Socioeconomic Heterogeneity in the Effect of Health Shocks on Earnings. Evidence from Population-Wide Data on Swedish Workers

P Lundborg, M Nilsson & J Vikström

20100906

1 Introduction

There is a strong socioeconomic gradient in health. It has been widely documented that people of higher socioeconomic position enjoy better health and this pattern seems to hold up irrespective of what health measures that are used (see e.g.van Doorslaer, Wagstaff, Bleichrodt, Calonge, Gerdtham, Gerfin, Geurts, Gross, Häkkinen, Leu, O'Donell, Propper, Puffer, Rodríguez, Sundberg, and Winkelhake (1997); Marmot (1999); Smith (1998); Mackenbach and Bakker (2002)). It is also believed that socio-economic inequalities in health have increased during the recent decades in most Western countries Mackenbach, Bos, Andersen, Cardano, Costa, Harding, Reid, Hemstrom, Valkonen, and Kunst (2003). The socioeconomic gradient in health is often a cause of social concern, since it is a common belief that an individual's health should not be determined by his or her economic or social position CSDH (2008).

While there is a general agreement about the existence of a socioeconomics gradient in health, there is surprisingly little agreement about its underlying causes. In the epidemiological literature, it has traditionally been assumed that differences in socioeconomic status causes differences in health. Lately, economists have instead explored the hypothesis that health outcomes influence socioeconomic status (e.g. Smith (1999)). This is a very different explanation than the traditional one, since it suggests that health outcomes are the mechanism through which the socioeconomic gradient partly arises. Such an explanation could also explain the commonly found empirical pattern that the socioeconomic gradient in health widens during the middle ages, since health shocks are more common in lower socioeconomic groups at older ages. If health shocks are partly causing the widening in socioeconomic status, the policy implications are obviously very different from the policy implications based on the traditional viewpoint.

In this paper, we contribute to the small but growing literature on the effects of health shocks on labor market outcomes. This literature has for the most part focused on average effects of health shocks on labor market outcomes. We believe that this is restrictive for a number of reasons. First, the impact of a given health shock may differ between socioeconomic groups. If people of lower socioeconomic status experience more severe effects of a given health shock, they would face a double penalty, since they already face an increased risk of experiencing various health shocks. It is not farfetched to assume that the ability to cope with health shocks partly depends on factors such as education and income. Such heterogeneity in the effects among adults has not, to our knowledge, been thoroughly explored in the literature so far. Yet, such heterogeneity could help to explain the widely documented increase in the socioeconomic gap over the lifecycle.

In this paper, we attempt to improve the understanding of heterogeneity in the impact of health shocks on labor market outcomes. For this purpose, we use large-scale and register-based data on the entire population of workers in Sweden and apply panel fixed-effects techniques in order to answer two related questions: Are there heterogeneous effects of given health shocks across socioeconomic groups? Does the heterogeneity of the effects change across the life-cycle?

Our data and analytical approach gives us a number of advantages. First, our huge sample, covering in total 1.2 million individuals, allows us to address heterogeneous effects by socioeconomic status and age with large precision. In our analyses, will focus on heterogeneous returns with respect to education, since education is not affected by health shocks during middle ages. Focusing on heterogeneous effects by income would be more problematic, since income in itself may be affected by health shocks. Our large sample also allow us to consider a large number of different types of health shocks, instead of limiting the analyses to a general measure of health shocks or to a specific type of health shock. The latter approach has been common in the literature.

Second, we are able to follow individuals for up to 14 years, which allow us to consider the long-term impact of health shocks. Examining the long-term impact is important, since some health shocks may permanently reduce work capacity. Moreover, there may exist heterogeneity in the long-term impact of health shocks by socioeconomic status, which may be masked by only focusing on short-term effects. It is for instance well known that adherence to medical therapies varies by socioeconomic status, which may also obviously imply that the long-term impact of a given health shock may vary by socioeconomic status (Goldman and Smith (2005)).

Third, the panel structure of the data allow us to employ panel-data fixed effects techniques and thereby account for time-invariant factors at the individual level that may be associated with both underlying health and labor market outcomes, such as chronic conditions, genes and early life environment.

Fourth, our estimates are based on detailed register-based on health shocks taken from the national inpatient hospital registers, whereas most previous literature use self-reported health shocks. Using register-based data is an advantage, since there is substantial evidence of reporting bias by socioeconomic status. People with higher education and income report worse health for a given condition (Etilé and Milcent (2006); d'Uva, Doorslaer, Lindeboom, and O'Donnell (2008)). This is problematic, since it will bias the estimates of the effect of health shocks on labor market outcomes downwards. Moreover, it would hide important heterogeneity in the "true" impact of health shocks on labor market outcomes across socioeconomic groups.

Fifth, our data allow us to distinguish between acute and planned hospitalizations. We are thus able to study the impact of health shocks that were unexpected from the individual's point of view. Together with our fixed-effects specification, focusing on unexpected health shocks facilitates a causal interpretation of our estimates. Finally, our data includes information on medical treatments, which allow us to investigate to what extent heterogeneity in the impact of health shocks is generated through differences in treatment received by socioeconomic status. This issue has, to the best of our knowledge, not been investigated in the literature so far.

We start our paper by documenting large and long-run average effects of health shocks on earnings. As expected, the effects differ by type of health shock and are largest for Mental and Behavioral diagnoses and Heart Diseases. We then show that there exists substantial heterogeneity in the effect of certain given health shocks on earnings. The results are most pronounced for the older age groups 50-59. Lower socioeconomic groups not only face worse outcomes from a given health shock, but also face more long-lasting effects compared to higher socioeconomic groups. These results are consistent with the idea that education is beneficial for health and that educated people are better enabled to cope with health shocks.

Our results are consistent with the idea that the socioeconomic gradient in health is partly caused by the impact of health shocks on socioeconomic status. We contribute to the literature by showing that the gradient arises not only because people of lower socioeconomic status are more likely to experience health shocks but also because the impact of a given health shock is more severe.

We believe that our results are relevant for a number of reasons. First, knowledge about heterogeneous impacts of health shocks may point to the possibility of targeted efforts towards groups who suffer disproportionally. Example of such policies are more intense screening for health markers among socioeconomics risk groups, regular health check-ups, and improving adhesion to treatment. Second, the results provide valuable information for evaluations of the cost effectiveness of various medical interventions designed to prevent or cure disease. In such evaluations, estimates of the value of production losses associated with health events are typically crudely measured through the average wage rate. Such an approach will disguise important difference in the cost-effectiveness across subgroups.

The paper is organized as follows. In Section 2, we review the literature that is connected to our study and in particular the recent literature on the effect of health shocks on socioeconomic outcomes. In Section 3, we describe our data and the constructed variables, while Section 4 and section 5 provides some descriptive patterns on health shocks and labor market outcomes. Section 6 describes our econometric method and the assumptions that underlie our identification strategy. Section 7 presents our results. In Section 8 we run some placebo estimations. Section 9 present estimates of the heterogeneous effects and Section 10 present some sensitivity analysis. Finally, Section 11 concludes.

2 Background

The socioeconomic gradient in health is one of the most widely replicated results in the social sciences. It dates back to at least the 19th century, where researchers have documented marked health differences across different groups in society, such as the royalty, the land-elite, and the working class (see Antonovsky (1967) for a review of the early literature on the socioeconomic gradient in health). The gradient is usually found to widen during working life but then narrows as people reach older ages. The causes of this pattern have been widely debated and there is currently no consensus in the literature. Clearly, however, the major theories in the public health literature all have in common that it is implicitly or explicitly assumed that socioeconomic status causes health and that the effect of health on socioeconomic status is negligible (for a discussion about this see Deaton (2002)).

Economists, typically interested in the determinants of earnings, have recently questioned this standard assumption and instead stressed that health events also have important causal effects on income through its effect on labor market outcomes. Smith (1999), using self-reported data on middle-aged and elderly Americans from the Health and Retirement Survey, found that onset of a new illness reduced household wealth substantially and that a large share of this reduction in wealth was attributable to a decline in labor earnings. In addition, negative health shocks have been found to strongly predict retirement and reduced labor force participation (Smith 1999, 2004, 2005; Case and Deaton (2003)). Surveys by Smith (1999) and Case and Deaton (2005) even conclude that the larger part of the association between health and income at middle and older ages likely reflects an impact of health on income.

There are several reasons to expect a negative effect of health shocks on labor market outcomes. A sudden negative health event, such as a work related injury, myocardial infarction, a stroke, or an accident means less time to allocate between work and leisure and potentially leads to decreased productivity. In addition, the health shock may affect the marginal utility of consumption, such that worse health for instance may decrease the value of consumption activities. There may also be important long-term effects. Some health events are so severe that they permanently reduce the work capacity. However, even less severe events may have long-term effects. The individual may not be able to perform their old work tasks, and frictions in the labor market may prevent a transition into a new occupation. Furthermore, the time out of work may imply that valuable experience and contacts are lost, income is lost, the workplace may have been rationalized, and future employers may be reluctant to hire the individual due to the risk of future health events. For all these reasons, we expect adverse long-term effects of negative health events on income, labor supply and productivity.

As discussed in the introduction, the recent literature has for the most part implicitly assumed that the impact of a health event on labour market outcomes is the same across subgroups of the population. This assumption can be contrasted to extensive evidence from the medical literature, which clearly suggests that there are marked socioeconomic differences in survival from cancers, stroke, coronary heart disease and acute myocardial infarction (Schrijvers and Mackenbach (1994); Smith et al. 1998; Peltonen, Rosen, Lundberg, and Asplund (2000); Tonne et al 2005). In addition, evidence from the psychological literature also suggests that there exists substantial heterogeneity in individuals' responses to sudden changes in life, such as health shocks (see e.g. Davidson (1992); Gross (1998)).

While the evidence cited above concerns survival and health outcomes, it points to the obvious possibility that the labor market consequences of health events also may differ according to socioeconomic status. Income and education may for instance be important resources in coping with adverse health events. In the health economics literature, education is often assumed to make people more productive in their health production, due to the better health knowledge that follow with education Grossman (1972). It would then follow that educated people could be assumed to better cope with adverse health shocks. Evidence consistent with this is reported in Goldman and Smith (2002), for instance, where educated people were found to better adhere to medical treatments for AIDS and diabetes, which are known to be quite demanding. In our analyses, we will therefore allow the effects of health shocks to vary be educational status.

In addition to focusing on average effects, most of the economic studies on the impact of health events treat them as exogenously given (see e.g. Currie, Madrian, Ashenfelter, and Card (1999)). A small number of recent studies have however addressed the endogeneity of health events and provided evidence suggesting a causal effect of health events on labour outcomes (see e.g. Riphahn (1999); Au, Crossley, and Schellhorn (2005) ; Disney, Emmerson, and Wakefield (2006); Gómez and Nicolás (2006)). The evidence about the causal impact of health shocks is still very limited, however, and in an extensive survey of the economic literature on the effect of health events on labour market outcomes, Currie, Madrian, Ashenfelter, and Card (1999) concluded that: "... there is no consensus about the magnitude of the effects or about their size relative to the effects of other variables." Our paper therefore contributes to this literature and provides new evidence using unique data at the population level.

It should also be noted that most previous studies have focused on the impact of one particular health shock at a time, (see e.g. Dano (2005) on the effects of accidents in Denmark). Other studies have used some general measure of health (see e.g. Stewart, 2001). The limitation with both these approaches is that they prevent assessing the relative importance of various types of health shocks in a given population. It is also apparent from the literature that the estimates of the effects of health events on labour market outcomes are quite sensitive to the measures used. This is not very surprising since the impact of a specific health shock may not alter one's productivity in certain occupations. It is therefore important to be able to include several measures of health events in the analyses. We will assess the effects of different diagnosis groups, which is based on the internationally standardized classification of diseases. The latter type of information is also important for policy purposes, since it may suggest for which type of health shocks interventions has the greatest potential of preventing adverse labour market consequences.

3 Data

The data used in this study comes from two different databases. The population register database Louise from Statistics Sweden provides us with information on the entire Swedish population aged 16-64. It includes a rich set of socioeconomic variables (e.g. age, sex, income, immigration status, marital status and employment status). We use yearly data for the period 1990-2004. We restrict our sample to individuals aged 30 - 59 when they suffer (or potentially could suffer) a health shock. The reason for this is that many of those younger than 30 have not yet finished their education and entered the labor market, and many of those older than 59 are about to retire from the labor market which prohibit an analysis of the long-term labor market outcomes.

The Swedish National Patient Register (NPR) is used to identify health shocks. The register provides yearly information on all in-patient care in Sweden from 1987 and onwards. The register includes patient information such as personal registration number and age, and administrative data including date of admission, acute/planned admission and length of stay as well as rich medical data including main and secondary diagnosis (through the International Classification of Diseases, ICD) and medical procedures.

We create our treatment group by selecting admissions from the NPR for the period 1992 - 2000. The reason for this is that we want to have information a couple of years before as well as a couple of years after the health shock. We restrict our analysis to acute admissions, since we wish to focus on health shocks that are unexpected from the individual's point of view. For some diseases there are long queues, which mean that some planned admissions are anticipated several years before the actual admission. For individuals with more than one admission during this time period, only the first one is used in the analysis, treating the following admissions as a result of the first. In total this gives us about 800,000 health shocks over a period of nine years.

We use the international standard and classify all these admissions into

20 major types of diseases. Of these we choose to focus on 10 most common (in terms of incidence)[1]. It leaves us with infectious diseases (ICD-9: 0010-139, ICD-10: A00-B99), cancer (ICD-9: 140-239, ICD-10: C00-D48), mental and behavioral problems (ICD-9: 290-319, ICD-10: F00-F99), diseases in the nerve system (and ICD-9: 320-359, ICD-10: G00-G99), respiratory diseases (ICD-9: 460-519, ICD-10: J00-J99), heart diseases (ICD-9: 390-459, ICD-10: I00-I99), diseases of the digestive organs (ICD-9: 520-579, ICD-10: K00-K93), disease of the musculoskeletal system and connective tissues (ICD-9: 710-739, ICD-10: M00-M99), diseases of the genitourinary system (ICD9: 580-629, ICD-10: N00-N99) and external accidents (ICD-9: 800-1000, ICD-10: S00-T98).

Our control group includes all individuals that *potentially* could suffer a negative health shock in a given year. In order to keep the empirical analysis manageable from a computational point of view we randomly sample 3 percent of the controls each given year. We follow the treated and non-treated over a long time period. It means that some of the individuals in the control group will suffer a health shock within our observation window, and then become treated instead of non-treated. This problem of dynamic treatment assignment is analyzed in Fredriksson and Johansson (2008). They show that the controls should be included in the sample up until the point they become treated. The controls that experience a health shock later in the sample window are therefore included as controls in the sample up until the year they are affected by the shock. This leaves us with approximately 750,000 controls divided over nine years.

In this paper we focus on labor supply effects of a health shock. For that reason we wish to exclude individuals that never are a part of the labor force. In our main specification we therefore only include individuals who participated in the labor force two years prior to the potential shock year. We take two years, since we do not want to make the restriction too far in the past, and since we do not wish to make the restriction shortly before the shock, because that may hide important pre-treatment trends. We have performed robustness analysis with respect to this restriction, and our results are quite insensitive to instead making the restriction three or four years before the shock year. Labor force participation is defined using yearly labor earnings. We define labor force participation as having a yearly labor income larger than one Price Basic Amount [2]. This restriction reduces the number of treated to approximately 630,000 and the number of controls to 600,000.

[1]with the exception of admission related to pregnancies

[2]between 33,000 SEk (€3,300) and 38,000 SEK (€3,800) depending on year

A potential outcome of a negative health shock is of course death. We link the National Causes of Death register to our sample in order to investigate how common this is. The National Causes of Death register records all deaths of individuals who have a permanent residence in Sweden. Around 8 percent of our shock group and around 1 percent of the controls die within our time period. We have performed estimations including only those individuals that are alive throughout the time period, and found that our results are quite insensitive to this restriction (see A, table 13).

For all the treated and controls we extract a set of socioeconomic variable. Most of these variables are measured only once per year (in November). In order to cope with the potential problem of socioeconomic status being affected by the actual health shock, we use the variables from the year prior to the health shock.

For each observation in the sample we add a number of economic outcomes variables for the time period 1990 - 2004. The variables are taken from the LOUISE database and are measured on yearly basis. Our main economic outcome is yearly labor income. It records all cash compensation paid by employers. We also utilize register data on income from the sickness-(SI) and unemployment insurance (UI).

4 Descriptive statistics

Table 1 provides statistics on the fraction of the population affected by a health shock in a given year (we present statistics for 1995). The population is divided by age and level of education. The first row presents the fraction affected by any health shock. The table reveals a pronounced age pattern. Health shocks are much more common in old ages. About 3.5 percent experience a health shock each year in the youngest age group (age 30 -39). The same number is 5.5 percent in the oldest age group. Similar patterns hold for both low and high educated. This is in line with the idea that health deteriorates with age.

As expected there are also large differences between individuals with high and low education. We present statistics for individuals with university education (high education) and individuals without university education (low education). Individuals with low education are much more likely to be affected by negative health shocks compared to individuals with high education. We see that this is the case for all age groups. For instance, in the youngest age group the likelihood of at least one health shock in a given year is about 40 percent larger among the low educated group.

Table 1 also reports the incidence of the ten most common diagnoses. For most diagnoses the incidence increases with age, most notably for heart diseases and cancer. Heart diseases goes from being the third rarest shock for the youngest age category to being the most common one for the oldest age group, while the cancer incidence is four times higher in the age 50 - 60 category compared to the age 30 - 39. We also see that the incidence is larger in the group with low education for almost all diagnoses. A notable exception is cancer. For this diagnose the incidence is more or less equal for both individuals in age 30 -39 and individuals in age 40 - 49, and there are only a small difference in the oldest age category.

Table 2 reports statistics on a number of the background variables used in the analysis. Note that all variables are measured one year prior to the "potential" shock, and that we use additional characteristics in our analysis (e.g. residence municipality and occupational sector). The statistics are divided by treatment status and age group. The background characteristics display some important and expected patterns. Individuals with lower education, males, a child in the household, immigrants and non-married are all more likely to experience a health shock. These patterns hold for all age categories.

The upper panel of table 3 presents similar characteristics, but now separately by individuals with high and low education. In the lower panel of the table we also report information on our main outcome variable, labor earnings. Here the labor earnings are measured one year prior to the shock. The pre-shock labor earnings are important to have in mind when interpreting the size of our estimates. As expected, labor earnings are higher for those that do not experience a health shock and for individuals with high education. This has at least two important implications for our empirical analysis. First, it is going to be important to control for this pre-shock differences in labor earnings. Second, when comparing the size of the effects for individuals with high and low education it is important to keep in mind that the starting level is very different for these two groups. It is also worth noticing that we have conditioned our population to have earnings larger than one Price Basic Amount two years prior to the shock. It means that the above observed differences are not entirely driven by individuals who never have been a part of the labor force.

5 Graphical evidence

In this section we provide graphical evidence on the long-term impacts of a health shock. Figure 1-3 show, for each age group, the average labor earnings for those that are not affected by a negative health shock and those affected by a shock. Figure 4 provides similar information, but here the sample is divided by level of education and all age groups are taken together. The average labor earnings are displayed by time from the "potential" shock year, i.e. the year the treated experience the shock and the year the controls potentially could have experienced a shock. We are able to follow all individuals at least 2 years before and at least 4 years after the shock, and some individuals we can follow as long as 10 years before and 12 years after the shock.

These four graphs directly reveal a number of interesting patterns. As already seen from the descriptive statistics, health shocks are more common in lower socioeconomic groups, i.e. individuals with lower labor earnings. This is particularly apparent in the older age categories. Taking this heterogeneity into account is therefore crucial. Furthermore, there are also differences in the pre-shock trends in labor earnings. Already, several years before the actual shock labor earnings increases more among the control group. We control for such pre-shock trends in labor earnings by including both linear and quadratic trends for the two groups in our analysis.

Besides these differences in the long-run trends there seems to be a small, but nevertheless apparent decline in labor earnings for the treated group already one year before the shock year. This suggests that there are some health shocks that are anticipated and/or some health shocks that affect the individual long before they forces the individual to seek medical help. We have therefore produced similar figures for each type of health shock, presented in the appendix. They show that individuals that suffer health shocks like cancer and mental and behavioral diseases have declining labor earnings several years before the admission. For other type of health shocks like infectious diseases and external accidents we do not see such pre-treatment effects. Considering the type of health shocks, these patterns are not surprising. For instance, mental health problems may affect the individuals long before they realize that they need to seek professional help. The existent of pre-treatment effects for some but not for other health shocks is something we have to take into account in our analysis.

Taking these pre-treatment trends into account it is still apparent that there is a large jump in the labor earnings following a negative health shock. This essentially reflects the short-term effect of a health shock. There are some exceptions to this. For instance infectious- and genitourinary diseases. For these shock types the drop in earnings seems to be fairly small. Beside this immediate jump, figures 1-4 also provide a first indication that there are long-term impacts of health shocks. The gap in labor income that emerges in the shock year never seems to be closed. The individuals that suffer from a negative health shock seem to end up on a lower long-term income trajectory, with long-lasting impacts on their labor earnings. This is the case for both low- and high educated but it looks from figure 4 that the recovery is faster and stronger among those with a high education.

One could also note a big difference between figure 1, 2 and 3. In Figure 3 there is a pronounced decrease in labor earnings at time five and onwards for the shock group and at time eight and onwards for the controls. This decline could be explained by the fact that our sample gets older and as a result leaves the work force (i.e. retire). As expected we do not see this pattern for the younger age groups in figure 1 and 2. Also note that the average earnings start to decline closer in time of the shock year for our treatment group. This could be an indicator that early retirement due to health reasons is a non negligible way to exit the labor force for those who suffer a negative health shock. The effect is noticeable for both the low- and high educated.

6 Empirical strategy

The aim of this paper is to estimate the effect of a negative health shock on labor earnings. In order to do this we compare the labor earnings of those affected by a shock in the given year to those not affected. We will start by estimating the average effect on earnings. We will then use the rich data and estimate potential placebo effects. Last, we will also estimate a range of heterogeneous effect models. We start to estimate a model without heterogeneous effects. The baseline model is as follow:

$$y_{it} = \beta_0 + \lambda_t + \lambda_z + \sum_{\tau=1}^{n} \delta_{t+\tau} D_i + \varepsilon_{ti} \tag{1}$$

here the subscripts denotes individual (i), time in terms of time from shock year (t), and calendar year (z). y_{it} is the labor earnings for individual i at time t. The λ:s denotes time from health shock- (λ_t) and calendar year (λ_z) time fixed effects. The calendar year effects control for general differences in the economy, and the time-fixed effects control for the timing with respect to the shock year. D is a dummy variable taking the value 1 if the individual suffer a health shock in in time t (treated) and 0 otherwise (control). The coefficient of interest is δ. This coefficient measures the effect on earnings between the treated and the controls at time t. Since we believe that the effect on earnings may be dynamic over time, i.e. the effect is most likely large at time $t = 1$ due to the health shock but then the effect may decrease (increase) as the affected's labor supply returns to the same level as before (drops even more), we allow the treatment to vary with time. The next step is to extend the model so that it also includes observed individual characteristics. Not controlling for these may lead to biased estimation results. The new model specification is:

$$y_{it} = \beta_0 + \lambda_t + \lambda_z + \sum_{\tau=1}^{n} \delta_{t+\tau} D_i + \mathbf{X}'\beta + \varepsilon_{ti} \tag{2}$$

where $\mathbf{X}'$ is a vector of observed characteristics including, gender, level of education, immigrant status, age, residence municipality, and sector of employment. All covariates are measured one year prior to the treatment in order to handle the potential problem of socioeconomic status being affected by the actual treatment. Note that, as a result of this, all observed characteristics are time constant. This specification gives us unbiased estimates under the assumption that there is no unobserved heterogeneity between our treatment- and control group. This is unfortunately a very strong assumption.

In order to take also unobserved heterogeneity into account we estimate an individual fixed effects model. This model specification allows us to control for unobserved, but time-invariant, heterogeneity. Under the assumption that the endogenous variables, such as labor preferences, early life environment or underlying ability are constant over time, the model yields unbiased estimates. The fixed effect approach thus allows us to control for time-invariant factors at the individual level that may be associated with both underlying health and labor market outcomes. This, in combination with the fact that we only focus on acute admissions, allow us to treat the health shock as unexpected / exogeneous from the individual's perspective. This in turn implies that we are able to give a causal interpretation to our estimations. We know that our shock- and control group seem to have different pre-shock trends in labor earnings. As a way to handle this issue we add linear trends to the model.

Given the rich panel data available we estimate the following model:

$$y_{it} = \beta_0 + \lambda_t + \lambda_z + \alpha_i + (\gamma_1 * D_i + \gamma_2(1 - D_i)) + \sum_{\tau=1}^{n} \delta_{t+\tau} D_i + \varepsilon_{ti} \tag{3}$$

where α_i denotes the individual fixed effect. The expression $(\gamma_1 * D_i + \gamma_2(1 - D_i))$ captures the linear trend. γ_1 takes the value one if the individual lack university education and zero otherwise, while γ_2 takes the value one if the individual have university education and zero otherwise. Thus allowing four different trends, depending on treatment and level of education.

7 Results

In this section we present results from a model without heterogeneous effects, and the next subsection presents different placebo estimates. Based on the results from these two sections we then perform and discuss an extensive set of heterogeneous effect models. In the last subsection we then perform important sensitivity analysis, where we investigate whether our estimates are due to differences in the severity of the health shock or if it is due to differences in the possibility to cope with a given health shock.

Table 4 shows the estimated short-term and long-term effects for different models. The four models successively become more and more extended. In the first specification, column (1), we only control for calendar year and time fixed effects. The calendar year effects control for general differences in the economy, and the time-fixed effects control for the timing with respect to the shock year. In model two we control for a long row of observed characteristics, including gender, level and type of education, immigrant status, age, residence municipality, and sector of employment. All covariates are measured one year prior to the shock in order to handle the potential problem of socioeconomic status being affected by the actual health shock. Not controlling for these observed characteristics may lead to biased estimation results. In the third model we instead control for individual fixed effects.[3] Finally, we know from figures 1-4 that our shock and control group seem to have different pre-shock trends. In order to handle this issue, linear trends are included in the fourth model specification.

Let us first focus on the short-term results and see how the results differ across the five models. We find large significantly negative effects on labor earnings of a health shock in all five specifications. The results from the basic specification in column one is that the average labor earning declines with approximately 24,000 SEK (€2,400) in the shock year. As expected, the estimate decreases when we include more and more controls. For both model two and three, with observed characteristics and fixed effects the estimate is about 20,000 SEK (€2,000). When trends are included into the model (column 4) the estimate further drops to about 16,000 SEK (€1,600).

[3] Note that all observed characteristics are time constant, since they are measured in the year before the shock year. This prohibits an analysis including both the observed characteristics and the fixed effects.

In comparison with the average yearly labor earnings of about 180,000 SEK (€18,000) it is a sizeable and economically significant effect.

Next consider the long-term effects from the most extended specification in model 4. They show that the average effect of a health shock increases with time. Remember that the direct effect was SEK 16,000. After five years this effect is SEK 24,000 and after ten years the effect is as large as 30,000 SEK. It confirms the results from previous studies. [Referenser] Health shocks have large and long-lasting labor market effects. Before investigating this average effect in more detail we turn to a series of placebo effect estimations.

8 Placebo effects

In order to verify that the effects estimated above are truly an effect of the negative health shock we have performed a series of placebo estimations. We move the time of the shock one or two years back in time. If the estimation of this pseudo-shock were to yield significant results, we might worry that the estimates from equation 6 do not represent an effect of the actual health shock but some other group specific characteristic influencing both the probability to suffer a negative health shock and declining labor earnings.

Table 5 shows the estimated placebo effects, using the same five models as in Section 7. The first model gives significant and large effect from the placebo health shocks two years before the actual health shocks. However, we see that the coefficient diminishes in size as we extend the model. In the fourth model specification, using fixed effects and linear trend, the estimated effect of the pseudo-shock is 1,920 SEK (€192). It is a fairly small effect compared to the sharp decline that occurs at the time of the actual shock. It is only significant due to the very large sample size. Table 6 present the same estimations but now the placebo treatment is given the year prior to the shock. We see that estimates are similar to those in table 5. This confirms the previous insights that it is important to have a long panel with extended information both before and after the health shocks. This permits detailed controls for unobserved characteristics as well as underlying trends.

Table 15 in the appendix present placebo estimates for each diagnose group. We use the most extended model specification, individual fixed effects and linear trends. The placebo treatment takes place two years prior to the actual shock. We see that the size of the placebo estimate differs quite a lot between the diagnoses. The biggest effect is for cancer

(3,836 SEK) and heart disease (3,416 SEK), while the smallest is found for infectious- and respiratory diseases. The only nonsignificant effect is for infectious diseaes. However, all effects are fairly small and the statistical significance could to a large extend be explained by the large sample size.

9 Heterogeneous effects

We know from section 2 and 4 that health shocks are more common in lower socioeconomic groups and at older ages. If individuals from lower socioeconomic groups not only suffer from more frequent health shocks but also suffer disproportionally hard for a given health shock it may provide a reason to why the socioeconomic gradient in health widens with age. It may also be the case that the differential effect varies by age. At older ages individuals from higher socioeconomic groups have accumulated more resources both in terms of wealth and health. We will therefore estimate the heterogeneous effects in this section. The effect will be measured in two dimensions; age and level of education. We focus on our most extended model with, individual fixed effects, linear- and quadratic trends. We estimate the effect for each age category separately, and measure the education level heterogeneity using a shock - low education interaction.

$$y_{it} = \beta_0 + \lambda_t + \lambda_z + \alpha_i + (\gamma_1 * D_i + \gamma_2(1 - D_i)) + \sum_{\tau=1}^{n} \delta_{t+\tau} D_i + \sum_{\tau=1}^{n} \theta_{t+\tau} D_i + \gamma_i + \varepsilon_{ti} \quad (4)$$

where γ_i is a dummy variable taking the value 1 if the individual lacks university education and 0 otherwise. δ measures the main effect of the shock while θ is the added effect for the low educated. All else is as in section 6. In order to make a valid comparison of the relative effects for different subgroups we have to keep in mind that the earnings before the treatment differs a lot between subgroups. Highly educated and older individuals have significantly higher earnings. For instance, if the decline in earnings is larger, in absolute numbers, for the high educated, the relative decline may still be larger for the low educated. In order to take this into account we construct a measure that sets the individual's current earnings in relation to the average labor earnings for his/hers age and level of education one year prior to the shock. More precisely, we divide the current earnings level with this pretreatment earnings level. The resulting earnings ratio could then be used in the analysis to capture the relative effects of a health shock.

Table 7 shows the estimated shock effect on yearly earnings by age and level of education. It directly reveals some very striking patterns. We find very strong heterogeneous effects. In the year of the health shock the labor earning effect is approximately 5.5 percent for those with high education in all age groups. The effect for the low educated is much larger. In fact, for those aged 50 - 59 the effect is almost twice as large for those with low education compared to the group with university education.

Besides these strong short-term heterogeneous effects there are also very interesting time patterns. Interestingely the time pattern differ a lot across age groups. For the youngest group the difference between the high- and low educated decreases over time. Directly after the health shock the effect is almost 80 percent larger for the group with low education. After two years this difference has decreased to 40 percent and after five years the effect for the high educated is basically as strong as for the low educated. At older ages completely different time patterns emerge. For both those in age 40 - 49 and those in 50 - 59 the difference in the effect for high- and low educated increases over time. For instance, five years after the shock the effect is almost twice as large among those with low education in the 40 - 49 group. In the oldest age category the ratio is almost six. It means that at old ages there are very large differences in long-run possibilities to cope with a negativ health shock. Most likely, it is an effect of that many individuals with low education chooses to retire entirely from the labor market when they suffer a negative health shock at such high age.

To verify the results we, again, run placebo regressions. The results are presented in table 14. We see that there is a significant but small pre-treatment effect for the high educated. The effect for low educated is much larger but still fairly small.

Next, we turn to separate analysis for each type of health shock. The results are shown in table 9. The first panel of the table show the effect on yearly earnings on the year of the shock. We see that the heterogenous effects from 7 is apparent for most shock types. For external accidents the effect is more than twice as large for the low educated compared to the high educated. There are also very large differences for musculoskeletal diagnosis and heart diseases. The last panel of the table present the long term effects for each health shock. We can clearly see that the difference between high- and low educated disappear in the long run for the youngest ages. At the same time the difference gets even more pronounced for the oldest age category. Actually, besides for heart diseases there seem to be no long run negative earnings effects for the high educated 50 - 59 year olds even though there are large negative effects for the low educated.

10 Sensitivity analysis

The results so far suggest that individuals with low education have much worse ability to cope with health shocks, especially at old ages. If correct, that provides two additional explanations to why the socioeconomic health gradient increases at high ages. We regard this as an important insight. There is however other competing explanations to the patterns that we have found in the previous subsections. Namely that individuals with low education suffer from much worse health shocks and / or receive worse treatment. This would imply that it is this differential severity rather than the differential ability to cope with health shocks that explain our results.

As a way to test this alternative explanation we look at the share of the affected that have a secondary or third diagnose and one or more medical procedures[4]. We find it plausible to assume that more diagnoses imply a worse health status. For number of medical procedures the story could go both ways. More procedures could be a sign of worse health but it could also be an indicator of better treatment which potentially could improve the long term outcome. If the low educated have more diagnosis or procedures than the high educated, the found heterogeneous effects could to some extent be explained.

The results for number of secondary and third diagnoses are shown in the top panel of table 10. The number of diagnoses are divided into four categories; no secondary diagnose, a secondary diagnose, a third diagnose, and three or more diagnosis. We see that the vast majority does not have any secondary diagnose even though the probability of more diagnosis increases with age. Among those who have a secondary diagnose, just one secondary diagnose is the most common. Most importantly, there are no differences between the low- and high educated in terms of secondary diagnosis.

The lower panel of table 10 show the number of medical procedures by age and level of education. As before, we divide them into four categories; no procedures, one procedure, two procedures, and three or more procedures. Unlike with secondary diagnose, there does not seem to be a clear age pattern when it comes to number of medical procedures. Most patients do not have any medical procedure and those who have usually have just one. There are only small differences between the high- and low educated. Table 12 (appendix?) present the same result but divided also by shock type. Again, there is no apparent difference between low- and high educated in terms of number of medical procedures.

[4]e.g. surgery

These results do not lead us to believe that the observed socioeconomic differences in labor earnings could be explained by neither the number of diagnosis nor the number of medical procedures.

An explanation related to the number of medical procedures is that individuals with lower socioeconomic status face a different supply of treatment than high socioeconomic individuals. This hypothesis could be tested by looking at to what extent the patients are admitted to university (teaching?) hospitals[5] . Beside clinical education and training of future and current health professionals these hospitals also are centers for experimental, innovative and advanced health procedures. Being admitted to a university hospital could thus indicate that the patient receive better and more advanced treatment.

Table 11 show the share of treated admitted to a university hospital. In the first row, all treated are included. We see that for all age categories, high educated are a lot more likely to have been admitted to a university hospital compared to the low educated. In the second row we restrict the sample to only contain those with a secondary or third diagnose. The idea is that a more complicated health shock should improve the possibility of being admitted to a university hospital. However, this restriction does not change the result. In the third row the sample is restricted to only those with one or more medical procedures. The idea is again that a more severe health shock improves the possibility of being admitted to a university hospital. From the table we clearly see that those who have one or more medical procedure to a larger extent are admitted to a university hospital. This is true for all age categories and regardless of the patient's socioeconomic status. However, a larger share of high educated are still admitted to a university hospital compared to the low educated.

Given that university hospitals are located in cities with a university (and thus have a more educated population), the above results are not surprising. In order to see if there is an education difference in being admitted to a university hospital all else equal, we estimate a simple linear probability model controlling for individual characteristics (including residence municipality) and shock specific information such as shock type, secondary diagnosis, and number of procedures. The results are shown in the lower panel of 11. The first row presents the likelihood when all shock types are included in the model. High educated are more likely than low

[5]Sweden had eight university hospitals during the period: Akademiska sjukhuset, Uppsala; Karolinska Universitetssjukhuset, Stockholm; Norrlands universitetssjukhus, Umeå; Sahlgrenska Universitetssjukhuset, Göteborg; Universitetssjukhuset MAS; Universitetssjukhuset i Lund; Universitetssjukhuset i Linköping; Universitetssjukhuset Örebro

educated to be admitted to a university hospital than low educated even when controlling for residence area. This is true for all ages. However, the effect is small. The following rows of table 11 present the likelihood given a certain shock type. The likelihood of admission to a university hospital is statistically significant and positive for most shock types and age groups. However, the effect is fairly small...

11 Conclusions

12 Tables

Table 2: Share of population affected by shock in 1995, by age and level of education

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Control	Shock	Control	Shock	Control	Shock
Age	34.3	34.6	44.6	44.9	54.0	54.3
Male	0.54	0.59	0.50	0.53	0.50	0.54
Level of education						
Primary and lower sec. edu. ,< 9 years	0.010	0.016	0.086	0.11	0.24	0.27
Primary and lower sec. edu. ,9 (10) years	0.13	0.17	0.13	0.15	0.078	0.076
Upper secondary edu. , < 3 years	0.41	0.44	0.33	0.35	0.29	0.30
Upper secondary edu. , 3 years	0.12	0.10	0.12	0.11	0.12	0.12
Post-secondary edu., <2 years	0.17	0.15	0.16	0.14	0.11	0.11
Post-secondary edu. , ≥ 2 years	0.13	0.100	0.16	0.13	0.14	0.12

mean coefficients

The individual is considered employed if he/she has performed labor in November each given year.

An immigrant is an individual born outside of Sweden.

Continued on next page

– continued from last page

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Control	Shock	Control	Shock	Control	Shock
Postgraduate education	0.0061	0.0039	0.0095	0.0067	0.011	0.0089
Education unknown	0.014	0.020	0.0068	0.0089	0.0045	0.0060
Married	0.45	0.43	0.65	0.60	0.70	0.67
Immigrant	0.096	0.11	0.10	0.12	0.093	0.10
Child in household	0.63	0.59	0.62	0.55	0.18	0.16
Children 0 - 3	0.32	0.27	0.066	0.057	0.0050	0.0055
Children 4 - 6	0.29	0.27	0.12	0.098	0.0093	0.0090
Children 7 - 10	0.27	0.27	0.23	0.20	0.030	0.027
Children 11 - 15	0.15	0.17	0.37	0.33	0.097	0.086
Children 16 - 17	0.027	0.036	0.19	0.18	0.079	0.069
Employed	0.92	0.90	0.95	0.93	0.95	0.94
Occupational area						
Agriculture, hunting & forestry	0.012	0.013	0.011	0.012	0.012	0.013
Fishing	0.00016	0.00021	0.00016	0.00016	0.00016	0.00015
Mining & quarrying	0.0021	0.0026	0.0024	0.0031	0.0031	0.0036
Manufacturing	0.20	0.21	0.18	0.19	0.19	0.20
Electricity, gas & water supply	0.0068	0.0067	0.0093	0.0088	0.0092	0.0094
Construction	0.059	0.065	0.056	0.061	0.059	0.063
Wholesale & retail trade	0.12	0.11	0.099	0.093	0.10	0.097

mean coefficients
The individual is considered employed if he/she has performed labor in November each given year.
An immigrant is an individual born outside of Sweden.

Continued on next page

– continued from last page	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Control	Shock	Control	Shock	Control	Shock
Hotels & restaurants	0.021	0.024	0.012	0.014	0.011	0.011
Transport, storage & communication	0.073	0.078	0.069	0.073	0.065	0.071
Financial intermediation	0.024	0.017	0.025	0.019	0.025	0.020
Real estate, renting & business act.	0.10	0.088	0.089	0.084	0.089	0.082
Public administration & defence	0.049	0.047	0.067	0.064	0.074	0.072
Education	0.052	0.046	0.085	0.077	0.099	0.089
Health & social work	0.20	0.20	0.22	0.22	0.19	0.19
Other community, social & personal service	0.037	0.039	0.039	0.041	0.042	0.046
Activities of households	0.000019	0.000037	0.000021	0.0000045	0.000036	0.000012
Extra-territorial org. & bodies	0.000057	0.000049	0.000025	0.000045	0.000064	0.000044
Observations	378392		434354		426466	

mean coefficients
The individual is considered employed if he/she has performed labor in November each given year.
An immigrant is an individual born outside of Sweden.

Table 1: Share of population affected by shock in 1995, by age and level of education

	Age 30-39			Age 40-49			Age 50-59		
	Low edu Affected	High edu Affected	Total Affected	Low edu Affected	High edu Affected	Total Affected	Low edu Affected	High edu Affected	Total Affected
Any Shock (%)	3.90	2.46	3.48	4.50	2.94	4.04	5.69	4.12	5.34
Infectious diseases (%)	0.19	0.17	0.18	0.16	0.14	0.15	0.19	0.16	0.18
Cancer (%)	0.08	0.08	0.08	0.18	0.18	0.18	0.30	0.29	0.30
Mental & behavioral (%)	0.79	0.27	0.64	0.91	0.36	0.75	0.73	0.39	0.65
Nerve system (%)	0.12	0.06	0.10	0.13	0.08	0.11	0.15	0.10	0.14
Heart diseases (%)	0.17	0.12	0.16	0.48	0.30	0.43	1.18	0.77	1.09
Respiratory diseases (%)	0.25	0.16	0.22	0.24	0.17	0.22	0.37	0.23	0.33
Digestive organs (%)	0.46	0.34	0.43	0.57	0.38	0.52	0.73	0.53	0.68
Musculoskeletal (%)	0.22	0.13	0.20	0.27	0.18	0.24	0.33	0.22	0.31
Genitourinary (%)	0.29	0.19	0.26	0.30	0.23	0.28	0.30	0.25	0.29
External accidents (%)	0.74	0.39	0.64	0.75	0.47	0.67	0.79	0.60	0.75

Table 3: Background variables, by age and level of education

	Age 30 - 39				Age 40 - 49				Age 50 - 59			
	Low edu		High edu		Low edu		High edu		Low edu		High edu	
	Control	Shock	Control	Shock	Control	Shock	Control	Shock	Control	Shock	Control	Shock
Age	34.3	34.5	34.4	34.7	44.7	44.9	44.5	44.9	54.1	54.4	53.8	54.1
Male	0.55	0.60	0.51	0.54	0.52	0.55	0.47	0.48	0.51	0.55	0.48	0.49
Married	0.43	0.41	0.49	0.48	0.63	0.58	0.69	0.66	0.69	0.66	0.72	0.69
Immigrant	0.092	0.10	0.079	0.086	0.10	0.12	0.091	0.095	0.094	0.10	0.084	0.090
Child in household	0.65	0.60	0.59	0.56	0.58	0.52	0.70	0.65	0.15	0.14	0.26	0.23
Employed	0.91	0.90	0.94	0.94	0.95	0.93	0.97	0.96	0.95	0.94	0.97	0.96
Earnings												
Mean	153,086	145,353	199,092	198,358	171,504	159,920	240,440	229,216	176,778	167,072	266,227	250,380
p25	104,700	93,800	128,400	129,700	124,100	109,000	163,100	154,700	123,400	110,800	184,100	172,000
Median	156,400	148,600	189,400	186,300	168,200	159250	212,600	206,900	171,200	164,300	233,600	225,400
p75	199,900	194,200	253,900	250,400	212,000	204,100	285,000	273,500	217,000	211,000	305,250	293,500
N	145,159	119,224	67,160	40,582	142,708	158,988	69,469	59,770	126,928	191,954	46,512	58,766

Mean coefficients. All variables measured one year prior to shock.
The individual is considered employed if he/she has performed labor in November each given year.
An immigrant is an individual born outside of Sweden.
All differences within age groups are statistically significant at the 1 percent level

Table 4: Treatment effect on yearly earnings, measured t years from treatment.

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings
time t	-24240.2***	-21217.7***	-21416.7***	-16152.7***
	(242.9)	(219.9)	(164.3)	(150.2)
time t+1	-27360.5***	-24119.2***	-24698.6***	-18018.4***
	(265.5)	(243.3)	(196.2)	(197.7)
time t+2	-28216.6***	-25104.6***	-25779.8***	-17655.5***
	(284.8)	(257.6)	(221.0)	(244.4)
time t+3	-31399.6***	-28553.1***	-29238.3***	-19667.9***
	(294.4)	(269.9)	(234.9)	(290.2)
time t+4	-35252.0***	-32646.3***	-33366.4***	-22352.6***
	(318.1)	(294.0)	(263.7)	(340.2)
time t+5	-38952.4***	-36105.6***	-36585.8***	-24336.8***
	(347.6)	(322.0)	(289.7)	(407.2)
time t+6	-42927.6***	-39993.8***	-39968.7***	-26454.0***
	(374.4)	(347.5)	(310.6)	(463.1)
time t+7	-46699.0***	-43636.4***	-43262.4***	-28450.0***
	(402.0)	(371.8)	(330.3)	(519.5)
time t+8	-50025.3***	-46805.8***	-46101.1***	-29998.1***
	(464.4)	(433.4)	(381.3)	(590.9)
time t+9	-52107.2***	-48812.4***	-47850.6***	-30400.8***
	(503.0)	(469.3)	(410.0)	(643.7)
time t+10	-53929.8***	-50841.7***	-49463.8***	-30701.6***
	(576.9)	(543.3)	(462.9)	(710.5)
Time variables	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Fixed Effects	No	No	Yes	Yes
Linear trends	No	No	No	Yes
N	17679410	16688491	17679410	17679410

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Placebo effect on yearly earnings. Placebo two years before actual shock.

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings
time t-2	-8815.0***	-6465.0***	-4831.9***	-1946.2***
	(206.0)	(189.8)	(156.6)	(180.0)
time t	-28050.1***	-23924.9***	-24025.4***	-19271.8***
	(248.3)	(225.8)	(212.8)	(298.8)
time t+2	-32024.2***	-27810.8***	-28376.8***	-21781.0***
	(289.4)	(262.7)	(261.9)	(422.4)
time t+5	-42757.7***	-38811.1***	-39254.1***	-29952.4***
	(351.4)	(326.1)	(321.1)	(633.1)
Time variables	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Fixed effects	No	No	Yes	Yes
Linear trends	No	No	No	Yes
N	17679410	16688491	17679410	17679410

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Placebo effect on yearly earnings. Placebo one year before actual shock.

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings
time t-1	-10489.7***	-7491.5***	-7062.8***	-3215.7***
	(223.7)	(200.9)	(154.3)	(155.7)
time t	-26130.9***	-22541.8***	-22831.3***	-17858.4***
	(244.7)	(222.0)	(186.6)	(205.8)
time t+2	-30106.0***	-26428.0***	-27187.3***	-19938.8***
	(286.3)	(259.4)	(238.8)	(307.7)
time t+5	-40840.5***	-37428.7***	-38031.6***	-27472.6***
	(348.9)	(323.4)	(302.9)	(487.7)
Time variables	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Fixed effects	No	No	Yes	Yes
Linear trends	No	No	No	Yes
N	17679410	16688491	17679410	17679410

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Treatment effect on yearly earnings by age and level of education. Measured t years from treatment.

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
t	-0.0528***	-0.0431***	-0.0531***	-0.0484***	-0.0581***	-0.0544***
	(0.00270)	(0.00283)	(0.00198)	(0.00203)	(0.00194)	(0.00191)
t+2	-0.0740***	-0.0415***	-0.0395***	-0.0674***	-0.0355***	-0.0876***
	(0.00495)	(0.00509)	(0.00311)	(0.00321)	(0.00315)	(0.00313)
t+5	-0.135***	-0.0220*	-0.0525***	-0.0900***	-0.0216***	-0.125***
	(0.00976)	(0.00992)	(0.00527)	(0.00539)	(0.00519)	(0.00508)
Time variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Linear trends	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Placebo effect on yearly earnings by age and level of education. Placebo two years before actual shock.

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
t-2	-0.0128***	0.00597	-0.00880***	0.00998***	-0.00756***	0.00624**
	(0.00332)	(0.00346)	(0.00218)	(0.00223)	(0.00195)	(0.00193)
t	-0.0598***	-0.0539***	-0.0622***	-0.0521***	-0.0644***	-0.0601***
	(0.00567)	(0.00593)	(0.00363)	(0.00373)	(0.00329)	(0.00328)
t+2	-0.0830***	-0.0563***	-0.0515***	-0.0727***	-0.0438***	-0.0955***
	(0.00858)	(0.00888)	(0.00521)	(0.00536)	(0.00470)	(0.00467)
t+5	-0.147***	-0.0426**	-0.0689***	-0.0976***	-0.0329***	-0.136***
	(0.0144)	(0.0147)	(0.00800)	(0.00818)	(0.00708)	(0.00695)
Time variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Linear trends	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Treatment effect on yearly earnings by age and level of education. Measured t years from treatment.

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
t						
Infectious diseases	-0.0137	-0.0426**	-0.0299***	-0.0361***	-0.0234*	-0.0471***
Cancer	-0.121***	-0.0472**	-0.111***	-0.0539***	-0.140***	-0.0731***
	(0.0143)	(0.0176)	(0.00735)	(0.00864)	(0.00651)	(0.00768)
Mental & behavioral	-0.158***	-0.0311**	-0.152***	-0.00802	-0.142***	-0.0128
	(0.0102)	(0.0111)	(0.00654)	(0.00741)	(0.00845)	(0.00968)
Nerve system	-0.0728***	-0.0343	-0.0936***	-0.0180	-0.0626***	-0.0620***
	(0.0176)	(0.0195)	(0.0126)	(0.0143)	(0.0106)	(0.0120)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Continued on next page

– continued from previous page						
	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
Heart diseases	-0.0382**	-0.0569***	-0.0807***	-0.0623***	-0.0984***	-0.0664***
	(0.0125)	(0.0138)	(0.00587)	(0.00662)	(0.00412)	(0.00462)
Respiratory diseases	-0.0198	-0.0197	-0.0379***	-0.0155	-0.0376***	-0.0404***
	(0.0131)	(0.0144)	(0.00951)	(0.0103)	(0.00703)	(0.00794)
Digestive organs	-0.0220***	-0.0282***	-0.0219***	-0.0379***	-0.0214***	-0.0388***
	(0.00634)	(0.00715)	(0.00447)	(0.00508)	(0.00482)	(0.00534)
Musculoskeletal	-0.0623***	-0.0940***	-0.0861***	-0.0792***	-0.0909***	-0.0778***
	(0.0110)	(0.0122)	(0.00616)	(0.00737)	(0.00938)	(0.0101)
Genitourinary	-0.00868	-0.0317***	-0.0104*	-0.0347***	-0.00863	-0.0336***
	(0.00759)	(0.00891)	(0.00493)	(0.00589)	(0.00624)	(0.00711)
External accidents	-0.0529***	-0.0686***	-0.0632***	-0.0702***	-0.0673***	-0.0708***
	(0.00586)	(0.00647)	(0.00455)	(0.00510)	(0.00423)	(0.00477)
t+2						
Infectious diseases	-0.0246	-0.0654**	-0.0272*	-0.0378*	-0.0192	-0.0552**
	(0.0221)	(0.0243)	(0.0137)	(0.0161)	(0.0188)	(0.0207)
Cancer	-0.129***	-0.0173	-0.0386***	-0.0572***	-0.0435***	-0.107***
	(0.0232)	(0.0278)	(0.0110)	(0.0130)	(0.00985)	(0.0120)
Mental & behavioral	-0.209***	-0.0524**	-0.147***	-0.0654***	-0.125***	-0.0574***
	(0.0170)	(0.0184)	(0.0109)	(0.0123)	(0.0123)	(0.0142)
Nerve system	-0.0602	-0.0940	-0.113***	-0.0443*	-0.0510**	-0.120***
	(0.0559)	(0.0574)	(0.0191)	(0.0219)	(0.0162)	(0.0188)
Heart diseases	-0.0504**	-0.0486*	-0.0770***	-0.0748***	-0.116***	-0.0950***
	(0.0177)	(0.0202)	(0.00934)	(0.0106)	(0.00744)	(0.00822)
Respiratory diseases	-0.00605	-0.0524*	-0.0373*	-0.0321	-0.0222	-0.0841***
	(0.0215)	(0.0237)	(0.0172)	(0.0185)	(0.0130)	(0.0143)
Digestive organs	-0.0492***	-0.0170	-0.00141	-0.0612***	0.0112	-0.0739***
	(0.0116)	(0.0128)	(0.00844)	(0.00935)	(0.00726)	(0.00820)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Continued on next page

– continued from previous page

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
Musculoskeletal	-0.0275	-0.0995***	-0.0417***	-0.0836***	-0.0281*	-0.109***
	(0.0176)	(0.0195)	(0.00979)	(0.0118)	(0.0114)	(0.0129)
Genitourinary	-0.106***	-0.00413	0.00448	-0.0557***	0.0248*	-0.0718***
	(0.0144)	(0.0162)	(0.00764)	(0.00941)	(0.0105)	(0.0120)
External accidents	-0.0603***	-0.0462***	-0.0338***	-0.0679***	-0.0146	-0.0864***
	(0.0102)	(0.0111)	(0.00728)	(0.00816)	(0.00811)	(0.00885)
t+5						
Infectious diseases	-0.0913*	-0.0374	-0.0350	-0.0696**	0.0181	-0.122**
	(0.0384)	(0.0418)	(0.0226)	(0.0267)	(0.0421)	(0.0442)
Cancer	-0.123**	-0.0185	0.00386	-0.0821***	0.0484**	-0.143***
	(0.0407)	(0.0460)	(0.0170)	(0.0200)	(0.0151)	(0.0180)
Mental & behavioral	-0.296***	-0.0204	-0.171***	-0.0650***	-0.0388*	-0.0985***
	(0.0269)	(0.0290)	(0.0161)	(0.0185)	(0.0175)	(0.0223)
Nerve system	-0.199***	-0.0160	-0.151***	-0.0689*	-0.0331	-0.164***
	(0.0444)	(0.0496)	(0.0298)	(0.0341)	(0.0274)	(0.0309)
Heart diseases	-0.100***	-0.0448	-0.110***	-0.0809***	-0.149***	-0.111***
	(0.0285)	(0.0321)	(0.0162)	(0.0180)	(0.0119)	(0.0129)
Respiratory diseases	-0.0679	-0.0110	-0.0402	-0.0724*	0.00855	-0.130***
	(0.0464)	(0.0492)	(0.0339)	(0.0356)	(0.0183)	(0.0208)
Digestive organs	-0.0938**	-0.0110	-0.0139	-0.0848***	0.0229	-0.112***
	(0.0335)	(0.0345)	(0.0133)	(0.0147)	(0.0134)	(0.0147)
Musculoskeletal	-0.111***	-0.0394	-0.0458**	-0.110***	-0.0280	-0.119***
	(0.0263)	(0.0295)	(0.0170)	(0.0199)	(0.0211)	(0.0233)
Genitourinary	-0.150***	-0.0150	0.0194	-0.104***	0.0743***	-0.128***
	(0.0236)	(0.0266)	(0.0129)	(0.0156)	(0.0182)	(0.0204)
External accidents	-0.116***	-0.0280	-0.0441***	-0.0789***	0.00497	-0.118***
	(0.0215)	(0.0226)	(0.0131)	(0.0143)	(0.0114)	(0.0127)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Number of bidiagnoses and medical procedures. By age and level of education

	Age 30-39		Age 40-49		Age 50-59	
	Low edu.	High edu.	Low edu.	High edu.	Low edu.	High edu.
# of diagnoses						
0	0.78	0.79	0.74	0.76	0.70	0.71
1	0.17	0.17	0.19	0.18	0.21	0.21
2	0.037	0.034	0.047	0.042	0.062	0.056
≥ 3	0.014	0.012	0.017	0.016	0.024	0.022
# of procedures						
0	0.69	0.68	0.70	0.68	0.72	0.71
1	0.22	0.24	0.21	0.22	0.18	0.19
2	0.060	0.063	0.066	0.070	0.064	0.066
≥ 3	0.025	0.024	0.028	0.030	0.032	0.036
Observations	119224	40582	158988	59770	191954	58766

mean coefficients

Table 11: Share and probaility of being admitted to a university hospital, by age and level of education

	Age 30-39		Age 40-49		Age 50-59	
	Low edu.	High edu.	Low edu.	High edu.	Low edu.	High edu.
Total (%)	20	25	19	25	20	25
Secondary diagnose (%)	21	26	21	25	20	26
#procedures >0 (%)	24	29	25	30	27	31
	Probability of being admitted to a university hospital					
	All shock types					
High edu.	0.00638***		0.00673***		0.0121***	
	(0.000457)		(0.000382)		(0.000390)	
	Infectious diseases					
High edu.	0.00998***		0.00297		0.00254	
	0.00179)		(0.00167)		(0.00186)	
	Cancer					
High edu.	0.0521***		0.00645*		0.0182***	
	(0.00430)		(0.00259)		(0.00289)	
	Mental & behavioral					
High edu.	0.00435**		-0.00293*		0.0142***	
	(0.00160)		(0.00125)		(0.00161)	
	Nerve system					
High edu.	-0.00601		0.00919**		0.00845**	
	(0.00337)		(0.00307)		(0.00292)	
	Heart diseases					
High edu.	0.00228		0.00803***		0.0133***	
	(0.00226)		(0.00129)		(0.000961)	
	Respiratory					
High edu.	0.0140***		0.00518***		0.00362*	
	(0.00165)		(0.00153)		(0.00162)	
	Digestive organs					
High edu.	0.00278*		0.00473***		0.00308***	
	(0.00108)		(0.000893)		(0.000898)	
	Musculoskeletal					
High edu.	0.00584**		0.00316*		0.0190***	
	(0.00193)		(0.00160)		(0.00175)	
	Genitourinary					
High edu.	-0.00139		0.00758***		0.0214***	
	(0.00170)		(0.00131)		(0.00158)	
	External accidents					
High edu.	0.00386***		0.00183		0.00529***	
	(0.00114)		(0.000997)		(0.00108)	

mean coefficients

Table 12: Number of medical procedures by shock type, age, and level of education.

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Low edu.	High edu.	Low edu.	High edu.	Low edu.	High edu.
Infectious diseases						
1	0.046	0.041	0.051	0.040	0.034	0.040
2	0.0055	0.0049	0.0054	0.0051	0.0066	0.0055
≥ 3	0.0017	0.0029	0.0037	0.0025	0.0051	0.0035
Cancer						
1	0.41	0.41	0.39	0.44	0.30	0.32
2	0.14	0.13	0.14	0.15	0.12	0.12
≥ 3	0.060	0.074	0.073	0.069	0.081	0.076
Mental & behavioral						
1	0.0024	0.0032	0.0018	0.0018	0.0037	0.0028
2	0.00030	0	0.00031	0.00026	0.00054	0.00039
≥ 3	0.00020	0	0.00086	0.00079	0.00043	0.00079
Nerve system						
1	0.065	0.070	0.068	0.056	0.060	0.056
2	0.0094	0.0042	0.0080	0.011	0.013	0.0081
≥ 3	0.0029	0.0031	0.0034	0.0052	0.0035	0.0029
Heart diseases						
1	0.15	0.15	0.11	0.12	0.093	0.11
2	0.043	0.039	0.045	0.045	0.041	0.039
≥ 3	0.017	0.019	0.020	0.023	0.025	0.022
Respiratory diseases						
1	0.15	0.16	0.12	0.13	0.089	0.097
2	0.018	0.018	0.025	0.022	0.016	0.016
≥ 3	0.0067	0.0074	0.0097	0.0059	0.0068	0.010
Digestive organs						
1	0.49	0.52	0.42	0.43	0.37	0.36

Continued on next page

– continued from previous page						
	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
2	0.11	0.12	0.12	0.13	0.12	0.12
≥ 3	0.030	0.028	0.037	0.037	0.041	0.044
Musculoskeletal						
1	0.28	0.29	0.25	0.27	0.26	0.25
2	0.073	0.075	0.072	0.068	0.082	0.067
≥ 3	0.022	0.025	0.022	0.017	0.024	0.022
Genitourinary						
1	0.32	0.36	0.34	0.35	0.34	0.34
2	0.11	0.11	0.15	0.14	0.13	0.13
≥ 3	0.037	0.037	0.054	0.054	0.047	0.055
External accidents						
1	0.30	0.35	0.32	0.35	0.33	0.34
2	0.11	0.12	0.11	0.12	0.13	0.13
≥ 3	0.060	0.053	0.061	0.059	0.067	0.069
Observations	88345	29419	119602	44625	145920	43997

13 Graphs

Figure 1: Shock effect, age group 30 - 39

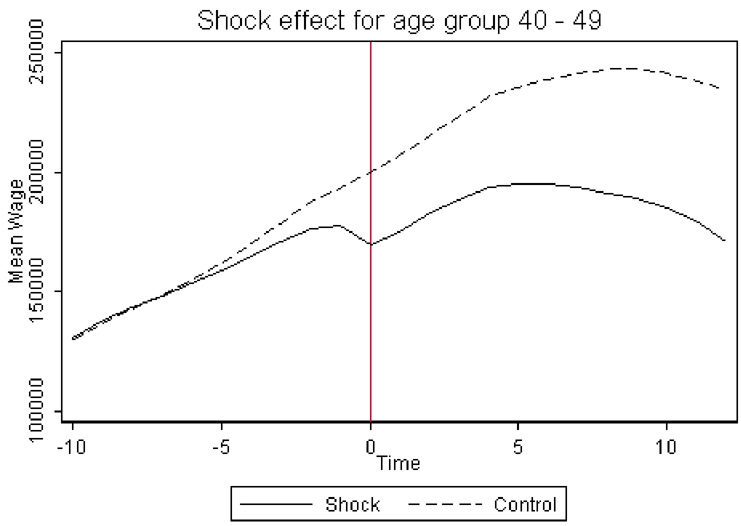

Figure 2: Shock effect, age group 40 - 49

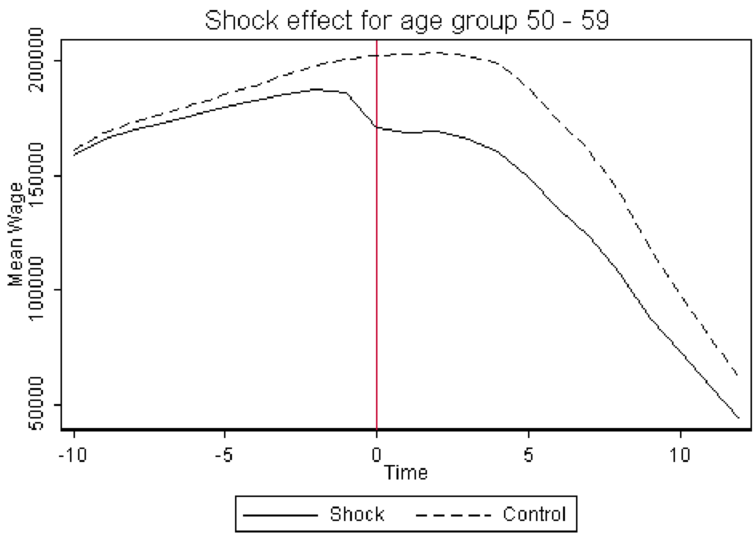

Figure 3: Shock effect, age group 50 - 59

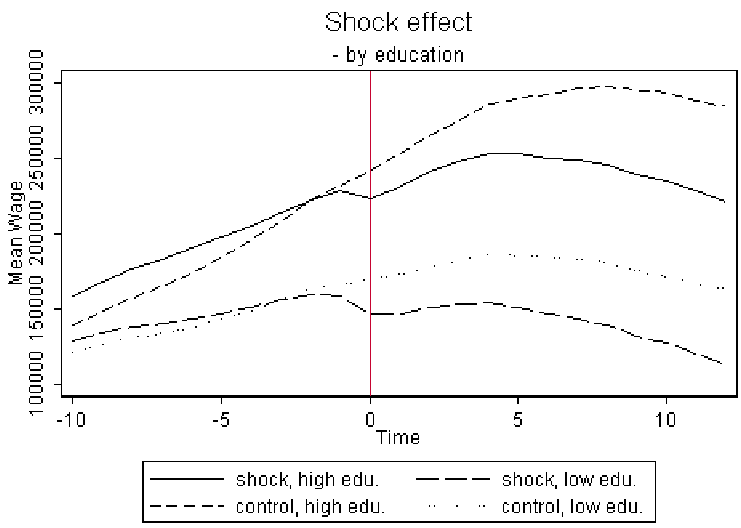

Figure 4: Shock effect, by level of education

A Appendix A

Table 13: Treatment effect on yearly earnings, measured t years from treatment. Conditional on survival throughout the time period.

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings
time t	-22479.6***	-21217.4***	-20086.4***	-14971.9***
	(249.0)	(219.9)	(168.1)	(153.5)
time t+1	-25471.0***	-24117.4***	-23076.8***	-16599.3***
	(271.3)	(243.3)	(199.9)	(201.6)
time t+2	-26576.2***	-25105.2***	-24244.5***	-16376.8***
	(289.8)	(257.7)	(224.7)	(248.9)
time t+3	-29948.9***	-28555.8***	-27738.3***	-18474.6***
	(299.1)	(269.9)	(238.5)	(295.8)
time t+4	-33999.1***	-32648.9***	-31895.9***	-21238.1***
	(322.5)	(294.0)	(267.3)	(346.5)
time t+5	-37730.2***	-36103.3***	-35038.2***	-23191.3***
	(352.2)	(322.1)	(293.6)	(414.9)
time t+6	-41760.6***	-39991.4***	-38358.6***	-25290.7***
	(378.8)	(347.5)	(314.3)	(471.5)
time t+7	-45738.9***	-43635.5***	-41698.9***	-27379.7***
	(405.7)	(371.8)	(333.6)	(528.5)
time t+8	-49248.7***	-46802.4***	-44545.3***	-28978.8***
	(467.9)	(433.4)	(384.5)	(600.6)
time t+9	-51503.4***	-48811.5***	-46307.4***	-29434.5***
	(506.1)	(469.3)	(412.9)	(653.9)
time t+10	-53662.6***	-50837.2***	-48022.9***	-29876.2***
	(579.5)	(543.3)	(465.5)	(721.0)
Time variables	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Fixed Effects	No	No	Yes	Yes
Linear trends	No	No	No	Yes
N	17131882	16686928	17131882	17131882

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B Appendix B

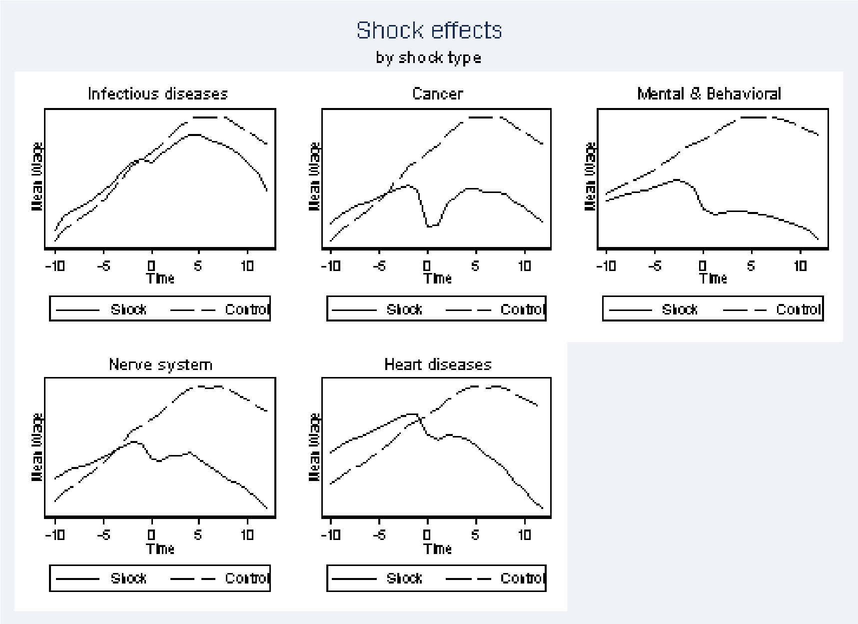

Figure 5: Shock effect, by shock type

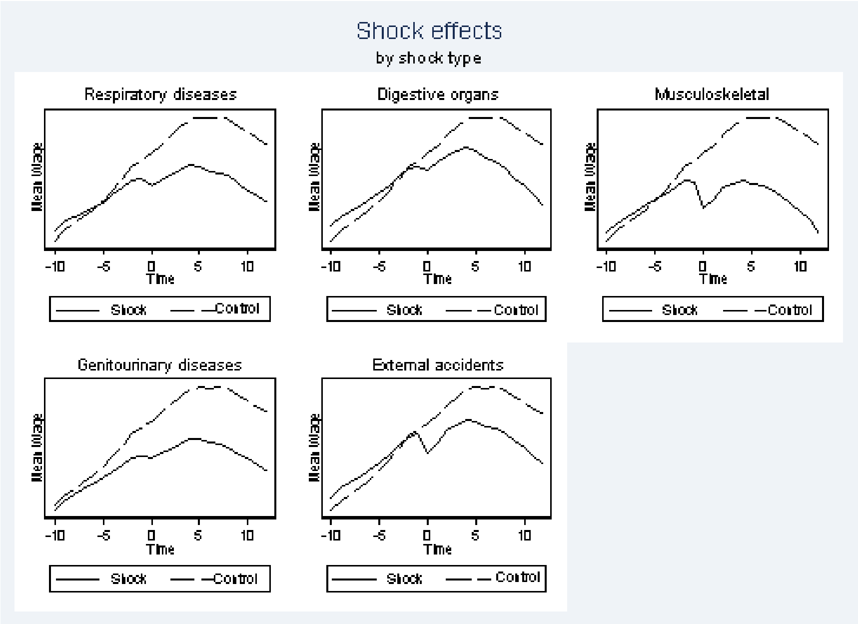

Figure 6: Shock effect, by shock type

C Appendix C

Table 14: Share of population who died within the timeperiod.

	Died		
	Full sample	High edu.	Low edu.
Control (%)	1.28	1.48	0.82
Shock (%)	8.24	8.90	6.33
Total (%)	4.87	5.43	3.39
Observations	1235679	882153	341552

D Appendix D

References

ANTONOVSKY, A. (1967): "Social Class, Life Expectancy and Overall Mortality," *The Milbank Memorial Fund Quarterly*, 45(2), 31–73, ArticleType: primary_article / Issue Title: Part 1 / Full publication date: Apr., 1967 / Copyright Â© 1967 Milbank Memorial Fund.

AU, D. W. H., T. F. CROSSLEY, AND M. SCHELLHORN (2005): "The effect of health changes and long-term health on the work activity of older Canadians," *Health Economics*, 14(10), 999–1018.

CASE, A., AND A. DEATON (2005): "Health and wealth among the poor: India and South Africa compared," Discussion paper, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.

CASE, A. C., AND A. DEATON (2003): "Broken Down by Work and Sex: How Our Health Declines," Discussion paper, National Bureau of Economic Research, Inc.

CSDH (2008): "WHO | Commission on Social Determinants of Health - final report," Discussion paper, World Health Organization, Geneva.

CURRIE, J., B. C. MADRIAN, O. C. ASHENFELTER, AND D. CARD (1999): "Chapter 50 Health, health insurance and the labor market," vol. Volume 3, Part 3, pp. 3309–3416. Elsevier.

DANO, A. M. (2005): "Road injuries and long-run effects on income and employment," *Health Economics*, 14(9), 955–970.

Table 15: Placeboeffect by shock type

	(1) Infectious diseases	(2) Cancer	(3) Mental & behavioral	(4) Nerve system	(5) Heart diseases	(6) Respiratory	(7) Digestive organs	(8) Musculoskeletal
time t-2	-825.7	-3836.9***	-3404.6***	-2475.3**	-3416.8***	-1170.8*	-1698.7***	-2419.4***
	(828.1)	(497.2)	(507.9)	(758.1)	(407.1)	(553.8)	(376.1)	(516.0)
time t	-10777.9***	-37173.8***	-38356.2***	-22180.8***	-29184.5***	-10884.3***	-10927.1***	-29641.5***
	(1434.5)	(892.9)	(846.0)	(1218.1)	(671.6)	(945.2)	(602.5)	(870.5)
time t+2	-13073.9***	-27473.7***	-48444.2***	-29032.3***	-38214.9***	-13737.9***	-12438.5***	-24954.2***
	(2006.7)	(1240.3)	(1153.8)	(1773.6)	(955.6)	(1357.4)	(842.2)	(1210.7)
time t+5	-19374.8***	-25640.2***	-55094.9***	-38993.3***	-56032.2***	-18708.7***	-20352.4***	-32946.9***
	(2812.3)	(1794.1)	(1682.8)	(2494.6)	(1403.5)	(2111.1)	(1333.5)	(1766.6)
Time variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8895871	8914301	9148557	8767169	9678413	9029573	9713311	9088398

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

DAVIDSON, R. J. (1992): "Emotion and Affective Style: Hemispheric Substrates," *Psychological Science*, 3(1), 39–43.

DEATON, A. (2002): "Policy Implications Of The Gradient Of Health And Wealth," *Health Aff*, 21(2), 13–30.

DISNEY, R., C. EMMERSON, AND M. WAKEFIELD (2006): "Ill health and retirement in Britain: A panel data-based analysis," *Journal of Health Economics*, 25(4), 621–649.

D'UVA, T. B., E. V. DOORSLAER, M. LINDEBOOM, AND O. O'DONNELL (2008): "Does reporting heterogeneity bias the measurement of health disparities?," *Health Economics*, 17(3), 351–375.

ETILÉ, F., AND C. MILCENT (2006): "Income-related reporting heterogeneity in self-assessed health: evidence from France," *Health Economics*, 15(9), 965–981, PMID: 16958084.

FREDRIKSSON, P., AND P. JOHANSSON (2008): "Dynamic Treatment Assignment," *Journal of Business and Economic Statistics*, 26(4), 435–445.

GOLDMAN, D., AND J. P. SMITH (2005): "Socioeconomic Differences in the Adoption of New Medical Technologies," *American Economic Review*, 95(2), 234–237.

GOLDMAN, D. P., AND J. P. SMITH (2002): "Can Patient Self-Management Help Explain the SES Health Gradient?," *SSRN eLibrary*.

GÓMEZ, P. G., AND A. L. NICOLÁS (2006): "Health shocks, employment and income in the Spanish labour market," *Health Economics*, 15(9), 997–1009.

GROSS, J. J. (1998): "Relations between Affect and Personality: Support for the Affect-Level and Affective-Reactivity Views," *Personality & social psychology bulletin*, 24(3), 279–288.

GROSSMAN, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223–55.

MACKENBACH, J. P., AND M. BAKKER (2002): *Reducing inequalities in health: a European perspective*. Routledge.

MACKENBACH, J. P., V. BOS, O. ANDERSEN, M. CARDANO, G. COSTA, S. HARDING, A. REID, O. HEMSTROM, T. VALKONEN, AND A. E. KUNST (2003): "Widening socioeconomic inequalities in mortality in six Western European countries," *Int. J. Epidemiol.*, 32(5), 830–837.

Marmot, M. (1999): "Multilevel approaches to understanding social determinants," in *Social Epidemiology*, ed. by L. Berkman, and I. Kawachi. Oxford: Oxford University Press.

Peltonen, M., M. Rosen, V. Lundberg, and K. Asplund (2000): "Social Patterning of Myocardial Infarction and Stroke in Sweden: Incidence and Survival," *Am. J. Epidemiol.*, 151(3), 283–292.

Riphahn, R. T. (1999): "Income and employment effects of health shocks A test case for the German welfare state," *Journal of Population Economics*, 12(3), 363–389.

Schrijvers, C. T., and J. P. Mackenbach (1994): "Cancer patient survival by socioeconomic status in seven countries: a review for six common cancer sites [corrected]," *Journal of Epidemiology and Community Health*, 48(5), 441–446, PMID: 7964352.

Smith, J. P. (1998): "Socioeconomic Status and Health," *The American Economic Review*, 88(2), 192–196, ArticleType: primary_article / Issue Title: Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association / Full publication date: May, 1998 / Copyright Â© 1998 American Economic Association.

——— (1999): "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status," *Journal of Economic Perspectives*, 13(2), 145–166.

van Doorslaer, E., A. Wagstaff, H. Bleichrodt, S. Calonge, U. Gerdtham, M. Gerfin, J. Geurts, L. Gross, U. Häkkinen, R. E. Leu, O. O'Donell, C. Propper, F. Puffer, M. Rodríguez, G. Sundberg, and O. Winkelhake (1997): "Income-related inequalities in health: some international comparisons," *Journal of Health Economics*, 16(1), 93–112.