer, Jacob (1970), The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach, Journal of Economic Literature 8, 1-26.

Mincer, Jacob (1974), Schooling, experience, and earnings, New York.

Mitchell, J.M., and R. Burkhauser (1990), Disentangling the effect of arthritis on earnings: a simultaneous estimate of wage rates and hours worked, Applied Economics 22, 1291-1310.

Mullahy, J. and J. Sindelar (1991), Gender differences in labor market effects of alcoholism", The American Economic Review 81, 161-5.

Mullahy, J. and J. Sindelar (1993), Alcoholism, work and income, Journal of Labor Economics 11, 494-520.

Mullahy, J. and J. Sindelar (1995), Health, income and risk aversion: assessing some welfare costs of alcoholism and poor health, Journal of Human Resources 30, 439-59.

Muller, L. Scott; Scott, Charles G., and Barry V. Bye (1996), Labor-Force Participation and Earnings of SSI Disability Recipients: A Pooled Cross-Sectional Times Series Approach to the Behavior of Individuals, Social Security Bulletin 59, 22-42.

Mushkin, S.J. (1962), Health as an investment, Journal of Political Economy 70(5), Part 2: Investment in Human Beings, 129-157.

Polachek, Solomon W. and W. Stanley Siebert (1993), The economics of earnings, Cambridge; New York and Melbourne: Cambridge University Press.

Puhani, P. (2000), The Heckman Correction for Sample Selection and Its Critique – A Short Survey, Journal of Economic Surveys 14, 53-68.

Schultz Theodore W. (1960), Capital Formation by Education, Journal of Political Economy 68, 571-83.

Schultz, Theodore W. (1961), Investment in human capital, American Economic Review 51, 1-17.

Skogman Thoursie, Peter (1999) Disability and work in Sweden, Swedish Institute for Social Research 39, Doctoral Dissertation, Stockholm University.

Smith, James P. (1998), Socioeconomic Status and Health, American Economic Review 88, 192-196.

Smith, James P. (1999), Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status, Journal of Economic Perspectives 13, 145-166.

Stafford, Frank P., and Marianne Sundstrom (1996), Time Out for Childcare: Signalling and Earnings Rebound Effects for Men and Women, Labour 10, 609-29.

Sundberg, Gun (1996), "Health, Work-Hours, and Wages in Sweden" in Essays on Health Economics, Doctoral dissertation, Department of economics, Uppsala University.

Tobin, James (1958), Estimation of Relationships for Limited Dependent Variables, Econometrica 26, 24-36.

Thomas, Duncan and John Strauss (1997), Health and wages: Evidence on men and women in urban Brazil, Journal of Econometrics 77, 159-185.

Willis, Robert (1986), "Wage determinants: a survey and reinterpretation of human capital earnings function" in Handbook of Labor Economics I, Edited by O. Ashenfelter and R. Layard, Elsevier Science Publishers BV.

Appendix

Table A1 Descriptive statistics, by sickness history, in 1988

Variable	No LT sick 1983-1988 (N=1503)		LT sick 1983-1988 (N=185)		Sickness cohorts (SC) SC1983 (N=37)		SC1984 (N=48)		SC1985 (N=42)		SC1986 (N=31)		SC1987 (N=27)	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std
Gender: *Woman*=1, Men=0	0.49	0.50	0.55	0.50	0.62	0.49	0.46	0.50	0.55	0.50	0.65	0.49	0.52	0.51
Age	40.54	11.09	44.28	12.16	44.59	10.80	44.44	12.93	42.21	12.40	44.26	12.38	46.81	12.27
Age-groups														
16-30 *years*	0.23	0.42	0.16	0.37	0.08	0.28	0.19	0.39	0.24	0.43	0.16	0.37	0.11	0.32
31-45 *years*	0.43	0.50	0.36	0.48	0.46	0.51	0.31	0.47	0.36	0.49	0.39	0.50	0.30	0.47
46-50 *years*	0.12	0.32	0.09	0.29	0.11	0.32	0.06	0.25	0.12	0.33	0.10	0.30	0.07	0.27
51-55 *years*	0.10	0.31	0.12	0.33	0.14	0.35	0.13	0.33	0.10	0.30	0.10	0.30	0.19	0.40
56-65 *years*	0.12	0.32	0.26	0.44	0.22	0.42	0.31	0.47	0.19	0.40	0.26	0.45	0.33	0.48
Citizenship														
Swedish born	0.88	0.33	0.90	0.30	0.84	0.37	0.92	0.28	0.86	0.35	0.94	0.25	0.96	0.19
Foreign born	0.08	0.27	0.04	0.19	0.08	0.28	0.02	0.14	0.05	0.22	0.03	0.18	0.00	0.00
Nationalized	0.05	0.21	0.07	0.25	0.08	0.28	0.06	0.25	0.10	0.30	0.03	0.18	0.04	0.19
Education														
Low	0.47	0.50	0.65	0.48	0.81	0.40	0.73	0.45	0.60	0.50	0.52	0.51	0.52	0.51
Medium	0.36	0.48	0.26	0.44	0.11	0.32	0.23	0.43	0.31	0.47	0.26	0.45	0.44	0.51
High	0.17	0.38	0.09	0.29	0.08	0.28	0.04	0.20	0.10	0.30	0.23	0.43	0.04	0.19
Marital status														
Unmarried	0.38	0.49	0.31	0.47	0.43	0.50	0.44	0.50	0.26	0.45	0.13	0.34	0.22	0.42
Married	0.54	0.50	0.58	0.50	0.41	0.50	0.52	0.51	0.64	0.49	0.77	0.43	0.59	0.50
Divorced	0.07	0.25	0.09	0.29	0.14	0.35	0.02	0.14	0.07	0.26	0.10	0.30	0.19	0.40
Widowed	0.01	0.10	0.02	0.13	0.03	0.16	0.02	0.14	0.02	0.15	0.00	0.00	0.00	0.00
Earnings (Th. SEK)	120.70	70.60	98.70	70.60	86.20	64.60	82.30	55.70	98.70	62.60	128.90	106.10	110.10	53.00
Yearly hours work	1478.8	838.4	1124.6	923.6	789.3	941.7	1014.3	920.8	1277.6	891.2	1293.3	872.6	1348.1	916.5

Variable	No LT sick 1983-1988 (N=1503) Mean	Std Dev	LT sick 1983-1988 (N=185) Mean	Std Dev	Sickness cohorts (SC) SC1983 (N=37) Mean	Std Dev	SC1984 (N=48) Mean	Std Dev	SC1985 (N=42) Mean	Std Dev	SC1986 (N=31) Mean	Std Dev	SC1987 (N=27) Mean	Std
Hourly wages	62.17	31.50	52.50	33.28	42.92	27.85	47.13	27.27	48.48	18.63	74.74	55.13	55.89	25.53
Compensated days of sickness, by year														
1983	5.92	11.14	22.48	36.66	65.27	54.52	16.79	24.84	6.76	9.67	12.23	19.47	10.15	18.69
1984	5.99	12.62	46.65	84.42	104.27	151.08	74.52	60.67	15.36	24.02	12.26	16.64	6.30	10.12
1985	6.76	13.17	56.94	116.21	93.92	217.10	80.73	96.86	53.36	57.08	17.84	29.70	14.41	22.17
1986	7.16	17.02	77.84	104.57	57.92	100.13	72.23	107.67	110.69	131.93	120.94	72.24	14.56	18.48
1987	7.53	14.72	56.44	83.94	33.38	67.02	37.10	81.63	49.40	78.92	81.68	109.31	104.41	57.53
1988	7.73	11.61	13.77	16.27	8.86	13.34	12.96	15.97	17.12	17.71	15.52	16.95	14.74	16.77
1989	12.35	34.63	37.94	86.20	41.03	98.17	27.06	73.82	32.26	71.80	46.81	93.34	51.70	103.14
1990	19.93	55.26	51.89	94.16	45.49	95.42	62.71	109.08	50.67	85.24	55.35	94.80	39.33	79.69
1991	17.85	48.19	36.78	83.21	33.62	80.84	37.50	80.71	36.31	80.22	46.87	104.85	28.96	71.66
Sickness days (1986-88), by diagnosis														
Musculoskeletal	5.43	25.75	90.48	238.15	92.76	335.41	83.90	244.87	128.95	246.25	89.68	163.39	40.15	75.24
Cardiovascular	1.16	25.75	25.03	176.11	0.89	3.78	58.44	319.66	24.17	111.00	8.84	48.11	18.67	96.99
Respiratory	8.73	14.20	20.31	105.41	20.70	64.99	11.67	18.60	8.36	13.11	13.55	16.86	61.52	263.40
Mental	1.02	8.05	30.47	152.66	77.76	298.15	4.73	19.80	37.12	113.08	13.03	71.27	21.11	100.89
General Symptoms	2.19	6.62	10.29	56.29	3.49	8.40	4.31	15.43	2.93	8.09	29.32	122.74	19.85	60.54
Injuries	2.92	25.21	23.18	87.20	3.11	8.28	18.63	110.08	24.86	88.41	44.58	117.29	31.59	48.73
Others	6.39	21.52	34.28	82.33	15.03	34.83	30.60	93.73	26.88	74.36	63.29	114.68	45.41	68.83
Spells of sickness (1986-88), by diagnosis														
Musculoskeletal	0.441	1.141	0.978	1.722	0.973	2.327	0.604	0.893	1.405	2.061	0.935	1.181	1.037	1.786
Cardiovascular	0.043	0.260	0.065	0.323	0.054	0.229	0.063	0.320	0.048	0.216	0.065	0.250	0.111	0.577
Respiratory	1.866	2.586	2.059	3.015	2.270	4.107	2.188	3.589	1.548	1.941	2.097	2.119	2.296	2.447
Mental	0.053	0.349	0.205	0.753	0.351	1.086	0.083	0.347	0.357	1.055	0.065	0.250	0.148	0.456
General symptoms	0.444	1.036	0.649	1.303	0.784	2.043	0.479	0.850	0.500	1.153	0.806	1.195	0.815	1.001
Injuries	0.195	0.579	0.459	0.840	0.216	0.479	0.375	0.815	0.476	0.740	0.645	1.018	0.704	1.103
Other	0.915	1.698	1.465	2.162	1.270	1.981	1.208	2.021	1.405	1.809	2.194	3.114	1.444	1.761

Table A2 OLS estimated parameters of the (log) earnings equation in 1988 considering the sickness history 1983-1988

Variable	All Insured			Men			Women			Not long-term Sick 1983-88			Long-term Sick 1983-1987		
	Param.	Robust Std. Err.	Marginal Effect [a]	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect
Female	**-0.54**	0.14	-41.58							**-0.58**	0.14	-44.25	-0.18	0.61	-16.32
Age	**0.39**	0.06	-2.08	**0.38**	0.08	-2.33	**0.40**	0.08	-1.76	**0.41**	0.06	-1.56	0.18	0.19	-4.83
Age-Squared/100	**-0.49**	0.07		**-0.48**	0.10		**-0.52**	0.10		**-0.52**	0.07		-0.25	0.23	
Citizenship															
Foreign	**-2.06**	0.38	-87.22	**-2.26**	0.64	-89.57	**-1.95**	0.47	-85.74	**-2.08**	0.40	-87.51	-0.42	1.33	-34.33
Nationalized	-0.57	0.39	-43.67	-0.86	0.58	-57.87	-0.46	0.49	-37.00	**-0.79**	0.42	-54.75	0.89	1.14	143.71
Education															
Medium	**0.88**	0.14	139.91	**0.93**	0.19	154.19	**0.78**	0.20	118.91	**0.82**	0.14	126.87	0.97	0.60	163.37
High	**0.90**	0.17	146.03	0.56	0.30	74.96	**1.11**	0.20	204.68	**0.86**	0.18	137.15	**1.91**	0.69	574.42
Married	0.05	0.11	5.36	**0.36**	0.18	43.46	-0.25	0.15	-21.95	0.08	0.12	8.35	-0.20	0.50	-18.33
Sickness cohorts															
Year 1983	**-1.93**	0.91	-85.53	-2.38	1.78	-90.76	-0.91	0.83	-59.61				-1.41	1.86	-75.63
Year 1984	-1.09	0.70	-66.38	-1.52	0.91	-78.08	-0.86	0.96	-57.84				-0.72	1.07	-51.28
Year 1985	-0.23	0.47	-20.71	0.14	0.53	14.86	0.04	0.69	4.12				-0.33	0.83	-27.90
Year 1986	-0.13	0.43	-12.06	1.34	0.83	280.22	-0.80	0.46	-55.11				0.09	0.90	9.34
Year 1987	-0.50	0.62	-39.52	**-1.78**	0.53	-83.19	0.91	0.74	147.90						
Sickness days (1986-88), by diagnosis															
Musculoskeletal	0.002	0.001	0.173	**0.003**	0.002	0.459	-0.002	0.002	-0.143	**0.003**	0.001	0.327	0.001	0.002	0.134
Cardiovascular	0.001	0.001	0.084	0.001	0.001	0.123	-0.001	0.001	-0.085	**0.001**	0.001	0.122	0.001	0.002	0.047
Respiratory	0.002	0.002	0.284	**0.002**	0.001	0.220	**0.007**	0.004	0.685	**0.012**	0.004	1.465	0.001	0.001	0.053
Mental	-0.001	0.002	-0.082	0.001	0.002	0.192	**-0.007**	0.003	-0.642	-0.006	0.004	-0.747	0.000	0.002	-0.028
Gen. Symptoms	0.000	0.005	-0.020	**-0.018**	0.008	-2.509	**0.004**	0.001	0.351	0.005	0.005	0.654	-0.003	0.006	-0.269
Injuries	-0.004	0.004	-0.506	0.000	0.002	-0.007	**-0.013**	0.005	-1.217	**0.003**	0.001	0.305	**-0.010**	0.004	-0.954
Others	0.000	0.002	0.038	0.003	0.003	0.393	-0.002	0.003	-0.193	0.002	0.001	0.202	-0.001	0.003	-0.107
Compensated days of sickness, by year															
1983	**0.019**	0.006	2.194	0.005	0.009	0.751	**0.017**	0.006	1.606	**0.011**	0.004	1.270	0.014	0.014	1.363
1984	-0.005	0.004	-0.535	0.008	0.004	1.082	**-0.010**	0.004	-0.974	0.003	0.004	0.342	-0.008	0.006	-0.795
1985	**-0.006**	0.003	-0.732	**-0.009**	0.002	-1.329	-0.003	0.005	-0.279	0.002	0.005	0.258	**-0.007**	0.003	-0.705
1986	-0.003	0.003	-0.298	**-0.015**	0.005	-2.155	**0.006**	0.003	0.607	0.000	0.002	0.048	-0.002	0.005	-0.173
1987	0.003	0.004	0.300	**0.015**	0.005	2.127	0.000	0.003	0.040	0.004	0.003	0.483	0.000	0.004	0.006
1988	**0.048**	0.005	5.626	**0.047**	0.007	6.719	**0.046**	0.006	4.294	**0.034**	0.004	4.074	0.061	0.015	6.046
Intercept	**3.443**	1.083		**3.438**	1.542		2.778	1.553		**2.959**	1.157		7.600	3.852	
N	1688			849			839			1503			185		
R-squared	0.20			0.23			0.24			0.18			0.32		
Root MSE	2.73			2.74			2.65			2.60			3.61		

Notes: The marginal effects show the effect of a one-unit change of a continuous variable X_j on the annual earnings (in thousand Swedish crowns). **Bolds** indicate parameters significant at the 5 % level. *Italics* show dummy variables, for which the marginal effect is the percentage change on earnings when category “1” is compared to “0”.

Table A3 OLS Estimated parameters of the (log) wage equation in 1988 (same *Note* as for Table A2)

	All Insured			Men			Women			Not long-term Sick 1983-88			Long-term Sick 1983-1987		
Variable	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect	Param.	Robust Std. Err.	Marginal Effect
Female	**-0.21**	0.02	-19.11							**-0.22**	0.02	-19.96	**-0.14**	0.06	-13.19
Age	**0.04**	0.01	0.26	**0.04**	0.01	0.33	**0.04**	0.01	0.16	**0.04**	0.01	0.34	0.04	0.02	
Age-Squared/100	**-0.05**	0.01		**-0.05**	0.01		**-0.04**	0.01		**-0.05**	0.01		-0.04	0.02	
Citizenship															
Foreign	**-0.14**	0.03	-13.00	**-0.09**	0.05	-8.63	**-0.16**	0.04	-15.20	**-0.12**	0.03	-11.30	-0.25	0.13	-21.74
Nationalized	0.01	0.05	1.19	0.03	0.07	2.60	-0.02	0.06	-1.57	0.04	0.05	4.02	-0.13	0.08	-12.00
Education															
Medium	**0.12**	0.02	13.04	**0.21**	0.03	23.45	0.03	0.03	2.88	**0.13**	0.02	14.33	0.02	0.06	2.01
High	**0.25**	0.03	28.03	**0.33**	0.05	39.14	**0.17**	0.03	19.07	**0.26**	0.03	30.22	0.09	0.14	9.17
Married	0.03	0.02	3.08	**0.09**	0.03	9.03	-0.01	0.02	-1.19	0.03	0.02	3.01	0.00	0.06	0.09
Cohorts of long-term sickness															
Year 1983	0.05	0.08	5.49	0.05	0.13	5.09	0.01	0.10	1.31				0.04	0.15	
Year 1984	0.05	0.05	5.38	0.00	0.07	0.14	0.10	0.08	11.02				0.02	0.09	
Year 1985	**-0.09**	0.04	-9.00	-0.11	0.08	-10.66	-0.06	0.06	-6.22				-0.13	0.09	
Year 1986	0.17	0.09	18.70	0.16	0.19	17.34	0.17	0.10	18.79				0.16	0.13	
Year 1987	-0.04	0.06	-3.62	-0.09	0.09	-8.29	0.02	0.08	2.43						
Sickness days (1986-88), by diagnosis															
Musculoskeletal	**-0.0002**	0.0001	-0.014	**-0.0005**	0.0002	-0.034	-0.0001	0.0001	-0.007	-0.0003	0.0002	-0.021	-0.0002	0.0001	-0.010
Cardiovascular	-0.0001	0.0001	-0.007	-0.0001	0.0002	-0.005	-0.0001	0.0001	-0.003	0.0000	0.0001	0.000	-0.0001	0.0002	-0.004
Respiratory	**0.0004**	0.0001	0.025	**0.0004**	0.0001	0.027	0.0009	0.0005	0.046	0.0008	0.0006	0.051	0.0003	0.0001	0.018
Mental	-0.0001	0.0001	-0.005	-0.0001	0.0001	-0.004	0.0000	0.0004	0.000	-0.0014	0.0007	-0.090	**0.0000**	0.0002	0.002
Gen. Symptoms	0.0008	0.0004	0.048	0.0026	0.0018	0.176	**0.0008**	0.0004	0.042	0.0026	0.0019	0.163	0.0007	0.0004	0.037
Injuries	0.0000	0.0002	-0.003	0.0002	0.0001	0.016	-0.0005	0.0003	-0.027	0.0001	0.0002	0.006	**-0.0005**	0.0004	-0.026
Others	-0.0002	0.0002	-0.012	**-0.0010**	0.0003	-0.065	0.0002	0.0002	0.009	-0.0003	0.0003	-0.019	-0.0004	0.0003	-0.020
Compensated days of sickness, by year															
1983	-0.0004	0.0006	-0.025	-0.0018	0.0011	-0.122	0.0001	0.0007	0.006	-0.0001	0.0009	-0.008	-0.0010	0.0012	-0.053
1984	-0.0002	0.0003	-0.012	-0.0003	0.0005	-0.018	0.0000	0.0005	-0.001	-0.0003	0.0008	-0.020	-0.0004	0.0004	-0.018
1985	-0.0004	0.0002	-0.022	0.0004	0.0005	0.026	**-0.0005**	0.0002	-0.026	-0.0003	0.0007	-0.019	**-0.0005**	0.0002	-0.024
1986	-0.0003	0.0003	-0.015	-0.0001	0.0008	-0.006	-0.0003	0.0004	-0.015	-0.0007	0.0004	-0.041	0.0001	0.0006	0.004
1987	-0.0002	0.0004	-0.010	0.0002	0.0010	0.016	-0.0005	0.0003	-0.029	0.0006	0.0007	0.038	-0.0006	0.0004	-0.030
1988	**0.0017**	0.0008	0.102	0.0007	0.0012	0.047	0.0020	0.0011	0.109	0.0011	0.0009	0.071	0.0021	0.0017	0.110
Intercept	**3.24**	0.12		**3.09**	0.19		**3.16**	0.15		**3.17**	0.13		**3.56**	0.38	
n	1571			790			781			1412			159		
R-squared	0.19			0.16			0.13			0.19			0.32		
Root Mean Sq. Error	0.34			0.37			0.31			0.35			0.31		

Table A4 Estimated parameters of the selection equation

Variables	All Insured		Men		Women		Not long-term Sick 1983-88		Long-term Sick 1983-88	
	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.	Param.	Std.Err.
Female	-0.11	0.12					-0.14	0.13	0.24	0.38
Age	0.21	0.06	0.27	0.10	0.23	0.08	0.24	0.06	0.07	0.19
Age-Squared/100	-0.28	0.07	-0.37	0.14	-0.33	0.11	-0.32	0.08	-0.12	0.23
Citizenship										
Foreign	-0.90	0.18	-0.75	0.30	-1.00	0.28	-1.01	0.19	-0.44	0.97
Nationalized	-0.20	0.26	-0.45	0.39	-0.45	0.50	-0.36	0.28	0.77	0.84
Education										
Medium	0.66	0.17	1.09	0.32	0.63	0.26	0.64	0.18	0.69	0.64
High	0.71	0.20	0.46	0.28	1.23	0.39	0.60	0.21		
Married	0.20	0.20	1.01	0.55	0.19	0.34	0.19	0.20	0.22	0.56
Sickness Cohorts										
Year 1983	-0.84	0.51	-2.97	1.15	3.68	1.70			-0.64	0.86
Year 1984	-0.98	0.38	-2.90	1.05	-0.24	0.81			-0.91	0.75
Year 1985	0.07	0.53	0.60	1.45	1.33	1.10			-0.32	0.80
Year 1986	-0.35	0.54			-1.53	0.95				
Year 1987	0.31	0.61	-8.87	4.62	2.40	1.23			-0.29	0.87
Sickness Days (1986-88), by diagnosis										
Musculoskeletal	0.002	0.001	0.004	0.005	-0.007	0.002	0.016	0.014	0.000	0.001
Cardiovascular	0.002	0.002			-0.004	0.001				
Respiratory	0.056	0.013	0.023	0.025	0.072	0.022	0.059	0.015	0.073	0.035
Mental	0.000	0.001	-0.148	0.098	-0.014	0.007	0.016	0.021	0.000	0.001
Gen. Symptoms	-0.002	0.002	0.021	0.018	0.112	0.059	0.084	0.047	-0.002	0.002
Injuries	-0.003	0.001	0.012	0.006	-0.018	0.004	0.041	0.022	-0.004	0.002
Others	0.002	0.003	0.026	0.015	-0.008	0.003	0.072	0.028	-0.001	0.002
Compensated days of sickness, by year										
1983	0.013	0.005	0.013	0.012	0.059	0.022	0.031	0.015	0.004	0.006
1984	-0.002	0.002	0.014	0.007	-0.013	0.005	-0.003	0.010	-0.004	0.002
1985	-0.001	0.001	-0.003	0.003	-0.004	0.003	0.010	0.008	-0.002	0.001
1986	-0.002	0.002	-0.031	0.014	0.019	0.007	-0.012	0.006	0.000	0.002
1987	-0.002	0.002	0.131	0.059	-0.002	0.003	0.010	0.014	0.001	0.002
1988	0.106	0.018	22.426	19.545	0.160	0.040			0.040	0.018
"*50 plus*"	0.02	0.31	-0.10	0.60	0.35	0.44	0.12	0.32	0.17	0.93
Intercept	-2.77	1.04	-4.26	1.80	-3.05	1.62	-3.13	1.10	0.43	3.83
Censored obs.	117		59		58		91		26	
Uncensored obs.	1571		790		781		1412		159	
Wald chi2	331.52		140.44		97.10		293.49		62.81	
Log likelihood	-797.64		-421.63		-274.89		-721.04		-69.10	

Notes: *Param* represents the estimates of the parameters in the semi-log hourly wage equation.