



IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

**Socioeconomic heterogeneity in the
effect of health shocks on earnings:
evidence from population-wide data on
Swedish workers**

Petter Lundborg
Martin Nilsson
Johan Vikström

WORKING PAPER 2011:11

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala

Visiting address: Kyrkogårdsgatan 6, Uppsala

Phone: +46 18 471 70 70

Fax: +46 18 471 70 71

ifau@ifau.uu.se

www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Socioeconomic heterogeneity in the effect of health shocks on earnings: evidence from population-wide data on Swedish workers^a

by

Petter Lundborg^b, Martin Nilsson^c and Johan Vikström^d

31st August, 2011

Abstract

In this paper, we test for the existence of socioeconomic heterogeneity in the effect of health shocks on labor market outcomes using register data on the total population of Swedish workers. We estimate fixed effect models and use unexpected hospitalizations as a measure of health shocks. Our results suggest large heterogeneity in the effects, where low educated individuals suffer relatively more from a given health shock. This result holds across a wide range of different health shocks and our results suggest that the heterogeneity increases by age. We test several potential explanations to these results. Extensive sensitivity analyses, including a difference-in-differences matching model, show that our estimates are robust to a number of potential threats. We conclude that socioeconomic heterogeneity in the effect of health shocks offers one explanation to why the socioeconomic gradient in health widens during middle ages.

Keywords: Health, Health Shocks, Socioeconomic Status, Life-cycle
JEL-codes: I10, I12

^aWe would like to thank Caroline Hall, Per Johansson and seminar participants at Lunds University and Uppsala University for helpful comments. The financial support of the Swedish Council of Working Life and Social Research FAS (dnr 2004-2005 and dnr 2009-0826) is acknowledged.

^bLunds University, VU University Amsterdam, HEP, Centre for Economic Demography, IZA, and Tinbergen Institute, petter.lundborg@nek.lu.se.

^cDepartment of Economics Uppsala University, IFAU-Uppsala and UCLS, martin.nilsson@ifau.uu.se

^dIFAU-Uppsala and UCLS, johan.vikstrom@ifau.uu.se

Table of contents

1	Introduction	3
2	Background	7
3	Data.....	10
3.1	Descriptive statistics	13
4	Graphical analysis	14
5	Empirical strategy.....	17
6	Results.....	20
6.1	Average effects	20
6.2	Heterogeneous effects by level of education and age	22
6.3	Heterogeneous effects by type of health shock.....	24
7	Robustness analysis.....	26
7.1	Placebo effects	26
7.2	DID-matching.....	29
7.3	Detailed measure of education	31
7.4	Analysis using survivals	33
7.5	Severity of the health shocks	35
8	What explains the heterogeneous effects?	36
9	Conclusions	41
	References	44

1 Introduction

It has been widely documented that people of higher socioeconomic position enjoy better health. This socioeconomic gradient in health seems to hold up irrespective of what health measures that are used (see e.g. van Doorslaer et al. 1997, Marmot 1999, Smith 1998, Mackenbach & Bakker 2002). The socio-economic gradient in health is commonly found to widen during the middle-ages, but then narrow again as people reach older ages (see e.g. van Kippersluis et al. 2009, Case & Deaton 2005*b*). It is also believed that socioeconomic inequalities in health have increased during the recent decades in most Western countries (Mackenbach et al. 2003).

While there is a general agreement about the existence of a socioeconomic gradient in health, there is surprisingly little agreement about its underlying causes. In the epidemiological literature, it has traditionally been assumed that socioeconomic status in terms of income and/or level of education affects health. Economists have instead explored the hypothesis that health outcomes influence socioeconomic status (e.g. Smith 1998). 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. Surveys by Smith (1999) and Case & Deaton (2005*a*) even conclude that a larger part of the association between health and socioeconomic status at middle and older ages likely reflects an impact of health on socioeconomic status.

In this paper, we test for the existence of socioeconomic heterogeneity in the effects of health shocks on labor market outcomes. The previous literature has for the most part focused on the average effects of health shocks on labor market outcomes, while heterogeneity in the effects of health shocks has not, to our knowledge, been thoroughly explored in the literature.

We find the focus on average effects restrictive for a number of reasons. First, a substantial literature has shown socioeconomic heterogeneity in the recovery and survival from medical conditions, such as cancers and heart diseases (e.g. Schrijvers & Mackenbach 1994, Smith et al. 1998, Peltonen et al. 2000). In line with this, results have shown that high educated individuals are better at adhering to medical treatments, such as AIDS

and diabetes treatments (Goldman & Smith 2002). Second, a recent literature has shown differences in access to medical technologies and treatments by socioeconomic status (Rosvall et al. 2008). Third, people of different socioeconomic status may face different incentives to return to the labor market after facing a health shock, due to the structure of the health insurance and social insurance systems. Fourth, the extent to which job tasks require good physical health may vary according to socioeconomic status.

If there exists substantial heterogeneity in the effect of health shocks on labor market outcomes, both by socioeconomic status and age, we believe that this could be an important explanation for the widely documented increase in the socioeconomic gradient in health over the life-cycle. If the *impact* of a given health shock is stronger for people of low socioeconomic status, they would face a double penalty as they already face an increased risk of experiencing negative health shocks. Moreover, the differential impact by socioeconomic status may vary by age. For instance, the cumulative advantage hypothesis suggests that some mediators of the socioeconomic status and health relationship (e.g. smoking or social capital) accumulates over the life cycle. This suggests that older individuals from lower socioeconomic groups may be especially sensitive to health shocks.

There are several reasons why knowledge about heterogeneity in the impact of health shocks on labor market outcomes are important for policy. First, such knowledge may point to the possibility of targeted efforts towards groups who suffer disproportionately from health shocks. Example of such policies are more intense screening for health markers among socioeconomic risk groups, regular health check-ups, and improving adherence 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 shocks are typically crudely measured through the average wage rate.

In order to investigate the presence of heterogeneous effect of health shocks, we use several large-scale register data sets on the entire population of Swedish workers. In the analysis, we use education as a measure of socioeconomic status. The main reason for this is that level of education is unaffected by health shocks during middle-ages. Our data

and analytical approach has a number of advantages. First, our large sample allow us to estimate heterogeneous effects by both socioeconomic status and age with large precision. The large sample size also allow us to consider a large number of different types of health shocks.

Second, most of the economic studies on the impact of health events treat them as exogenously given (see e.g. the survey by Currie & Madrian 1999). Only a small number of recent studies have addressed the endogeneity of health events and provided evidence suggesting a causal effect of health events on labor outcomes (see e.g. Riphahn 1999, Au et al. 2005, Disney et al. 2006, Gómez & Nicolás 2006). In our study, 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. Our data also 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. Moreover, by comparing the responses across educational groups we are able to difference away any remaining anticipations effects. This holds if the anticipation effects are similar for individuals with high and low education, which seems plausible from inspection of the raw data. Together with the fixed-effects specification, focusing on unexpected health shocks and studying heterogenous responses facilitates a causal interpretation of our estimates. We also perform an extensive set of placebo estimates as well as sensitivity analysis using DID-matching (see e.g. Heckman et al. 1997).

Third, our estimates are based on detailed register data on health shocks taken from the national inpatient hospital registers, while most previous literature use data on self-reported health shocks. Using register 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 & Milcent 2006, d'Uva et al. 2008). This is problematic, since it will bias the estimates of the effect of health shocks on labor market outcomes.

Fourth, our data allow us to investigate some possible mechanisms that may give rise

to socioeconomic heterogeneity in the impact of health shocks. We are able to test to what extent the heterogeneity arises from differential access to health-care and treatments, differences in the severity of health shocks, differences in occupations, and differences in the incentives to return to work after a health shock across socioeconomic groups. Knowledge about the mechanisms is important for policy purposes. For instance, if differential access to medical care is the main reason for the observed gradient, then improved access to health care for low socioeconomic groups would be one policy option one could use in order to weaken the gradient.

Fifth, we are able to follow individuals for up to 14 years. This allow us to consider both the short and long-term impact of health shocks. Since some health shocks may permanently reduce the work capacity, examining the long-term impact is important. 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. Since adherence to medical therapies have been found to vary by socioeconomic status, this may also imply that the long-term impact of a given health shock may vary by socioeconomic status (Goldman & Smith 2002).

We start our paper by documenting large and long-run average effects of health shocks on yearly earnings. We then show that these average estimates mask substantial heterogeneity in the effects of health shocks across socioeconomic groups. In the short-run, the effect of a health shock are much larger for individuals from lower socioeconomic groups. The difference is most pronounced for older individuals (aged 50-59) were the effect for individuals with low socioeconomic status is more than twice that for individuals with high socioeconomic status.

Our results also suggest some interesting time patterns. For young individuals (aged 30-39) the difference between individuals with low and high socioeconomic status decreases with time. At older ages the picture is completely different. For both those aged 40-49 and 50-59, the difference in the effect by socioeconomic status instead increases with time since the health shock. This suggests that at old ages there are very large differences in the long-run possibilities to cope with a negative health shock. Interestingly, we find similar heterogeneous effects by socioeconomic status and age across all types of

health shocks. These results are consistent with the idea that the socioeconomic gradient in health is partly caused by the impact of health shocks on socioeconomic status. They also offer one explanation for why the socioeconomic gradient in health widens during middle ages.

The paper is organized as follows. Section 2 discuss the recent literature on the effect of health shocks on socioeconomic outcomes and provides a background to the observed correlation between health and income over the life-cycle. Section 3, describes our unique population data and provides descriptive statistics. In section 4 we provide initial graphical evidence on the effects of health shocks. Section 5 presents our empirical strategy. Section 6 presents our main results, and section 7 presents an extended set of robustness analysis. Section 8 tries to explain the heterogeneous results. Finally, section 9 concludes.

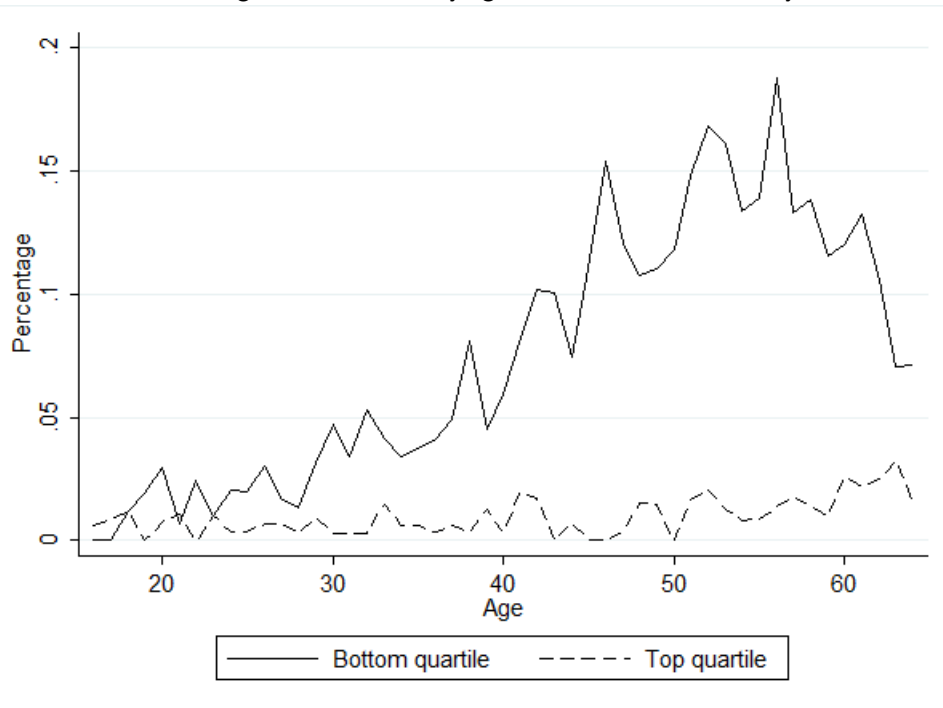
2 Background

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 the 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 (see e.g. van Kippersluis et al. 2009, Case & Deaton 2005b). *Figure 1* illustrates this pattern using Swedish survey data.¹ It shows the fraction in the upper income quartile and bottom income quartile at different ages that states that they have bad health. It confirms that there is a strong socioeconomic gradient in health, and that this gradient widens during working life but narrows at old ages.

Even if the same pattern emerges in many countries the causes of this pattern have been widely debated and there is currently no consensus in the literature. 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

¹The survey data comes from the Swedish database ULF (Survey of living standards). It is conducted on a yearly basis and covers a random sample of about 3000 individuals.

Figure 1: The socioeconomic gradient in health by age. Based on Swedish survey data from 1975-2007



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 (socioeconomic status) through their effect on labor market outcomes. A sudden negative health event, such as a work related injury, myocardial infarction, a stroke, or an accident leads to a reduced labor supply in the short-term. There may also be important long-term effects, however. Some health events are so severe that they permanently reduce the work capacity. Even less severe events may have long-term effects, since the initial time out of work may imply that valuable experience and contacts are lost and since future employers may be reluctant to hire the individual due to the risk of future health events. For all these reasons, one can expect important short-term and long-term effects of negative health events on labor supply and income.

As discussed in the introduction, the recent literature has for the most part implicitly assumed that the impact of a health event on labor market outcomes is the same across subgroups of the population. 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 events have been found to strongly predict retirement and reduced labor force participation (Smith 2005, 2007, Case & Deaton 2005b)²

As discussed in the introduction, there are several reasons to expect substantial heterogeneity in the impact of health shocks on labor market outcomes. First, evidence from the medical literature clearly suggests that there are marked socioeconomic differences in survival from cancers, stroke, coronary heart disease and acute myocardial infarction (Schrijvers & Mackenbach 1994, Smith et al. 1998, Peltonen et al. 2000, Tonne et al. 2005). In addition, evidence from the psychological literature 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). Such evidence points to the possibility that the labor market consequences of health events also may differ according to socioeconomic status.

The evidence from the medical literature suggest that income and education may be important resources in coping with adverse health shocks. 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). Evidence consistent with this is reported in Goldman & Smith (2002) where educated people were found to better adhere to medical treatments for AIDS and diabetes, which are known to be quite demanding. Educated people may also be able to acquire more information and better handle contacts with the health care system, and thereby allow them to get more appropriate treatment (see e.g. Rosvall et al. 2008).

Second, it is likely that individuals with high education more easily could change occupation or in other ways adjust their work conditions in response to a health shock. In contrast, occupations that only require low education may often be more physical in nature, which means that reduced health may have a stronger impact on the possibility to

²One exception is Smith (1999), who estimates heterogenous effects by level of income of a new illness on wealth and medical expenses. Smith (1999) finds that households whose pre-shock household income places them above the median income faces similar medical expenses but larger wealth losses as the below median income households.

perform ones job. For such reasons we expect that educated people could be assumed to better cope with adverse health shocks.

Third, there may exist differential incentives to return to work after a health shock by socioeconomic status (SES). Replacement rates in the Swedish social insurance systems are capped at a relatively low ceiling, which means that low SES people may have weaker incentives to return to work after a health shock compared to high SES people.³

The differential effect by socioeconomic status may also vary by age. The cumulative advantage hypothesis argue that some mediators of the socioeconomic status and health relationship (e.g. smoking or social capital) accumulates over the life cycle (see e.g. Ross & Wu 1996, Lynch 2003, Willson et al. 2007). This does not only affect the onset of new health conditions, but may also make individuals from lower socioeconomic groups in middle and old ages especially sensitive to health shocks. In middle ages many individuals with high SES also have accumulated large economic resources that make them better equipped to handle health shocks. This suggests that older individuals from lower socioeconomic groups are especially sensitive to health shocks. If that is the case it could explain why the socioeconomic gradient in health increases in the middle-ages. In our analyses, we will therefore allow the effects of health shocks to vary by educational status and by age. This is in line with the arguments in e.g. Case & Deaton (2005*b*) and Smith (2007) that it is important to take a life-cycle perspective in order to understand the socioeconomic gradient in health.

It should also be noted that most previous studies have focused on the impact of one particular health event 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. We will assess the

³The Swedish sickness insurance provides economic compensation when a worker is too sick to carry out his or her regular job. This insurance automatically covers all of the employed workers. The benefits in the Swedish sickness insurance are income related, the size of the benefits depends on the persons wage prior to the sick spell. The insurance consists of two main benefits, sickness benefit (SI) and disability benefit (DI). The SI is supposed to cover part of the income loss due to temporary illness. DI compensates individuals whose work capacity is permanently reduced. The replacement rates have changed over time, but the rates have been capped at a relatively low ceiling throughout our observed time period. (About 25 percent of the workers have income above the ceiling).

effects of different diagnosis groups, which is based on the internationally standardized classification of diseases (ICD). 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 labor market consequences.

3 Data

Our data are created from a combination of three Swedish population register data sets. The first register, called LOUISE (from statistics Sweden) covers the entire Swedish population in age 16-64. It contains a rich set of socioeconomic variables recorded on a yearly basis (e.g. age, sex, immigration status, marital status and employment status) as well as economic outcomes such as yearly labor earnings. We use data for the observation period 1990-2004. The second register, the Swedish National Patient Register (NPR), includes information on all in-patient care in Sweden from 1987 and onwards. It 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 detailed information on medical procedures. Note that, in Sweden, the nominal fee for obtaining treatment is very low. It is therefore unlikely that a large number of individuals choose not to seek help in case of an important health shock.

We create our treatment group by selecting admissions from the NPR for the period 1992-2000. This time period was chosen since it allow us to use information a number of years before as well as a number of years after the health shock. We further restrict our sample to individuals who are aged 30-59 when they suffer a health shock. The reason 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 prohibits an analysis of the long-term labor market outcomes. We also restrict the analyses to acute admissions and admissions that are not related to pregnancies. The former restriction is made 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 means

that some planned admissions are anticipated several years before the actual admission. For individuals with more than one acute admission during our observation period, only the first one is used in the analysis. We use the international standard and classify all the admissions into 19 major types of diseases. Of these we choose to focus on the ten most common (in terms of incidence).⁴

Since our focus in this paper is on labor supply effects of health shocks, we exclude individuals that are never 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 have performed robustness analysis with respect to this restriction, and our results are insensitive to instead making the restriction three 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 Base Amount (between 33,000 SEK (€3,300) and 38,000 SEK (€3,800) depending on year).⁵

The control group consists of all individuals that are a part of the labor force and *potentially* could have suffered a negative health shock in each given year but who did not. For all treated and non treated individuals, we record yearly labor earnings a number of years before and after the shock year.⁶ The number of time periods are restricted by our observation period 1990-2004. Note that this sampling implies that an individual that do not suffer a health shock is included in the control group more than one year. In order to keep the empirical analysis manageable from a computational point of view we randomly

⁴We exclude Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (ICD-9: 780 - 799, ICD-10: R00-R99). 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).

⁵The price base amount is a measure set by the Swedish Government a year at a time. The amount is calculated based on changes in the consumer price index. The price base amount has various uses, including ensuring that sickness benefits, study support, etc., do not decline in value because of an increase in the general price level.

⁶Labor earnings records all, gross, cash compensation paid by employers. Beside salary, this include for instance compensation paid by the employers during the first 14 days of a sick spell and subsistence allowance. Sickness insurance benefits paid from the 15th sick day and onward, unemployment insurance benefits, disability insurance benefits and other forms of social benefits are not included in this measure.

sample 3 percent of the individuals in the control group each given year.⁷

For all individuals in the treatment and control group we extract a set of socioeconomic variable and economic outcomes from the LOUISE database. 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.

A potential outcome of a negative health shock is obviously death. To control for this we use a third register, the National Causes of Death register. This register records all deaths of individuals who have a permanent residence in Sweden.

3.1 Descriptive statistics

Table 1 provides descriptive statistics on the fraction of the population affected by a health shock in a given year (we present statistics for 1995).⁸ The population is divided by age and level of education. The table reveals a pronounced age pattern, where health shocks become more common as people reach old ages. About 3.5 percent experience a health shock each year in the youngest age group (age 30-39). In the oldest age group this number increases with almost 40 percent to 5.5 percent. Similar patterns appear for both low and high educated. This confirms that health in general deteriorates with age. For instance, heart diseases goes from being one of the rarest health shocks among the youngest, to being the most common one among the oldest age group, and the cancer incidence is four times higher in the age 50-59 category compared to the age 30-39.

There are also large differences by level of education. In the table, we show statistics for individuals with an university education (high education) and individuals without an 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. This pattern holds for all age groups and for almost all types of health shocks. For

⁷We follow the treated and non-treated over a long time period. That 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. In order to handle this we follow the dynamic treatment assignment methods developed in Fredriksson & Johansson (2008) and include the controls up until the time they suffer their first health shock.

⁸The population here excludes individuals that never are part of the labor force, as defined above.

Table 1: Sample statistics for health shocks

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

Notes: The table reports the fraction affected by any health shocks and the ten most common types of health shocks in 1995 (excluding the group symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified). High education is defined as having some kind of university education and low education less than university education.

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 compared to the group with high education. A notable exception is cancer, for which the incidence is the same regardless of educational background.

Table 2 reports sample statistics for a number of background characteristics used in the analysis and for our main outcome variable yearly labor earnings, measured one year before the potential health shock. The background characteristics display some expected patterns. Individuals with low education, males, immigrants, and individuals with a child in the household or who are single are all more likely to experience a health shock. These patterns hold for all age categories. As expected, labor earnings are greater for those that do not experience a health shock. They are also greater for individuals with high education (not shown).

Table 2: Sample statistics for background characteristics and earnings

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Control	Shock	Control	Shock	Control	Shock
<i>Background characteristics:</i>						
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
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
Primary and lower sec. edu.	0.010	0.016	0.086	0.11	0.24	0.27
Primary and lower sec. edu.	0.13	0.17	0.13	0.15	0.078	0.076
Upper secondary edu.	0.41	0.44	0.33	0.35	0.29	0.30
Upper secondary edu.	0.12	0.10	0.12	0.11	0.12	0.12
Post-secondary edu.	0.17	0.15	0.16	0.14	0.11	0.11
Post-secondary edu.	0.13	0.100	0.16	0.13	0.14	0.12
Postgraduate education	0.0061	0.0039	0.0095	0.0067	0.011	0.0089
<i>Labor earnings:</i>						
Employed	0.92	0.90	0.95	0.93	0.95	0.94
Mean	166,273	156,713	193,406	177,882	200,434	186,006
P25	108,800	98,200	133,700	118,000	135,200	120100
Median	164,300	155,700	180,500	170,100	185,200	176100
P75	213,800	205,300	231,900	220,900	240,000	230,000
# observations	378,392		434,354		426,466	

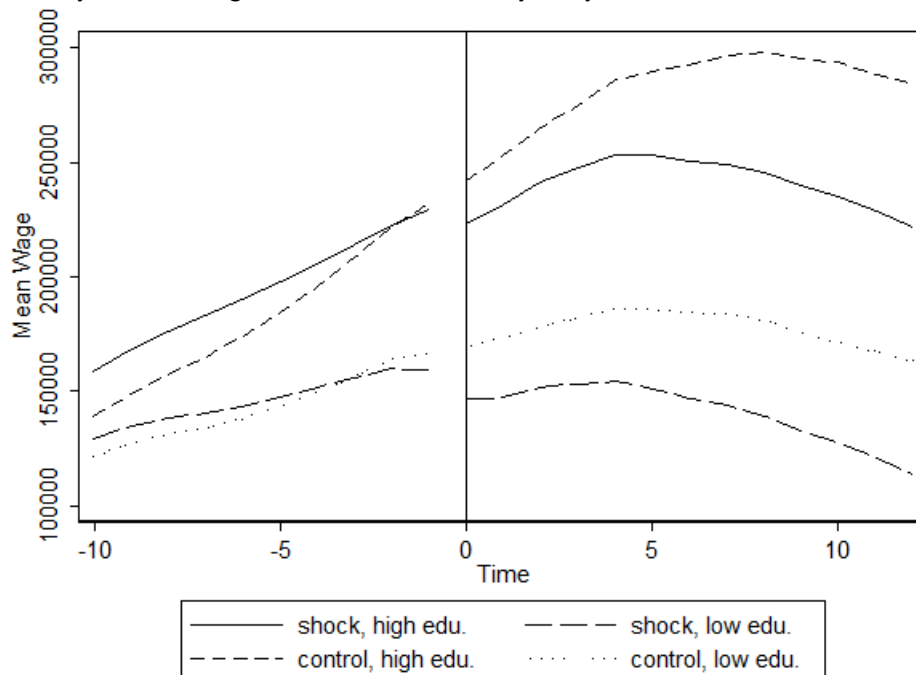
Note: The table reports background statistics for our analysis sample. The individual is considered employed if he/she has a job in November each given year. An immigrant is an individual born outside of Sweden. Earnings is reported in SEK (10 SEK \approx 1 €).

4 Graphical analysis

In this section we illustrate some of the most interesting patterns in the data graphically. This also serves as a background for our empirical analysis. *Figure 2* shows, for high and low educated, the average labor earnings for the treatment and the control group, respectively. The average labor earnings are displayed by time from the potential shock year (time=0), i.e. the year in which the treated individuals experience a health shock and the year in which the individuals in the control group potentially could have experienced a shock.

This figure reveals a number of interesting patterns. For both high and low educated the level of income as well as the pre-shock trends in labor earnings differ between

Figure 2: Yearly labor earnings before and after shock year by level of education



treated and controls. Already several years before the actual shock, labor earnings increase more in the control group. There also seems to be a small decline in labor earnings among the treated already one year before the actual shock. This suggests that there are some health shocks that are anticipated and/or affect the individuals labor earnings before they actually forces the individual to seek medical help. We conclude that taking pre-shock trends and anticipation effects into account will be important for our empirical strategy. However, even taking pre-treatment trends into account, it is still apparent that there is a large decrease, for both high and low educated, in the labor earnings in the year of the negative health shock. This indicates that health shocks have important effects on earnings.

We now turn to heterogenous effects. *Figure 3-Figure 5* present, for each age group, the ratio between average labor earnings for individuals with high and low education in the treatment and control group, respectively. Since income on average are higher among individuals with high education all ratios are above one.⁹ For the treated the earnings ratio

⁹One exception is 10 years before the shock year for individuals in age 30-39. This reflects the fact that some individuals with longer university education have not yet completed their education at this time point.

jumps up in the shock year in all age groups. For the control groups we see no such jumps. This means that a health shock decreases income relatively more for individuals with low education compared to individuals with high education. This captures socioeconomic heterogeneity in the short-term effect of health shocks. Besides this immediate difference, the figures also provide a first indication of substantial heterogeneity in the long-run. Several years after the health shock the earnings ratio between treated with high and low education is still much higher compared to the same ratio in the control group.

Figure 3: Ratio between mean labor earnings for high educated shock (control) and low educated shock (control). Age 30 - 39

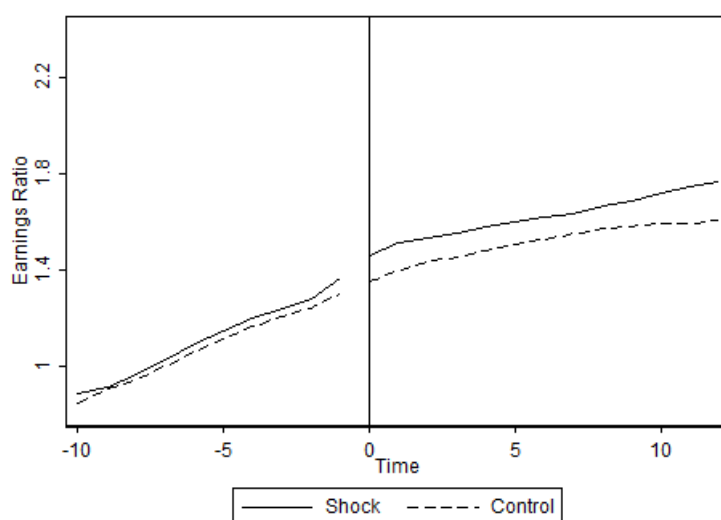


Figure 4: Ratio between mean labor earnings for high educated shock (control) and low educated shock (control). Age 40 - 49

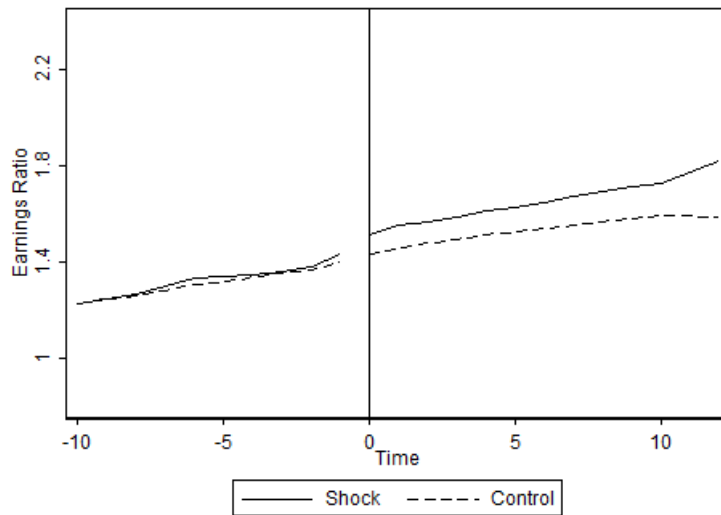


Figure 5: Ratio between mean labor earnings for high educated shock (control) and low educated shock (control). Age 50 - 59

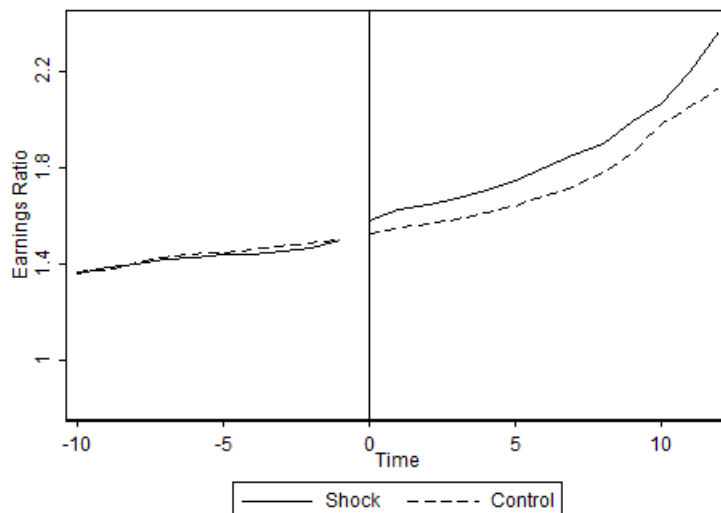


Figure 3-Figure 5 also have important implications for our empirical strategy. In *Figure 2*, illustrating average labor earnings, we saw important differences in pre-shock trends and anticipation effects. In *Figure 3-Figure 5* that focus on earnings ratios between mean labor income for high and low educated there are no important differences in pre-shock trends. Before the shock year the earnings ratio is very similar for the treatment and control groups. Moreover, there are no decline in the earnings ratio among the treated one year before the actual shock. In other words, the pre-shock trends and anticipation effects are remarkably similar for individuals with high and low education, so that focusing on heterogenous effects clearly mitigates one of the main issues with estimating the effects of health shocks. As a further illustration, *Figure A-1* in the appendix presents similar figures for each type of health shocks. They show that the earnings ratios are fairly stable before the shock year for almost all types of health shocks, even for cancer and mental and behavioral diseases. This further supports the focus on heterogenous effects.

5 Empirical strategy

The aim of this paper is to estimate the short-run and long-rung heterogenous effects of a negative health shock on labor earnings. To this end we focus on acute admissions, since there are good reasons to assume that they are more or less unanticipated from the individual's perspective. However, even if acute admission are unanticipated the probability of experiencing an acute admission may be correlated with observed and unobserved individual characteristics like labor preferences, early life environment and/or underlying ability. We will therefore include an extended set of fixed effects as well as include controls for differences in pre-shock trends in all our empirical models.

Our baseline heterogenous effects model, for labor earnings for individual i in time period t in calendar year z is:

$$y_{izt} = \lambda_t + \lambda_z + \lambda_i + \sum_{\tau=0}^T \delta_{\tau} I(t = \tau) D_i + I^{LE} \left[\sum_{\tau=0}^T \delta_{\tau}^{LE} I(t = \tau) \right] D_i + \gamma_1 D_i t + \gamma_2 (1 - D_i) t + I^{LE} \left[\gamma_1^{LE} D_i t + \gamma_2^{LE} (1 - D_i) t \right] + \varepsilon_{izt}. \quad (1)$$

We have normalized time so that year 0 is the shock year. D_i is an indicator variable taking the value 1 if the individual suffer a health shock in year 0 and zero otherwise, and I^{LE} is an indicator variable taking the value one if the individual has low education. The coefficients of interest are $\delta_0, \dots, \delta_T$ which capture the main effect of a health shock in the shock year, one year after the shock and so on, and in particular $\delta_0^{LE}, \dots, \delta_T^{LE}$ which capture the additional effect for individuals with low education. In the analysis we take age into account by running separate regressions for three different age groups.

This model controls for an extended set of fixed effects, including the timing with respect to the shock year (or potential shock year for the controls), λ_t , calendar time fixed effects, λ_z , as well as individual fixed effects¹⁰, λ_i . These fixed effects control for changes over time, aggregated changes in the economy, and all time-invariant factors at the individual level, respectively.

We also control for underlying pre-shock trends in labor earnings by including linear trends that are allowed to vary by both treatment status and level of education. That is in total four separate linear trends. We have also run models using quadratic and even cubic trends but the results were insensitive to including more flexible controls for trends.

After controlling for fixed effects and general pre-shock trends there may remain some pre-treatment effects since some health shocks could be anticipated. That is one reason for focusing on heterogenous effects, which compare the responses to a health shock across educational groups. Our model could therefore be viewed as a Difference-in-Difference-in-Differences model where we compare the change in labor earnings across treated and controls with high and low education. If the anticipation effects are similar for individuals with high and low education our heterogenous effects estimate could be given a causal interpretation. Based on the figures presented in section 4 this seem highly plausible.

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 differs across the groups. This means that even if the effect in absolute numbers is larger for the high educated, the relative

¹⁰In one specification we also include a set of covariates instead of individual fixed effects. 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. For that reason we cannot include these background characteristics and individual fixed effects at the same time.

effect on labor income may still be larger for the low educated. We therefore construct a relative income measure that sets the individual's current earnings in relation to the average earnings level among his peers (i.e. those of same age and with equal level of education). More precisely, we divide the population into six groups by age and level of education, and construct a relative outcome measure by dividing the individual earnings rate with the average earnings within the group.¹¹

We believe that this extended fixed-effects specification, where we focus on unexpected health shocks, and compare the size of the change in earnings across educational groups facilitates a causal interpretation of our estimates. However, since we rely on observational data we will perform an extensive set of robustness analyses. We will therefore (1) perform placebo estimates to test for any significant pre-shock responses two years before the actual shock. The placebo estimates will be performed jointly and separately for our ten types of health shocks. (2) Use DID-matching in the spirit of Heckman et al. (1997) instead of running fixed effects models. (3) Use detailed data on the number of medical procedures and number of diagnoses in order to investigate whether our results are driven by differences across groups in the severity of the health shocks. (4) Estimate models where we divide the population into finer educational groups. (5) Run regressions only using the individuals that survives throughout the entire observation window in order to assess if our results is affected by differential survival rate across level of education and age.

As an illustration we will also estimate a model without heterogenous effects:

$$y_{izt} = \lambda_t + \lambda_z + \lambda_i + \sum_{\tau=0}^T \delta_{\tau} I(t = \tau) D_i + \gamma_1 D_i t + \gamma_2 (1 - D_i) t + \varepsilon_{izt}. \quad (2)$$

¹¹We divide with the average earnings one year prior to year the treated experience a health shock and the controls potentially could have experienced a health shock. The reason for this is that this earnings level should be unaffected by the health shocks.

6 Results

6.1 Average effects

As a background to the analysis of heterogeneous effects *Table 3* presents baseline estimates on the average effects of health shocks. The model in column (1) includes basic controls for calendar year and time fixed effects. In model (2) we then add an extended set of observed characteristics.¹² Model (3) includes individual fixed effects, and model (4) adds separate linear trends for the treated and non-treated. In all four specifications we find large and significant effects of a health shock on labor earnings during the year of the health shock. Note that we use our relative income measure, so that the coefficients should be interpreted in terms of relative effects. For instance, the estimate of the average effect in the shock year in model 4 in *Table 3* suggests that income on average decreases with 9 percent directly after the health shock.

Interestingly, the long-term effects are larger than the short-term effects. From model 4 in *Table 3*, we see that the effect is 13 percent five years after the shock, compared to 9 percent in the shock year. It clearly indicates that health shocks have sizeable, long-lasting and economically significant effects on labor outcomes.

¹²We include gender, level and type of education, immigrant status, age, residence municipality, marital status, and sector of employment.

Table 3: Estimates of the short-run and long-run effects of health shocks

	(1)	(2)	(3)	(4)
Shock year	-0.132** (0.00121)	-0.119** (0.00113)	-0.119** (0.000861)	-0.0898** (0.000759)
Shock year+1	-0.149** (0.00135)	-0.135** (0.00127)	-0.138** (0.00105)	-0.100** (0.00102)
Shock year+2	-0.153** (0.00144)	-0.140** (0.00134)	-0.143** (0.00117)	-0.0974** (0.00126)
Shock year+3	-0.171** (0.00150)	-0.159** (0.00141)	-0.162** (0.00125)	-0.108** (0.00151)
Shock year+4	-0.192** (0.00162)	-0.182** (0.00152)	-0.185** (0.00138)	-0.123** (0.00176)
Shock year+5	-0.212** (0.00178)	-0.201** (0.00168)	-0.203** (0.00153)	-0.133** (0.00212)
Shock year+6	-0.233** (0.00193)	-0.222** (0.00182)	-0.221** (0.00164)	-0.145** (0.00242)
Shock year+7	-0.253** (0.00209)	-0.243** (0.00196)	-0.240** (0.00176)	-0.156** (0.00273)
Shock year+8	-0.271** (0.00243)	-0.261** (0.00229)	-0.256** (0.00203)	-0.164** (0.00311)
Shock year+9	-0.282** (0.00266)	-0.273** (0.00250)	-0.266** (0.00219)	-0.166** (0.00340)
Shock year+10	-0.292** (0.00308)	-0.285** (0.00291)	-0.275** (0.00249)	-0.168** (0.00377)
Time variables	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No
Individual fixed	No	No	Yes	Yes
Calender time	No	No	Yes	Yes
Time fixed	No	No	Yes	Yes
Linear trends	No	No	No	Yes
# observations	17,679,410	16,688,491	17,679,410	17,679,410

Note: The outcome is yearly labor earnings divided by the mean earnings in the control group. Controls include gender, marital status, number of kids in different age groups, level of education, immigrant status, age, residence municipality, and sector of employment (2 digits). Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

6.2 Heterogeneous effects by level of education and age

We now turn to the main purpose of our paper, which is to investigate heterogeneous responses to health shocks by level of education and age. Note that in all models in this subsection we control for individual fixed effects as well as flexible linear trends by treatment status and level of education. We run separate regressions for the age groups 30-39, 40-49 and 50-59, and measure the difference in the effects between low (no university education) and high educated (some university education) using an interaction effect. The results from this exercise are shown in *Table 4*. From these results we can see important heterogeneous effects. In the year of the health shock, labor earnings decline by between 5 and 6 percent for high educated in all three age groups. The effect for low educated is almost twice as large in all age groups (between 9 and 12 percent). This means that individuals from lower socioeconomic groups not only suffer from more frequent health shocks but also suffer disproportionately hard for a given health shock.

Beside these clear, short-term, differences in effects across groups, there are also some interesting time patterns. For the youngest group the difference between the high- and low educated decreases with time after the shock. During the shock year the effect on earnings is almost 80 percent larger among the low educated. After two years this difference has decreased to 40 percent, and five years after the shock, the effect is basically the same for high and low educated. For individuals in middle and old ages the time pattern is completely different. At these ages the difference in the effect for high- and low educated increases with time since the shock. For individuals in middle and old ages the effect in the shock year for low educated is about twice the effect for high educated. Five years after the shock the effect is almost three times as large among the low educated in the 40-49 group, and almost seven times as large in the 50-59 age group.

Table 4: Estimates of heterogenous effects by level of education and age

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

Note: The outcome is yearly labor earnings divided by the mean earnings among the non-treated with high respectively low education. High education is defined as some kind of university education and low education less than university education. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

6.3 Heterogeneous effects by type of health shock

Table 5 presents separate estimates for our ten major types of health shocks. As in the previous section, we show separate estimates by level of education and age. The first, second and third panel present the effect on yearly earnings in the shock year, two years after the shock and five years after the shock, respectively. In order to make the presentation of the results more transparent, we do not report standard errors, and only indicate significance using stars.

As expected, the main effects differ to a quite large extent across types of health shocks. The difference in the effects between the low and high educated are, however, surprisingly similar across all types of health shocks. In order to illustrate this pattern, we have produced a number of figures (*Figure 6* and *Figure 7*), which show how the low educated interaction effect evolves over time in different age groups and for different health shocks. These figures show that in the short-run, low educated experience more severe effects on labor earnings in all age categories and for all health shocks. In the long-run, the difference between high- and low educated in the youngest age group disappears for almost all types of health shocks. For the two older age groups we see the opposite pattern, with an increasing difference between high- and low educated over time for almost all types of health shocks. This means that the pattern we documented in the previous subsection show up for almost all types of health shocks. We view this as an very interesting insight.

Table 5: Estimates of heterogenous effects by type of health shock, level of education and age

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
<i>Shock year</i>						
Infectious	-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**
Mental	-0.158**	-0.0311*	-0.152**	-0.00802	-0.142**	-0.0128
Nerve system	-0.0728**	-0.0343	-0.0936**	-0.0180	-0.0626**	-0.0620**
Heart diseases	-0.0382*	-0.0569**	-0.0807**	-0.0623**	-0.0984**	-0.0664**
Respiratory	-0.0198	-0.0197	-0.0379**	-0.0155	-0.0376**	-0.0404**
Digestive organs	-0.0220**	-0.0282**	-0.0219**	-0.0379**	-0.0214**	-0.0388**
Musculoskeletal	-0.0623**	-0.0940**	-0.0861**	-0.0792**	-0.0909**	-0.0778**
Genitourinary	-0.00868	-0.0317**	-0.0104	-0.0347**	-0.00863	-0.0336**
External	-0.0529**	-0.0686**	-0.0632**	-0.0702**	-0.0673**	-0.0708**
<i>Shock year+2</i>						
Infectious	-0.0246	-0.0654*	-0.0272	-0.0378	-0.0192	-0.0552*
Cancer	-0.129**	-0.0173	-0.0386**	-0.0572**	-0.0435**	-0.107**
Mental	-0.209**	-0.0524*	-0.147**	-0.0654**	-0.125**	-0.0574**
Nerve system	-0.0602	-0.0940	-0.113**	-0.0443	-0.0510*	-0.120**
Heart diseases	-0.0504*	-0.0486	-0.0770**	-0.0748**	-0.116**	-0.0950**
Respiratory	-0.00605	-0.0524	-0.0373	-0.0321	-0.0222	-0.0841**
Digestive organs	-0.0492**	-0.0170	-0.00141	-0.0612**	0.0112	-0.0739**
Musculoskeletal	-0.0275	-0.0995**	-0.0417**	-0.0836**	-0.0281	-0.109**
Genitourinary	-0.106**	-0.00413	0.00448	-0.0557**	0.0248	-0.0718**
External	-0.0603**	-0.0462**	-0.0338**	-0.0679**	-0.0146	-0.0864**
<i>Shock year+5</i>						
Infectious	-0.0913	-0.0374	-0.0350	-0.0696*	-0.0181	-0.122**
Cancer	-0.123*	-0.0185	0.00386	-0.0821**	-0.0484*	-0.143**
Mental	-0.296**	-0.0204	-0.171**	-0.0650**	-0.0388	-0.0985**
Nerve system	-0.199**	-0.0160	-0.151**	-0.0689	-0.0331	-0.164**
Heart diseases	-0.100**	-0.0448	-0.110**	-0.0809**	-0.149**	-0.111**
Respiratory	-0.0679	-0.0110	-0.0402	-0.0724	0.00855	-0.130**
Digestive organs	-0.0938*	-0.0110	-0.0139	-0.0848**	0.0229	-0.112**
Musculoskeletal	-0.111**	-0.0394	-0.0458*	-0.110**	-0.0280	-0.119**
Genitourinary	-0.150**	-0.0150	0.0194	-0.104**	0.0743**	-0.128**
External	-0.116**	-0.0280	-0.0441**	-0.0789**	0.00497	-0.118**

Note: The outcome is yearly labor earnings divided by the mean earnings in the non-treated with high respectively low education. We report estimates for the ten most common types of health shocks. High education is defined as some kind of university education and low education less than university education. The models include controls for individual fixed effects, calendar time fixed effect, and time with respect to the shock year fixed effects. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

Figure 6: Difference between high educated and low educated in the effect of health shocks. By age and type of health shock

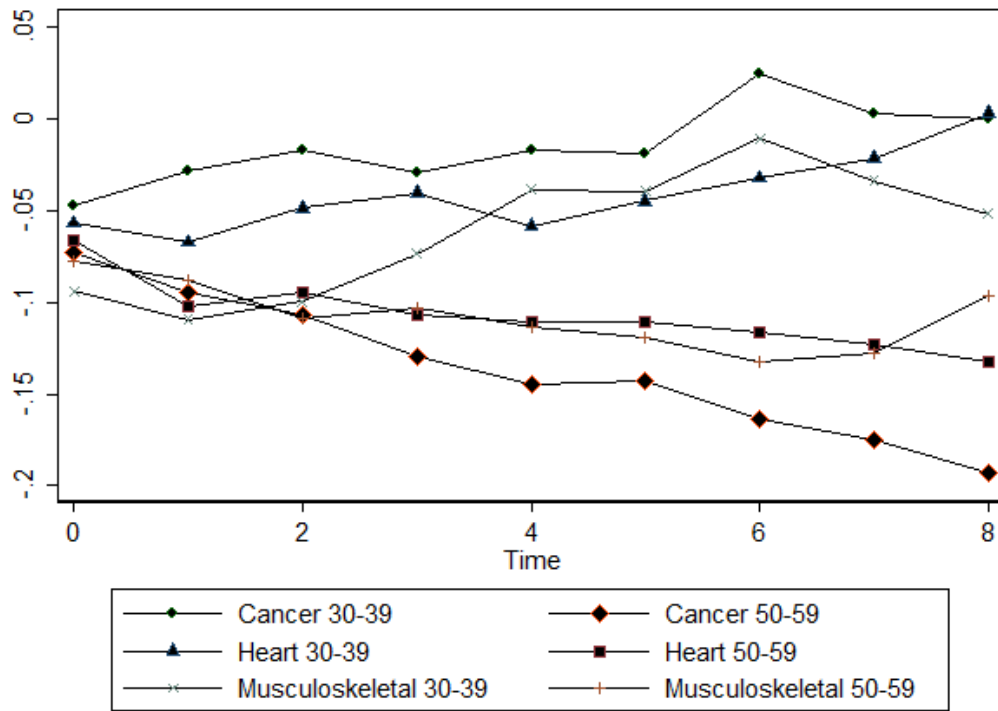
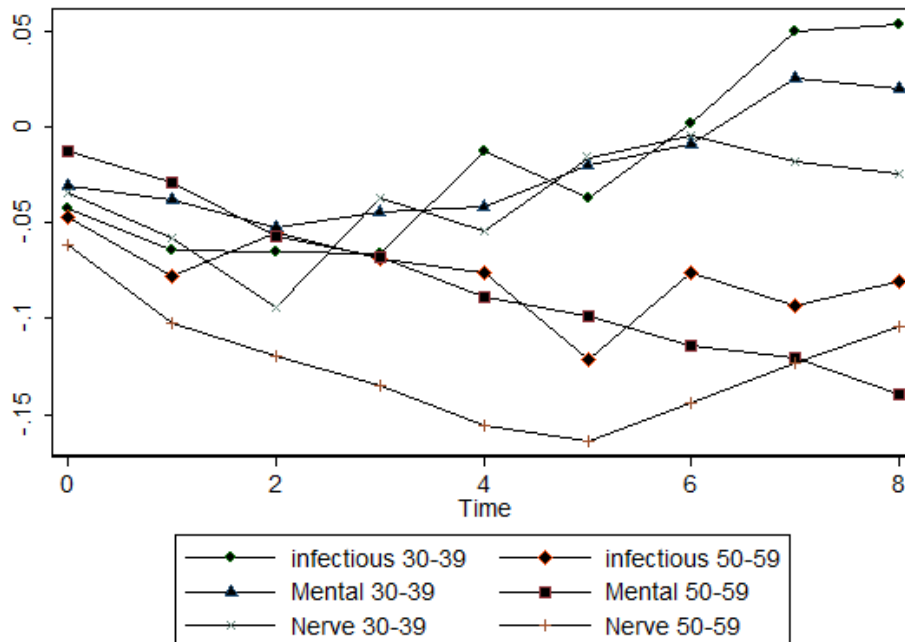


Figure 7: Difference between high educated and low educated in the effect of health shocks. By age and type of health shock



7 Robustness analysis

7.1 Placebo effects

We start our robustness analysis by providing placebo estimates. More specifically, we show placebo estimates, where we move the time of the shock two years back in time, that is before the actual health shock took place. If such placebo effects come out as significant, this serves as an indication that our estimated effects do not represent an effect of the actual health shock but rather the effect of some other time-variant group specific characteristic influencing both the probability to suffer a negative health shock and a decline in labor earnings.

Table 6 presents placebo estimates of both the main effect and the interaction effect that measure the difference in the effects between low and high educated, using our most extended model specification with individual fixed effects and linear trends. These placebo estimates indicate significant pre-treatment effects for both high and low educated. However, all pre-shock effects are small, and the statistical significance could to a large extent be explained by the large sample size. For instance, for ages 50-59 is the placebo estimate of the interaction effect ten times smaller than estimate of the effect in the shock year. Moreover, *Table 7* shows that most pre-treatment effects disappear when we also separate the estimates by type of health shock. In fact, out of 30 estimates of the interaction effect that measure the difference in the effects between low and high educated we find only one that is significant at 5 percent significance level. All these placebo estimates support our empirical strategy and confirm that it is important to have a long panel with extended information both before and after the health shocks. We therefore see no reason to change our main conclusions.

Table 6: Placebo estimates by level of education and age

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

Note: The placebo estimates is created by artificially moving back the treatment two years. The outcome is yearly labor earnings divided by the mean earnings among the non-treated in high and low education group respectively. Standard errors robust to heteroscedasticity and serial correlation. *, * and ** indicate significance at 5 and 1 percent level, respectively.

Table 7: Estimates of placebo effects by type of health shock, level of education and age

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
<i>Shock year-2</i>						
Infectious	0.00777	-0.0159	0.00258	-0.00225	-0.00625	-0.00346
Cancer	-0.0142	-0.00635	-0.00949	0.00417	-0.0136	0.00754
Mental	-0.00867	-0.00297	-0.0138	0.00800	-0.0151	0.0101
Nerve system	-0.0260	0.00722	-0.0102	0.0162	-0.00564	-0.00132
Heart diseases	-0.0186	0.0155	-0.0113	0.00989	-0.0105	0.0112
Respiratory	-0.0278	0.0227	-0.00225	0.00884	-0.0112	0.0117
Digestive organs	-0.00966	-0.000881	-0.00836	0.00700	-0.00297	0.00347
Musculoskeletal	-0.0116	0.00133	-0.0174	0.0112	-0.00168	-0.00128
Genitourinary	-0.00584	-0.00514	0.00237	-0.00699	-0.00651	0.00473
External	-0.0233**	0.0181	-0.0115	0.0161*	-0.00614	0.00410
<i>Shock year</i>						
Infectious	0.00322	-0.0775	-0.0258	-0.0484	-0.0234	-0.0702**
Cancer	-0.140**	-0.0626	-0.123**	-0.0627**	-0.157**	-0.0840**
Mental	-0.180**	-0.0709*	-0.189**	-0.0148	-0.166**	-0.0222
Nerve system	-0.0910*	-0.0462	-0.108**	-0.0102	-0.0693**	-0.0818**
Heart diseases	-0.0454	-0.0532	-0.0938**	-0.0608**	-0.106**	-0.0662**
Respiratory	-0.0453	-0.00700	-0.0408	-0.0152	-0.0405*	-0.0484**
Digestive organs	-0.0213	-0.0478*	-0.0273*	-0.0431**	-0.0249*	-0.0430**
Musculoskeletal	-0.0617*	-0.119**	-0.105**	-0.0871**	-0.0977**	-0.0938**
Genitourinary	-0.00454	-0.0509*	-0.00159	-0.0547**	-0.00896	-0.0388*
External	-0.0694**	-0.0647**	-0.0742**	-0.0686**	-0.0706**	-0.0755**
<i>Shock year+2</i>						
Infectious	-0.00187	-0.112	-0.0218	-0.0543	-0.0190	-0.0862**
Cancer	-0.155**	-0.0382	-0.0549*	-0.0694**	-0.0667**	-0.122**
Mental	0.240**	-0.107**	-0.196**	-0.0752**	-0.157**	-0.0706**
Nerve system	-0.0840	-0.110	-0.132**	-0.0343	-0.0599	-0.146**
Heart diseases	-0.0597	-0.0440	-0.0945**	-0.0730**	-0.127**	-0.0950**
Respiratory	-0.0399	-0.0359	-0.0413	-0.0320	-0.0258	-0.0953**
Digestive organs	-0.0479*	-0.0434	-0.00850	-0.0685**	0.00668	-0.0796**
Musculoskeletal	-0.0263	-0.134**	-0.0666**	-0.0946**	-0.0372	-0.131**
Genitourinary	-0.100**	-0.0302	0.0164	-0.0827**	0.0245	-0.0788**
External	-0.0820**	-0.0416	-0.0484**	-0.0662**	-0.0188	-0.0928**
<i>Shock year+5</i>						
Infectious	-0.0601	-0.102	-0.0275	-0.0925	0.0185	-0.164**
Cancer	-0.159	-0.0476	-0.0185	-0.0994**	0.0167	-0.164**
Mental	-0.338**	-0.0963	-0.239**	-0.0792*	-0.0830**	-0.117**
Nerve system	-0.231**	-0.0392	-0.177**	-0.0557	-0.0450	-0.200**
Heart diseases	-0.112	-0.0389	-0.134**	-0.0787*	-0.163**	-0.111**
Respiratory	-0.114	0.0112	-0.0456	-0.0725	0.00391	-0.146**
Digestive organs	-0.0918	-0.0475	-0.0234	-0.0951**	0.0168	-0.120**
Musculoskeletal	-0.109*	-0.0873	-0.0799*	-0.126**	-0.0405	-0.150**
Genitourinary	-0.142**	-0.0512	0.0359	-0.141**	0.0741*	-0.138**
External	-0.145**	-0.0221	-0.0640*	-0.0771**	-0.000669	-0.127**

Note: The outcome is yearly labor earnings divided by the mean earnings in the non-treated with high respectively low education. We report estimates for the ten most common types of health shocks. High education is defined as some kind of university education and low education less than university education. The models include controls for individual fixed effects, calendar time fixed effect, and time with respect to the shock year fixed effects. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

7.2 DID-matching

In our main analysis, we use a flexible specification with an extended set of fixed effects and controls for differences in underlying trends. One alternative to this specification is to use a Difference-in-Differences (DID) matching approach (see e.g. Heckman et al. 1997). This provides a non-parametric estimate of the same parameters as in our main model. Another potential benefit is that the treatment and control group becomes balanced in terms of covariates. In this subsection, we check the robustness of our main results, when using this alternative specification. More specifically, we first perform a one-to-one propensity score matching for each shock year using all our observed covariates.¹³ For this matched sample we then compare the difference in labor earnings two years before the shock with the labor earnings in the shock year, one year after the shock and so on. *Table 8* presents DID-matching estimates of the short-run and long-run effects. These estimates are very similar to our main estimates, including the increase in effect over time. If anything the effects are in general somewhat larger in the short-run and somewhat smaller in the long-run compared to the estimates from our main empirical strategy. All in all, we believe the results using the DID matching approach support our main conclusions.

¹³Note that all covariates are measured one year before the shock.

Table 8: DID-matching estimates of the short-run and long-run effects of health shocks

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
Shock year	-0.0269** (0.00429)	-0.0745** (0.00433)	-0.0631** (0.00320)	-0.0402** (0.00329)	-0.0805** (0.00332)	-0.0272** (0.00339)
Shock year+2	0.0189* (0.00584)	-0.141** (0.00590)	-0.0357** (0.00366)	-0.0806** (0.00377)	-0.0612** (0.00382)	-0.0634** (0.00393)
Shock year+5	0.0562** (0.00838)	-0.217** (0.00846)	-0.0314** (0.00448)	-0.132** (0.00459)	-0.0503** (0.00463)	-0.111** (0.00468)

Note: The outcome is the difference in yearly labor earnings between 2 years before the shock and the current year divided by the mean earnings among the controls. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

7.3 Detailed measure of education

In the analysis so far, we focused on heterogeneity in the impact of health shocks, using only two educational groups; individuals with and without university education. Do our conclusions change if we use a more detailed measure of education? To investigate this, we next present estimates using a more detailed measure of education, where five different educational groups are indicated.¹⁴ The estimates of the effects for each educational group and for each age group are presented in *Table 9*. For all educational groups and for all age groups, except for individuals with long university education aged 30-39, there is a large initial drop in the shock year. In the long-run we see no difference between the high school educated and those with a university education in the youngest age group, whereas for individuals in middle and old ages the difference between the lowest and highest educated widens with time since the health shock. That is the same patterns as in our main analysis and our conclusions do thus not change.

¹⁴The five groups are individuals with no high school education, short high school education (2 years or less), long high school education (3 years), short university education (2 years or less) and individuals with long university education (more than 2 years).

Table 9: Estimates of heterogenous effects using a detailed measure of level of education

	Main	Level 2	Level 3	Level 4	Level 5
<i>Age 30 -39</i>					
Shock year	-0.120**	0.0177**	0.0921**	0.0204**	0.130**
Shock year+2	-0.151**	0.0314**	0.117**	0.0178*	0.157**
Shock year+5	-0.209**	0.0505**	0.160**	-0.00412	0.180**
<i>Age 40 - 49</i>					
Shock year	-0.116**	0.0137**	0.0445**	0.0415**	0.0798**
Shock year+2	-0.131**	0.0254**	0.0701**	0.0655**	0.112**
Shock year+5	-0.178**	0.0408**	0.0998**	0.0894**	0.154**
<i>Age 50 - 59</i>					
Shock year	-0.127**	0.0121**	0.0481**	0.0543**	0.0763**
Shock year+2	-0.146**	0.0210**	0.0758**	0.0907**	0.121**
Shock year+5	-0.179**	0.0342**	0.0990**	0.123**	0.175**

Note: The outcome yearly labor earnings divided by the mean earnings among the non-treated in each educational and age group. The excluded category is individuals with no high school education. Level 2 to 5 is short high school education (2 years or less), long high school education (3 years), short university education (2 years or less) and individuals with long university education (more than 2 years). The models include controls for individual fixed effects, calendar time fixed effect, and time with respect to the shock year fixed effects. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

7.4 Analysis using survivals

One potential threat to our estimates is that some individuals die during our observation period. If the fractions that die differ in a systematic way across treatment groups, level of education and age, this may bias our estimates.¹⁵ In order to try to assess the extent of this potential problem, we can examine the fraction of individuals that dies within our observation period for the different groups. These sample statistics, presented in *Table 10*, indicate large differences in the survival rate. Low educated and old individuals who experience a health shock are more likely to die. We have therefore re-estimated our main heterogeneous effects model, now only including individuals that survive the *entire* observation period. The estimates from this exercise are presented in *Table 11*. Again, we find that in the long-run the difference between low and high educated disappears for the younger group, but widens for the two older cohorts. Our conclusions thus remain unaffected.

Table 10: Sample statistics on fraction of deaths

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Low edu.	High edu.	Low edu.	High edu.	Low edu.	High edu.
Control (%)	0.49	0.29	1.08	0.62	3.09	1.90
Shock (%)	3.25	2.26	7.51	5.49	13.55	9.98
Total (%)	1.74	1.04	4.48	2.88	9.42	6.43
Observations	264,007	107,643	300,638	128,937	317,419	104,937

Note: The table reports the fraction in our analysis sample that dies within the observation period.

¹⁵Note, however, that if low-educated people are more likely to die following a health shock, this may result in estimates that underestimate the difference in effects across low and high-educated groups. The reason is that the fraction of "frail" individuals in the low-educated group will decrease over time.

Table 11: Estimates of heterogenous effects using only those that survives throughout the entire observation period

	Age 30 -39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
Shock year	-0.0508** (0.00272)	-0.0416** (0.00285)	-0.0487** (0.00201)	-0.0464** (0.00207)	-0.0506** (0.00197)	-0.0508** (0.00196)
Shock year +2	-0.0717** (0.00498)	-0.0391** (0.00513)	-0.0352** (0.00314)	-0.0628** (0.00325)	-0.0306** (0.00317)	-0.0809** (0.00318)
Shock year +5	-0.132** (0.00982)	-0.0186 (0.00509)	-0.0488** (0.00533)	-0.0835** (0.00547)	-0.0188** (0.00527)	-0.119** (0.00520)
Time variables	Yes		Yes		Yes	
Controls	Yes		Yes		Yes	
Individual fixed	Yes		Yes		Yes	
Calendar time	Yes		Yes		Yes	
Time Fixed	Yes		Yes		Yes	
Linear trends	Yes		Yes		Yes	

Note: The outcome is yearly labor earnings divided by the mean earnings among the non-treated with high respectively low education. High education is defined as some kind of university education and low education less than university education. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

7.5 Severity of the health shocks

The results so far show that low educated individuals face much more severe effects of a health shock, especially at older ages. It suggests that individuals with low education have worse ability to cope with a health shock. An alternative explanation is that individuals with low education experience more intense shocks compared to individuals with high education. If so, we are not really measuring the same thing among groups with low and high education. In order to test for this alternative explanation, we use information on the number of diagnoses and number of medical procedures as a proxy for the severity of the health shock. Needless to say, this does not provide a perfect measure, but it is not farfetched to assume that a secondary (or third) diagnose and the number of medical procedures (e.g. surgeries) may imply a more severe health shock.¹⁶

Table 12 displays descriptive statistics for number of diagnoses (including the main diagnose) and number of medical procedures by level of education and age. We see that the vast majority only have a main diagnose, even though the probability of a secondary diagnose increases with age. Most importantly, there are no differences between the low- and high educated in terms of number of diagnoses. Unlike with number of diagnoses, there does not seem to be a clear age pattern when it comes to number of medical procedures. There are also only small differences between the high- and low educated.

We have also re-estimated our main heterogenous effects model taking the number of diagnoses and number of medical procedures into account. More specifically, we interact our treatment variable with number of diagnoses and number of medical procedures.¹⁷ If including these additional interaction effect significantly change our estimate of the difference in the effect of a health shock between high and low educated then differences in the severity of the health shocks explain an important part of the heterogenous effects. The results from this exercise are presented in *table 13*. For comparison, we have also included the baseline results from *table 4*. We only report the low education

¹⁶For the number of medical procedures, the interpretation is somewhat more difficult, since many medical 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.

¹⁷These variables are coded as dummy variables. The number of diagnoses are 1, 2, 3, or ≥ 4 ; Number of operations are 0, 1, 2, or ≥ 3 . The baseline categories are 1 diagnose and 0 operations.

Table 12: Sample statistics on number of diagnoses and medical procedures

	Age 30-39		Age 40-49		Age 50-59	
	Low edu.	High edu.	Low edu.	High edu.	Low edu.	High edu.
<i># of diagnoses</i>						
1	77.76	78.58	74.47	75.67	69.99	71.50
2	17.15	16.77	19.11	18.48	21.43	20.68
3	3.719	3.445	4.675	4.238	6.181	5.595
≥ 4	1.373	1.203	1.745	1.620	2.399	2.227
<i># of medical procedures</i>						
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	119,224	40,582	158,988	59,770	191,954	58,766

Note: The table reports the fraction of the treated in the analysis sample with a certain number of diagnoses and medical procedures. Number of diagnoses reports includes the main diagnose. Number of medical procedures counts as registered medical procedures for the current hospitalization.

interaction effect, since the main effects in the new models now represents the effect for the baseline category. They are therefore not comparable to our baseline results.

We find that the heterogenous effects estimate of the additional effect for low educated is distinctly similar across all three models. This holds for all age groups and for the short-term as well as the long-term heterogenous effects. Even if these measures are imperfect proxies for the severity of the health shock, these sample statistics show that it is unlikely that the observed heterogeneous effects purely are an effect of differences in the severity of the health shock.

8 What explains the heterogeneous effects?

So far we have documented that health shocks have sizeable and long-lasting effects on labor earnings. As discussed in the introduction and in the background sections, there are several potential explanations to this. First of all, the economic incentive to return to work after a health shock is stronger among the high educated, due to their higher earnings and due the fact that, as in most countries, there are maximum benefit levels in the Swedish social insurance. Hence, high and low educated may have different moral hazard profiles. It could also be the case that highly educated are able to acquire more information and better handle contacts with the health care system, which allow them to get more appro-

Table 13: Estimates of heterogenous effects controlling for number of diagnoses and medical procedures

	Shock year Low edu.	Shock year+2 Low edu.	Shock year+5 Low edu.
Age 30-39			
Baseline	-0.0431**	-0.0415**	-0.0220
#diagnoses	-0.0426**	-0.0409**	-0.0213**
#operations	-0.0426**	-0.0404**	-0.0206
Age 40-49			
Baseline	-0.0484**	-0.0674**	-0.09000**
#diagnoses	-0.0477**	-0.138**	-0.0886**
#operations	-0.0483**	-0.0661 **	-0.0881**
Age 50-59			
Baseline	-0.0544**	-0.0876**	-0.125**
#diagnoses	-0.0533**	-0.0861**	-0.124**
#operations	-0.0546	-0.0871**	-0.125**
Time variables	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Individual f.e	Yes	Yes	Yes
Calendar time	Yes	Yes	Yes
Time f.e	Yes	Yes	Yes
Linear trends	Yes	Yes	Yes

Note: The outcome is yearly labor earnings divided by the mean earnings among the non-treated with high respectively low education. Each row reports results from one specification. The specification with number of diagnoses includes an interaction effect between our treatment effect and number of diagnoses number of diagnoses (1, 2, 3, or ≥ 4), and the number of operations specification includes and an interaction between the treatment effect and number of medical procedures (0, 1, 2, or ≥ 3). High education is defined as some kind of university education and low education less than university education. Standard errors are robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

priate treatment. Moreover, low educated are to a higher extent employed in physical / blue collar occupations. We find it plausible that a decline in health complicates the return to work for these individuals compared to those within white collar professions due to the character of the jobs.

In the previous section we found no significant differences by level of education in the severity of the health shocks. Given this fact and if differences in economic incentives are an important factor we expect to find that individuals with high education more quickly leave the hospital and return to work. For that reason we test the economic incentives explanation by including interaction effects between the length of the stay at the hospital (1-5 days, 6-10 days, 11-15 days, 16-20 days, and ≥ 21 days). If the heterogenous effects

changes when including this interaction economic incentives offers one important explanation to our heterogeneous effects estimates. To test whether the result could be explained by differences in the quality of the treatments we interact the treatment with an indicator for whether the treating hospital was an university hospital. In Sweden and many other countries the university hospitals usually are the most advanced hospitals in each region. It is therefore not farfetched to assume that university hospital status provides a quite good measure of the quality of the treatment. Finally, in order to test whether the result could be explained by sector of occupation we include an interaction effect between the individuals sector of employment the year prior to the shock and the shock.¹⁸ If the heterogeneous effect diminishes or disappears with this specification, sector of employment is an important mechanism behind the previously found results.

The results from these different estimations are presented in table *Table 14*. The top, middle and last panel show the effect on yearly earnings for the age groups 30-39, 40-49, and 50-59, respectively. For comparison are the baseline results from table *Table 4* also included in the table. These new estimates are in most cases similar to the baseline estimates, including the long-run differences across age groups. Given this and the fact that the other results are in line with the baseline estimates we find no support that different moral hazard profiles, differences in treatment quality or differences in sector of employment could explain the observed heterogeneous effects.

¹⁸We use the Statistic Sweden's *Svensk Näringsgrensindelning* on a two digit level as indicator of sector of employment.

Table 14: Estimates of heterogenous effects controlling for length of stay, type of hospital and occupation

	Shock year Low edu.	Shock year+2 Low edu.	Shock year+5 Low edu.
Age 30-39			
Baseline	-0.0431**	-0.0415**	-0.0220
length of stay	-0.0417**	-0.0402**	-0.0202
university hosp.	-0.0433**	-0.0414**	-0.0214
occupation	-0.0426**	-0.0465**	-0.0356**
Age 40-49			
Baseline	-0.0484**	-0.0674**	-0.0900**
length of stay	-0.0172**	-0.0658**	-0.0880**
university hosp.	-0.0490**	-0.0677**	-0.0900**
occupation	-0.0578**	-0.0813**	-0.112**
Age 50-59			
Baseline	-0.0544**	-0.0876**	-0.125**
length of stay	-0.0506**	-0.0840**	-0.122**
university hosp.	-0.0553**	-0.0881**	-0.125**
occupation	-0.0616**	-0.0892**	-0.115**
Time variables	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Individual f.e	Yes	Yes	Yes
Calendar time	Yes	Yes	Yes
Time f.e	Yes	Yes	Yes
Linear trends	Yes	Yes	Yes

Note: The outcome is yearly labor earnings divided by the mean earnings among the non-treated with high respectively low education. Each row reports results from one specification. Each specification includes an interaction between the health shock indicator and either the length of stay in the hospital (1-5 days, 6-10 days, 11-15 days, 16-20 days, and ≥ 21 days), admission to University hospital or sector of employment (two digit level) and shock, respectively. High education is defined as some kind of university education and low education less than university education. Standard errors are robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

There are many mechanisms through which a health shock may affect earnings. Individuals that experience a health shock may for instance retire entirely from the labor market, decrease their working time, stay on sickness absence or disability insurance benefits, become unemployed and/or experience a decreased wage rate. In order to obtain some insights on how health shocks affect labor earnings we therefore perform separate estimates of heterogeneous effects of health shocks on the take-up rate of unemployment insurance benefits (UI), sickness insurance benefits (SI), of disability insurance benefits (DI), respectively.

The results from this exercise are presented in *Table 15*, where the top, middle and bottom panels show the effect on the take-up rate of SI, DI, and UI. We find evidence of heterogeneous effects for all types of benefits. Starting with the SI, we see that the take-up rate is between 10 and 30 percent higher for the low educated compared to the high educated during the shock year. Five years after the shock the difference between high and low educated has increased in the youngest group and decreased in the 40-49 age group. For the oldest age group there is basically no difference between high- and low educated five years after the shock. The heterogeneity in effects on SI take-up, thus, appears to be the opposite of what we found for labor income. The effect on the take up rate of DI is, in all age groups, more than 2 times as high for low educated compared to high educated in the shock year. With time this difference increases for all ages except for the 50-59 group. Again, this is a very different time pattern compared to the heterogeneous effects for labor income. The heterogeneous effects estimates for UI are all small and could hardly explain the observed differences between high- and low educated in the impacts on labor income.

All in all this means that the effect on labor income partly is compensated by increased income from the social insurances. None of the heterogeneous effects for SI, DI and UI could, however, explain the striking differences across level of education and age groups that we found for labor income.

Table 15: Estimates of heterogenous effects on social insurance usage

	Age 30 - 39		Age 40 - 49		Age 50 - 59	
	Main	Low edu.	Main	Low edu.	Main	Low edu.
SI						
Shock year	0.274** (0.00293)	0.0633** (0.00324)	0.327** (0.00251)	0.0231** (0.00276)	0.340** (0.00257)	0.0307** (0.00272)
Shock year+2	0.00640 (0.00323)	0.0769** (0.00355)	0.0361** (0.00285)	0.0270** (0.00309)	0.0468** (0.00301)	0.0229** (0.00311)
Shock year+5	-0.0545** (0.00454)	0.109** (0.00494)	-0.0186** (0.00399)	0.0157** (0.00429)	-0.0371** (0.00410)	0.00842 (0.00418)
DI						
Shock year	0.00126** (0.000370)	0.00266** (0.000462)	0.00436** (0.000499)	0.00343** (0.000600)	0.00687** (0.000774)	0.00768** (0.000838)
Shock year+2	0.0120** (0.000819)	0.0144** (0.00102)	0.0258** (0.00103)	0.0211** (0.00125)	0.0465** (0.00152)	0.0449** (0.00169)
Shock year+5	0.0254** (0.00131)	0.0337** (0.00163)	0.0514** (0.00163)	0.0467** (0.00197)	0.0819** (0.00238)	0.0711** (0.00258)
UI						
Shock year	-0.0119** (0.00170)	0.0183** (0.00188)	-0.00508** (0.00115)	0.00966** (0.00126)	-0.00483** (0.000989)	0.00582** (0.00100)
Shock year+2	-0.00980** (0.00252)	0.0176** (0.00281)	-0.00111 (0.00177)	0.00610* (0.00194)	-0.00728** (0.00155)	0.00191 (0.00158)
Shock year+5	0.000328 (0.00359)	-0.000417 (0.00402)	0.00921** (0.00250)	-0.0130** (0.00275)	-0.00718* (0.00220)	-0.00813** (0.00223)
Time variables	Yes		Yes		Yes	
Controls	No		No		No	
Individual f.e	Yes		Yes		Yes	
Calendar time	Yes		Yes		Yes	
Time f.e	Yes		Yes		Yes	
Linear trends	Yes		Yes		Yes	

Note: The outcome is the take-up of sickness insurance, disability insurance and unemployment insurance, respectively. Standard errors robust to heteroscedasticity and serial correlation. * and ** indicate significance at 5 and 1 percent level, respectively.

9 Conclusions

In this paper we have used a unique data set to estimate the effect of health shock on labor earnings. Our large-scale register data set covered the entire population of Swedish workers over a period of 15 years, which gives our results an unusually high degree of representativeness. Using panel fixed-effect techniques, and focusing on unexpected health

shocks, we argued that we came close to estimating the causal impact of health shocks on earnings. We find that negative health shocks have sizeable and long-term effects on labor earnings. This largely confirms the results from recent studies (see e.g. Riphahn 1999, Au et al. 2005, Disney et al. 2006, Gómez & Nicolás 2006, Smith 1998, 1999), which show that a non-negligible part of the relationship between income and health could be explained by the effect of health shocks on income. From a policy perspective this is an important conclusion.

Our main contribution to the literature, however, is to document important heterogeneous responses to health shocks by socioeconomic status and age. We find that individuals with low education suffer disproportionately hard from a given health shock. In the year of the health shock, the effect of the shock is almost twice as large for individuals without a university education compared to individuals with a university education. This result holds for all age groups. This substantiate previous findings in the literature regarding the importance of education for health production. For instance, Goldman & Smith (2002) show that more educated people better adhere to medical treatments for AIDS and diabetes, and Rosvall et al. (2008) finds that highly educated receives better treatment for a given health problem.

In the long-run, we find important differences across age groups. For the youngest cohorts, age 30-39, the difference between high and low educated diminishes over time, whereas the difference increases over time for the older cohorts, 40-49 and 50-59. This shows that individuals with low education suffer most at middle and old ages. This is in line with the cumulative advantage hypothesis, that some mediators of the socioeconomic status and health relationship (e.g. smoking or social capital) accumulates over the life cycle (see e.g. Ross & Wu 1996, Lynch 2003). In addition, at middle ages, individuals with high education have accumulated large economic resources that make them better equipped to cope with health shocks.

The fact that individuals from lower socioeconomic groups suffer disproportionately hard, especially at middle and old ages, offers one explanation as to why the socioeconomic gradient in health widens during middle ages. Our results show that at least parts of this widening is due to the fact that individuals in lower socioeconomic groups, who

experience a health shock, loses a disproportionately large amount of income. This makes them fall further down in the income distribution, which further strengthens the socioeconomic gradient in health during middle ages. Our findings also supports the arguments by e.g. Case & Deaton (2005a) and van Kippersluis et al. (2009) that in order to understand the socioeconomic gradient in health we need to take a life-cycle perspective.

What could then explain the heterogeneity in the impact of health shocks? We have attempted to test whether different moral hazard profiles differences in treatment quality and/or different sector of employment are important explanations. Different moral hazard profiles are potentially important since high educated face stronger economic incentives to quickly return to work. This is because of their higher earnings and because of the maximum benefit levels in the Swedish social insurance. Highly educated individuals may also be able to acquire more information and better handle contacts with the health care system, and thereby get more appropriate treatment. Finally, are low educated to a higher extent employed in physical / blue collar occupation, for which it in many cases are more complicated to return to work. Our estimates suggests that none of these three explanations seem to explain a major part of the observed heterogenous effects. We should, however, stress that none of our test offers conclusive evidence. We view this as an important area for future research.

We believe that the existence of large heterogeneity in the impact of health shocks on labor outcomes is an important insight for policy-makers. As discussed in the introduction, such heterogeneity means that policy advice that is based on average estimates may be severely misguided. Our results show that there may be gains in considering heterogenous effects, for instance, when evaluating new medical technologies and treatments, where the outcome measures include economic outcomes, such as income. By considering heterogenous effects, it may be possible to identify groups where the treatments have a beneficial cost-benefits ratio, whereas average effects may mask such heterogeneity in the cost-benefits ratios. Targeted interventions towards such groups may thus lead to a more efficient use of health care resources.

References

- Antonovsky, A. (1967), 'Social class, life expectancy and overall mortality', *The Milbank Memorial Fund Quarterly* **45**(2), 31–73.
- Au, D. W. H., Crossley, T. F. & Schellhorn, M. (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. & Deaton, A. (2005a), 'Health and wealth among the poor: India and south africa compared', *American Economic Review* **95**(2), 229–233.
- Case, A. & Deaton, A. S. (2005b), Broken down by work and sex: How our health declines, in 'Analyses in the Economics of Aging', NBER Chapters, National Bureau of Economic Research, Inc, pp. 185–212.
- Currie, J. & Madrian, B. C. (1999), Health, health insurance and the labor market, in O. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics', 1 edn, Vol. 3, Part C, Elsevier, chapter 50, pp. 3309–3416.
- Dano, A. M. (2005), 'Road injuries and long-run effects on income and employment', *Health Economics* **14**(9), 955–970.
- 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 Affairs* **21**(2), 13–30.
- Disney, R., Emmerson, C. & Wakefield, M. (2006), 'Ill health and retirement in Britain: A panel data-based analysis', *Journal of Health Economics* **25**(4), 621–649.
- d'Uva, T. B., Doorslaer, E. V., Lindeboom, M. & O'Donnell, O. (2008), 'Does reporting heterogeneity bias the measurement of health disparities?', *Health Economics* **17**(3), 351–375.

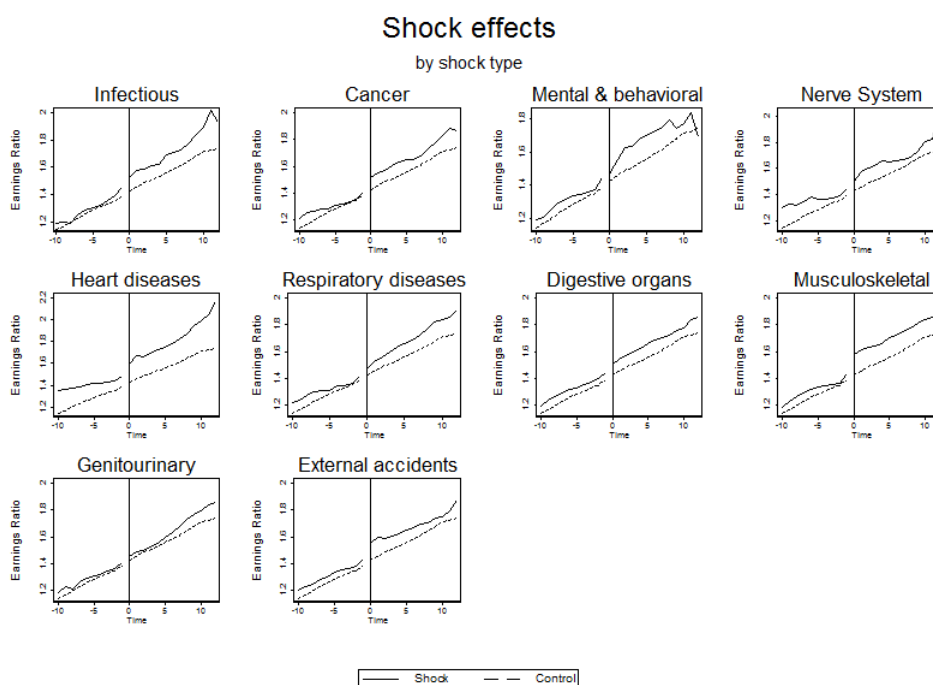
- Etilé, F. & Milcent, C. (2006), 'Income-related reporting heterogeneity in self-assessed health: evidence from france', *Health Economics* **15**(9), 965–981.
- Fredriksson, P. & Johansson, P. (2008), 'Dynamic treatment assignment', *Journal of Business and Economic Statistics* **26**(4), 435–445.
- Goldman, D. P. & Smith, J. P. (2002), 'Can patient Self-Management help explain the SES health gradient?', *Proceedings of the National Academy of Sciences of the United States of America* **99**(16).
- Gómez, P. G. & Nicolás, A. L. (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.
- Heckman, J. J., Ichimura, H. & Todd, P. E. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies* **64**(4), 605–54.
- Lynch, S. M. (2003), 'Cohort and life-course patterns in the relationship between education and health: A hierarchical approach', *Demography* **40**(2), pp. 309–331.
- Mackenbach, J. P. & Bakker, M. (2002), *Reducing inequalities in health: a European perspective*, Routledge.
- Mackenbach, J. P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemstrom, O., Valkonen, T. & Kunst, A. E. (2003), 'Widening socioeconomic inequalities in mortality in six western european countries', *International Journal of Epidemiology* **32**(5), 830–837.

- Marmot, M. (1999), Multilevel approaches to understanding social determinants, *in* L. Berkman & I. Kawachi, eds, 'Social Epidemiology', Oxford: Oxford University Press.
- Peltonen, M., Rosen, M., Lundberg, V. & Asplund, K. (2000), 'Social patterning of myocardial infarction and stroke in sweden: Incidence and survival', *American Journal of Epidemiology* **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.
- Ross, C. E. & Wu, C.-L. (1996), 'Education, age, and the cumulative advantage in health', *Journal of Health and Social Behavior* **37**(1), pp. 104–120.
- Rosvall, M., Chaix, B., Lynch, J., Lindstrom, M. & Merlo, J. (2008), 'The association between socioeconomic position, use of revascularization procedures and five-year survival after recovery from acute myocardial infarction', *BMC Public Health* **8**(1), 44.
- Schrijvers, C. T. & Mackenbach, J. P. (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.
- Smith, G. D., Hart, C., Watt, G., Hole, D. & Hawthorne, V. (1998), 'Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: The renfrew and paisley study', *Journal of Epidemiology and Community Health* (1979-) **52**(6), pp. 399–405.
- Smith, J. P. (1998), 'Socioeconomic status and health', *The American Economic Review* **88**(2), 192–196.
- Smith, J. P. (1999), 'Healthy bodies and thick wallets: The dual relation between health and economic status', *Journal of Economic Perspectives* **13**(2), 145–166.
- Smith, J. P. (2005), Unraveling the ses-health connection, Labor and demography, Econ-WPA.

- Smith, J. P. (2007), 'The impact of socioeconomic status on health over the life-course', *Journal of Human Resources* **42**(4).
- Stewart, J. M. (2001), 'The impact of health status on the duration of unemployment spells and the implications for studies of the impact of unemployment on health status', *Journal of Health Economics* **20**(5), 781–796.
- Tonne, C., Schwartz, J., Mittleman, M., Melly, S., Suh, H. & Goldberg, R. (2005), 'Long-Term survival after acute myocardial infarction is lower in more deprived neighborhoods', *Circulation* **111**(23), 3063–3070.
- van Doorslaer, E., Wagstaff, A., Bleichrodt, H., Calonge, S., Gerdtham, U., Gerfin, M., Geurts, J., Gross, L., Häkkinen, U., Leu, R. E., O'Donnell, O., Propper, C., Puffer, F., Rodríguez, M., Sundberg, G. & Winkelhake, O. (1997), 'Income-related inequalities in health: some international comparisons', *Journal of Health Economics* **16**(1), 93–112.
- van Kippersluis, H., Van Ourti, T., O'Donnell, O. & van Doorslaer, E. (2009), 'Health and income across the life cycle and generations in europe', *Journal of Health Economics* **28**(4), 818–830.
- Willson, A. E., Shuey, K. M. & Elder, Jr., G. H. (2007), 'Cumulative advantage processes as mechanisms of inequality in life course health', *American Journal of Sociology* **112**(6), 1886–1924.

Appendix

Figure A-1: Ratio between mean labor earnings for high educated shock (control) and low educated shock (control). By type of health shock



Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

- 2011:1** Hall Caroline and Linus Liljeberg "En jobbgaranti för ungdomar? Om Arbetsförmedlingens ungdomsinsatser"
- 2011:2** Angelov Nikolay, Per Johansson, Erika Lindahl and Elly-Ann Lindström "Kvinnors och mäns sjukskrivningar"
- 2011:3** Eliason Marcus "Undersköterskor och sjukvårdsbiträden i kristider: inkomst- och sysselsättningseffekter av friställningar inom den offentliga sektorn under 1990-talet"
- 2011:4** Brandén Maria and Sara Ström "För vems skull flyttar par? Kön, karriärmöjligheter och pars regionala rörlighet i Sverige"
- 2011:5** Sjögren Anna and Helena Svaleryd "Nitlott i barndomen – familjebakgrund, hälsa, utbildning och socialbidragstagande bland unga vuxna"
- 2011:6** Mörk Eva "Från försörjningsstöd till arbete – Hur kan vägen underlättas?"
- 2011:7** Forslund Anders and Johan Vikström "Arbetsmarknadspolitikens effekter på sysselsättning och arbetslöshet – en översikt"
- 2011:8** Eliason Marcus, Petter Lundborg and Johan Vikström "Massuppsägningar, arbetslöshet och sjuklighet"
- 2011:9** Lundin Daniela and Martin Lundin "Arbetsförmedlingens service och arbetsmetoder: om processtyrning i en målstyrd myndighet"
- 2011:10** Lagerström Jonas "Vilken betydelse har arbetsförmedlare för arbetslösas framtida sysselsättning och arbetsinkomst?"
- 2011:11** Lundborg Petter, Martin Nilsson and Johan Vikström "Hur påverkar socioekonomisk status och ålder arbetsmarknadseffekterna av olika hälsoproblem?"

Working papers

- 2011:1** Eliason Marcus "Assistant and auxiliary nurses in crisis times: earnings and employment following public sector job loss in the 1990s"
- 2011:2** Forslund Anders, Peter Fredriksson and Johan Vikström "What active labor market policy works in a recession?"
- 2011:3** Brandén Maria and Sara Ström "For whose sake do couples relocate?: Gender, career opportunities and couples' internal migration in Sweden"
- 2011:4** Bergemann Annette, Marco Caliendo, Gerard J. van den Berg and Klaus F. Zimmermann "The threat effect of participation in active labor market programs on job search behaviour of migrants in Germany"
- 2011:5** van den Berg Gerard J., Petter Lundborg, Paul Nystedt and Dan-Olof Rooth "Critical periods during childhood and adolescence: a study of adult height among immigrant siblings"
- 2011:6** Ridder Geert and Johan Vikström "Bounds on treatment effects on transitions"

- 2011:7** Vikström Johan, Michael Rosholm and Michael Svarer “The relative efficiency of active labour market policies: evidence from a social experiment and non-parametric methods”
- 2011:8** Carlsson Mikael and Oskar Nordström Skans “Evaluating microfoundations for aggregate price rigidities: evidence from matched firm-level data on product prices and unit labor cost”
- 2011:9** Carlsson Mikael, Julián Messina and Oskar Nordström Skans “Wage adjustment and productivity shocks”
- 2011:10** Lagerström Jonas ”How important are caseworkers – and why? New evidence from Swedish employment offices”
- 2011:11** Lundborg Petter, Martin Nilsson and Johan Vikström ”Socioeconomic heterogeneity in the effect of health shocks on earnings: evidence from population-wide data on Swedish workers”

Dissertation series

- 2010:1** Johansson Elly-Ann “Essays on schooling, gender, and parental leave”
- 2010:2** Hall Caroline “Empirical essays on education and social insurance policies”
- 2011:1** Hensvik Lena “The effects of markets, managers and peers on worker outcomes”