Effects of Education on Adult Health in Sweden: Results from a Natural Experiment

by

JASMINA SPASOJEVIC

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degreee of

Doctor of Philosophy

The City University of New York

2003

UMI Number: 3103172

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

8/27/03 Michael Grossman

Date — Chair of Examining Committee

8/27/03

Date — Executive Officer

Professor Michael Grossman

Professor Theodore Joyce

Professor Salih Neftci

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Effects of Education on Adult Health in Sweden:
Results from a Natural Experiment

by

JASMINA SPASOJEVIC

Adviser: Professor Michael Grossman

Understanding the determinants of health and the mechanisms through which they affect health is an important social policy issue. Empirical tests in health literature abound with the undisputed finding that years of formal schooling completed is the most important correlate of good health. This finding emerges whether health levels are measured by mortality rates, morbidity rates, self-evaluation of health status, or physiological indicators of health and whether the units of observation are individuals or groups. There is much less consensus as to whether this correlation reflects causality from more schooling to better health. A number of investigators have argued that omitted "third variables" such as a future orientation may cause both schooling and health to rise.

This study uses instrumental variables technique to estimate the causal effect of schooling on adult health outcomes in Sweden. The main aim of this study is to capitalize on a unique social experiment, the 1950 Swedish comprehensive school reform. Between 1949 and 1962, the school system created by the 1950 act was implemented randomly

and in stages. Because of this, persons born between 1945 and 1955 went through two different school systems, one of which implied at least one year of prolonged compulsory schooling. The instrumental variable for schooling generated from compulsory school reform yields a consistent estimate of the causal impact of schooling on various health measures.

The overall finding is that schooling directly and positively causes health as measured by index of bad health and body mass index in healthy range. Additional schooling generated by the compulsory school reform in Sweden produces better adult health outcomes controlling for cohort and regional effects, family background characteristics and individual income.

Acknowledgements

I want to thank some people for their help and assistance, without which this paper would not be possible.

First, I want to thank my advisor Prof. Michael Grossman for all his support and help, starting with getting the idea for the paper, gathering data, providing contacts with Swedish economists, and for his guidance, academic, professional and personal support. I would also like to thank Professors Ted Joyce and Salih Neftci for being on my Dissertation Committee.

My gratitude to Prof. Adriana Lleras-Muney for pointing out the topic and the Swedish data set that was used in Costas Meghir and Marten Palme's work on returns to education in Sweden.

Thanks also to: Marten Palme for his kindness in making available to me his programs and results; Janne Jonsson for being helpful in interpreting data codes, answering my questions and making available necessary information regarding the respondent's place of residence of the Swedish Level of Living Survey (SLLS); Robert Erikson, the Principal Investigator for the SLLS which allowed the SLLS data to be released for my research; Berglund Torbjörn for helpful assistance.

I want to thank Economics Department CUNY Graduate Center for giving me the Gilleece Fellowship and all professors for the fine academic experience I had while taking courses. I want to thank Professors Elizabeth Field-Hendrey and Robert Kaestner for writing recommendation letters for me that helped in getting me a job. I want to thank Prof. Robert Kaestner for expanding my knowledge about labor economics.

I want to thank Prof. Ingmar Nyman and Maja Turniski for helping with translation from the Swedish language.

I want to thank Michael and Sybille Goldberg. They were a great source of strength and support while I was in Durham, NH. I want to thank the University of New Hampshire Economics Department for the financial support that enabled me to come to the U.S. to pursue Ph.D. studies. I want to thank all professors and staff at the UNH Economics Department for unbelievable academic and personal experience.

I want to thank Prof. Steven Pressman for his support and help.

I want to thank the Soros Foundation for the fellowship which enabled me to come to Prague and Budapest to study economics and gave me an opportunity of a lifetime.

I want to thank my classmates John Bridges and Dhaval Dave for their help and support during this unbelievable process of writing a dissertation. Thank you to Cristina Yunzal for proofreading this manuscript.

I want to thank the New York Office of the National Bureau of Economic Research for kind assistance and support.

I want to thank my friends and family for everything and Zokica for being next to me while I was working on my dissertation.

TABLE OF CONTENTS

LIST OF TABLES

The Set of Tables Showing the Analysis of the Total 1981 and 1991 Data Set

Appendix Tables

Introduction

Understanding the determinants of health and the mechanisms through which they affect health is an important social policy issue. Empirical tests in health literature abound with the undisputed finding that years of formal schooling completed is the most important correlate of good health. Traditionally, returns to education are seen as private income returns, that is, higher wages. Returns to education, however, may be reflected in a number of social benefits as well. More educated people are more productive in the market and nonmarket sector. More educated people are likely to have higher earnings. More educated people are likely to be better producers of commodities that enter their utility functions. This means that income returns to education may understate the overall private returns expressed as monetary wage people earn in the labor market and the overall social benefits to education.

One aspect of the nonmarket return to education refers to its impact on health. If more educated people are healthier they will be better producers of both market and nonmarket commodities. It is more likely that these people will live longer and contribute to society more. They will have more time available to produce wages and/or commodities or pursue various productive activities. Hence, the finding that more schooling produces healthier people and reduces morbidity has significant public policy implications.

Schooling is the most important correlate of good health, irrespective of the measure of health employed, such as mortality rates, morbidity rates, self-evaluation of health status, or psychological indicators of health, or whether the units of observation are individuals or groups[1]. Grossman and Kaestner (1997) provide an extensive literature review of the effects of education on health.

However, there is much less consensus as to whether the correlation between health and schooling reflects causality from more schooling to better health. A number of different explanations exist as to why there is an observed correlation between health and schooling. One of the explanations emphasizes that the correlation between health and schooling does not reflect a causal relationship and that this correlation is instead due to omitted "third variables", such as a future orientation that affect schooling and health at the same time and in the same direction.

Initially, most studies of the relationship between education and health employed U.S. data. A number of recent studies have been conducted not only with U.S. data but also with data from European and other countries as well[2]. Here, the Swedish Level of Living Survey (SLLS) is analyzed.

The SLLS is a rich data set for analyzing the relationship between schooling and health at a point in time, as well as how that relationship has changed over time. The Surveys contain variables on the main outcomes of interest, such as the number of years

---

[1] The literature on the correlation between education and health is indeed voluminous and to mention some of the papers dealing with this issue such as Auster et. al., 1969; Grossman, 1975; Wagstaff, 1986; Kemna, 1987; Behrman and Wolfe, 1989; Ross and Wu, 1995; Elo and Preston, 1996; Gilleskie and Harrison, 1998; Hartog and Oosterbeek, 1998; Deaton and Paxson, 1999; Goldman and Lakdawalla, 2001.

[2] For example, Jacob Nielsen Arendt uses Danish data (The Danish National Work Environment Cohort Study) in his paper "Education Effects on Health: A Panel Data Analysis Using School Reforms for Identification", and Damien de Walque used HIV data from Uganda for his paper "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda.

of formal schooling completed and different health measures over time. This study will add new findings to the existing literature in this regard[3].

This study also adds to the body of recent work by using the instrumental variables technique to estimate the causal effect of schooling on adult health outcomes in Sweden. The characteristics of the Swedish compulsory school reform are used to construct the instrumental variable for schooling.

The main aim of this study is to capitalize on a unique social experiment, the 1950 Swedish comprehensive school reform, in order to uncover the causal effect of schooling on health. Between 1949 and 1962, the school system created by the 1950 act was implemented randomly and in stages. Persons born between 1945 and 1955 thus went through two different school systems, one of which implied at least one year of prolonged compulsory schooling. This study investigates whether the increase in compulsory schooling was associated with better health outcomes. The instrumental variable for schooling generated from compulsory school reform yields a consistent estimate of the causal impact of schooling on various health outputs. Health outputs include an index of health symptoms and body mass index variables.

[3] Other researchers also employed the SLLS in their work, such as studies by Gerdtham and Johannesson, Lindahl, Meghir and Palme. Mikael Lindahl, for example, investigates the effect of income on health using the SLLS surveys.

Chapter I

Literature Review

The correlation between schooling and health may be explained in several ways. One explanation is the productive efficiency argument (Grossman, 1972a; Grossman 1972b; Grossman, 2000). If more educated people are more efficient producers of health, then increases in schooling will lead to better health outcomes. In other words, more educated people are better able to obtain better health outcomes from given quantities of health inputs.

Another explanation of the causality from schooling to health is the allocative efficiency argument (Kenkel, 1991; Kenkel 2000). This argument refers to a situation where schooling improves individuals' health knowledge, implying that more educated people are more able to select better health inputs. For example, more educated people are better informed about the harmful effects of smoking and they are more likely to smoke less or stop smoking altogether. Schooling improves people's understanding of the relationship between health and the lifestyles they choose.

The second explanation argues that the relationship between health and schooling is causal, yet the direction of causality runs the opposite way, not from more schooling to better health but rather from better health to more schooling. This reverse causality argument emphasizes the importance of past health, which is reflected in the positive relationship between health and schooling. It is likely that healthier persons will attend school longer and will be more efficient in gaining knowledge.

The third explanation argues that the observed correlation between health and schooling does not reflect a causal relationship. This is most commonly referred to as the time preference hypothesis (Farrell and Fuchs, 1982). Instead, omitted third variables, such as an orientation towards the future, causes both health and schooling to change in the same direction. While this is a valid explanation of the observed correlation between health and schooling, however its difficulty is that time preference is very hard to measure.

The instrumental variables technique can be used to investigate whether the positive correlation between schooling and health reflects causality from the former to the latter. The idea is to predict schooling with variables that are correlated with it but are unlikely to be correlated with omitted determinants of health. The predicted values of schooling replace the actual values in the health equation.

One of the first studies that employed instrumental variables estimation for obtaining consistent estimates of the causal effect of schooling on health is Berger and Leigh (1989). They suggest that the observed correlation between health and schooling is due to the direct effect of schooling on health and that the effect of unobserved third variables is less important. Results strongly indicate that higher schooling has a positive and significant impact on disability, functional limitations and the two measures of blood pressure. However, one problem with this study is that family background variables are employed as instruments for schooling. These variables may be correlated with time preference.

Sander (1995) is another study among the earliest studies that employ the instrumental variables technique. This study estimates the effect of schooling on the odds that smokers quit smoking. The main result is that schooling significantly increases the odds that men and women twenty-five and older quit smoking. However, this study, similarly to the previous Berger and Leigh study, employs instruments that are based on family background characteristics and that may be correlated with omitted third variables.

A significant number of recent studies on the causality from schooling to health employ instruments that are likely to be correlated with schooling but unlikely to be correlated with unobserved determinants of health to obtain consistent estimates of the effects of schooling on health. These instruments include compulsory schooling laws and school reforms, unemployment rates during a person's teenage years, or availability of college openings (Lleras-Muney, 2002; Adams, 2002; Arendt, 2002; Currie and Moretti, 2002; Arkes, 2001)[4]. These studies are supportive of a direct and causal effect of formal schooling on various measures of health.

Overview of Recent Studies

Lleras-Muney (2002) investigates the relationship between education and adult mortality in the U.S. Compulsory education laws implemented from 1915 to 1939 that were state-specific are employed as instruments. For example, Lleras-Muney (2001) demonstrates that for individuals who were 14 years old between 1915 and 1939

[4] In essence this idea comes from the labor literature and econometric issues that pertain to estimation of the wage equations; just to mention some of the studies, such as Angrist and Krueger, 1991; Blackburn and Neumark, 1995; Acemoglu and Angrist, 1999. Extensive work on this topic is done by David Card. One of the most quoted papers on these issues is his chapter on the causal effect of education on earnings in the Handbook of Labor Economics (1999). There also exists a noticeable parallel between estimation of the wage equation with omitted ability ("ability bias") and estimation of the health equation with omitted time preference (Michael Grossman also pointed out this similarity).

compulsory education laws requiring a child to attend school for one more year increased educational attainment by about 5 percent. Since compulsory school laws employed as instruments are likely to be uncorrelated with unobserved determinants of health the effect of education on mortality rates is consistently estimated. This study follows synthetic cohorts compiled from U.S. Censuses of Population for 1960, 1970 and 1980. Lleras-Muney (2002) finds that average education was higher for individuals in states where more education was compulsory and that overall the laws increased educational attainment. Ordinary least squares results suggest that an additional year of schooling lowers the probability of dying in the next ten years by 1.3 percentage points, while instrumental variables estimates show this effect to be larger (3.6 percentage points).

Adams (2002) provides evidence that the positive relationship between schooling and health among older individuals represents an independent direct effect of schooling on health. The sample was restricted to U.S.-born individuals between the ages of 51 and 61 from the first wave of the 1992 Health and Retirement Study. Health measures employed are self-rated health and functional ability. Quarter of birth is used to identify exogenous variation in education with different identification strategies (with and without a set of parental and sibling characteristics). This instrument, employed in two stage least squares estimation, is based on variation in years of schooling that come from compulsory schooling laws in the U.S., which identifies exogenous variation in education as within-year differences in educational attainment by the quarter of birth[5]. The study indicates that healthier outcomes come with higher educational attainment. The OLS

[5] The first study that exploits the nature of compulsory school laws in the U.S. as an exogenous source of variation in years of schooling as an identification strategy is Angrist and Krueger (1991).

approach shows positive and significant effects of education on measures of health for older adults. If education is instrumented, education coefficients become larger.

Arendt (2002) investigates the causal relationship between schooling and health using the Danish National Work Environment Cohort Study for 1990 and 1995. The paper capitalizes on two Danish school reforms that took place in 1958 and 1975 in order to create an instrument for education. The people interviewed were Danish workers between the ages of 18 and 59 in 1990. Arendt obtains separate logit estimates for men and women for self-reported health, body mass index, never-been-smoking indicator and high blood pressure[6].

The simple ordered logit model of self-reported health includes exogenous regressors, such as occupation variables and age. The coefficient on education for men is –0.041 (for women –0.046). This negative and significant coefficient on education is suggestive of better self-reported health with increases in education. Coefficients on education from the two stage least squares, a method that accounts for potential endogeneity of education, are larger in size than simple ordered logit estimates. Arendt could not reject exogeneity of education for men but he could for women in self-reported health regressions. This implies that education is endogenous for women but not for men for this health outcome. However, this result is opposite in case of BMI and never-been-smoking indicator. For both BMI and never-been-smoking indicator, Arendt rejects the null of exogeneity of education for men.

Currie and Moretti (2002) estimate the effect of maternal education on various birth outcomes, such as low birth weight, use of prenatal care, probability of marrying a

---

[6] Arendt does not report estimations for high blood pressure regressions since education effects were small in size and insignificant.

highly educated man, probability of smoking during pregnancy and fertility. The study uses U.S. data on white women from individual birth certificates from the Vital Statistics Natality files for 1970 to 1999. To estimate the causal effect of education on health, the study employs two empirical strategies, longitudinal models estimation and the instrumental variables method.

The instrumental variable is created based on information about new college openings between 1940 and 1990. This is used to construct a variable of availability of two and four-year colleges in the county where a woman lived when she was 17, and used as an instrument for woman's schooling. The authors find that college openings increase maternal education. More specifically, an increase of one four-year college per one thousand persons aged 18 to 22 would result in almost one year more of mother's education among women 24 and over at the time of their first birth. Furthermore, they estimate that an increase in education due to new college openings reduces the risk of low birth weight by about one percent. The IV estimates are approximately double the OLS estimates for almost all outcomes and the estimates by either method are significant.

The main finding is that mother's education has a positive impact on infant's health. This effect is transmitted via channels through which mother's schooling affects child's health. In sum, this study finds that more educated mothers will: (1) have a child that is at a lower risk to be of low birth weight or be a pre-term infant; (2) be more likely to be married and married to a more educated men; (3) they will use prenatal care more and; (4) are less likely to smoke during pregnancy. These findings suggest some of plausible mechanisms via which more schooling causes better child health outcomes.

Currie and Moretti is one of a few studies that demonstrates the validity of their instruments[7].

Arkes (2001) uses intra-state differences in unemployment rates during a person's teenage years as an instrumental variable for schooling and a source of exogenous variation in educational attainment. Higher teenage unemployment rate translates into greater educational attainment since the opportunity costs of attending school are lower at that time. He uses data from the 1990 U.S. Census of Population and includes white males aged 47 to 56 in 1990. He finds from two-stage least squares probit models that an additional year of schooling lowers the probability of having a work-limiting health condition by 2.6 percentage points and lowers the probability of requiring personal care by 0.67 percentage points. These estimates are significant and large relative to the percentage of people who have such conditions. The two-stage probit model coefficients are larger than estimates obtained by the standard probit model.

All the studies employ the instrumental variables technique to obtain consistent estimates of the causal effect of schooling on health. Results are supportive of a direct and causal effect of formal schooling on various measures of health. Common finding is that the IV coefficients are larger than the ordinary least squares coefficients. The IV results are inconsistent with the time preference hypothesis that both schooling and health are determined by some omitted unobserved common factors. This is due to the assumption that instruments are correlated with schooling but are unlikely to be correlated with unobserved variables represented by error term in the health structural equation.

[7] Validity of instruments is measured with high first stage F-statistic especially in regressions with four year college openings when the mother was 17.

This study will complement and take advantage of methods used in recent work on the relationship between education and health. This study investigates whether more schooling induced by the school reform causes better adult health outcomes using the Swedish Level of Living Survey[8].

Other Studies Employing Swedish Data

A number of recent studies investigate the health-schooling correlation and employ SLLS or other Swedish data[9] (Gerdtham and Johannesson, 1999; Gerdtham et. al., 1999; Bolin et. al., 2002a; Bolin et. al., 2002b). Gerdtham and Johannesson (1999) estimate a "Grossman" model of demand for health using the 1991 Swedish Level of Living Survey. They use ordered probit model to estimate the demand for health equation. The dependent variable is stock of health and measured with self-rated health status. Results are that income and education significantly enhance health. Age, being a male, being overweight, living in big cities and being single significantly decreases health. Although the study finds that the effect of education on health is positive and significant, the authors recognize that the lack of instruments prevents further investigation into the issue of endogeneity of education.

Gerdtham et. al. (1999) find results similar to those of Gerdtham and Johannesson (1999). Different measures of health capital are employed, one of which is a measure of health status on a 0 (dead) to 1 (full health) cardinal scale by the rating scale method, and

---

[8] The 1991 SLLS and a joint 1981 and 1991 SLLS set are analyzed in this study.

[9] Bolin et. al., (2002a) employs data from the Swedish biannual survey of living conditions ULF for the years 1980/81, 1988/89 and 1996/97. Gerdtham et. al. (1999) uses

the other is a categorical measure of overall health status on a five-point scale. Gerdtham et. al. also include variables such as alcohol consumption, sport activities, and smoking activities in their estimation of demand for health equation. These represent health inputs and indicators of travel cost and distance when visiting a doctor. This study is based on a random sample of the population in Uppsala county that is not representative of Sweden as a whole, while Gerdtham and Johannesson employ the 1991 SLLS that is a random sample of the entire Swedish population.

Gerdtham and Sundberg (1996) employ the Swedish Level of Living Surveys in a study on measuring income-related health inequalities in Sweden. They use 44 different health conditions to measure health. These are further used to calculate an illness concentration index. They find that there are no income-related inequalities in health in Sweden. However, there are inequalities in diseases and disorders that are unfavorable to the poor. Inequalities in diseases and disorders are more pronounced in 1990s compared with the 1980s.

Work done by Meghir and Palme (1999 and 2001) investigate the effect of schooling on earnings in Sweden. Meghir and Palme (1999) take advantage of the 1950 Swedish compulsory school reform to predict years of formal schooling in the wage equation. They find that persons who went through longer compulsory schooling completed more years of formal schooling and had higher wages compared to those who did not complete more compulsory schooling. This study employs similar research design originated in Meghir and Palme's work with their 1999 study. Their study investigates if the increase in compulsory schooling is associated with better wage outcomes. This study

---

data that was collected in Uppsala County that is located in eastern Sweden north of Stockholm during 1995.

investigates if the increase in compulsory schooling is associated with better health outcomes. Meghir and Palme (1999) include cohort and county indicators in all regressions and construct the instrument for education based on whether the individual was born in a municipality which implemented the reform. They find returns to education of 2.8% with OLS and 3.6% with IV. However, the IV estimates are not significantly different from the OLS coefficients.

Lindahl (2002) investigates the effect of income on health and mortality using monetary lottery prices as an exogenous source of variation in income. He constructs a data set from the SLLS surveys and combines 48 health symptoms into an index of bad health. The SLLS shows if a respondent has ever won money betting in a lottery of any kind. Lindahl uses this information to construct instrument for the endogenous income variable. He finds that higher income enhances health and that this relationship is causal. The IV estimates are three to five times larger than the OLS estimates. Finally, Lindahl concludes that an increase in income by 10% increases good health by 0.01 to 0.02 standard deviations and life expectancy by 5 to 8 weeks.

In summary, this study will complement recent work on the relationship between education and health and take advantage of methods used by previous recent studies. It will investigate whether increase in schooling induced by the Swedish comprehensive school reform causes better adult health outcomes. Using Swedish data in the empirical analysis also adds to the growing body of recent work that studies non-US data.

Chapter II

Empirical Model and Econometric Method

The relationship between schooling and health is empirically investigated by the following system of two equations:

$$\text{Health} = \alpha_1 \text{Schooling} + \alpha_2 X + \varepsilon \qquad (1)$$

$$\text{Schooling} = \beta_1 Z + \beta_2 X + \mu \qquad (2)$$

The coefficient α_1 in the equation (1) represents the causal effect of schooling on health outcomes. Estimation of this effect must account for the potential correlation between the unobserved disturbance term (ε) in the health equation and the corresponding unobserved disturbance term (μ) in the schooling equation. If these terms are correlated, schooling is correlated with ε, and the first equation cannot be estimated by ordinary least squares (OLS).

The presence of variable Z in the schooling equation facilitates estimation of the health equation by the instrumental variables (IV) method. It is required that Z is highly correlated with schooling and uncorrelated with ε, which in turn incorporates unobservable variables in the health equation. This means that variable Z is excluded from the health equation and only belongs in the schooling equation (2).

The system is estimated by the two-stage least squares method. Estimation of the schooling equation by OLS is the first stage. Predicted schooling from the first stage is used in place of actual schooling in the estimation of health equation (1) in the second stage. The vector of variables X represents common determinants of both education and health, such as age or cohort effects, regional effects, gender and family background characteristics.

Chapter III

Swedish Natural Experiment

1. Compulsory Schooling Reform

Sweden is divided into 25 counties, which roughly correspond to cities and towns[10]. There are 1,031 smaller geographical areas called municipalities[11]. Swedish comprehensive schooling reform extended compulsory schooling from 7 or 8 years (depending on the municipality) to 9 years of basic education. It was enacted by the Swedish parliament in 1950. This unique social experiment, reflected by the 1950 compulsory schooling legislation, was conducted gradually over time, between 1949 and 1962, randomly, in stages, and by municipality.

Implementation actually began in 1949 in 14 municipalities. The number of municipalities with the new education system gradually increased over time to 50 percent in 1961. Municipalities had the option to implement the new system from the first grade only or for all grades up to the fifth grade. In 1962 all remaining municipalities adopted the new nine-year comprehensive compulsory schooling. Meghir and Palme (1999) provide an overview of the 1950 Swedish comprehensive school reform. For this research, the key aspect of the 1950 reform is that there is a cohort of people that went through two different school systems, one of which implied at least one year of longer compulsory schooling.

---

[10] The capital of Sweden, Stockholm, accounts for 2 counties.

[11] In 1949 Sweden consisted of 2,501 municipalities. In 1952 this division was changed and Sweden was divided into 1,031 municipalities. Residential variables in this study rely

The cohort of people most affected by the compulsory schooling reform is that born between 1945 and 1955. This cohort went through two different compulsory school systems. For instance, people born in 1945 were six or seven years of age when the new school system was in its first years of implementation. The majority of these children attended the old pre-reform school unless they resided in one of the earliest municipalities to adopt the reform. However, the situation is different for the people born in 1955. They were six or seven years of age by 1961 or 1962 (when the reform was implemented nationwide), and almost all of them attended the new post-reform compulsory schools. Note that in any one year the characteristics of municipalities that adopted reform were representative of Sweden as a whole (Meghir, Palme, 1999). This means that reform status should not be correlated with unobserved variables in the health status equation.

2. Reform Assignment as Instrumental Variable

The instrumental variable for schooling generated from compulsory school reform yields a consistent estimate of the causal impact of schooling on various health outputs. In econometric terminology, the compulsory schooling law serves as an instrument for years of formal schooling completed in health output equations.

An instrumental variable should be highly correlated with the endogenous variable, schooling, and unlikely to be correlated with omitted variables in the structural equation of health. Compulsory schooling reform will be used to predict years of formal schooling completed, which is then substituted in the health equation.

---

on the 1952 division of municipalities.

Other researchers have found that individuals completed more years of formal schooling if they were in the system with longer compulsory schooling. It is also shown that people with more years of formal schooling completed (due to longer compulsory schooling) earn higher wages. This study investigates whether the same effect could be translated to health, that is, if people with more years of formal schooling completed (due to longer compulsory school) have better health.

This is possible if there is a way to allocate people to the new reformed schooling. In this way reform assignment, that an individual was affected by the school reform, is the instrumental variable for years of schooling completed[12]. Computation of this variable requires knowledge of the municipality where individuals lived while they were in their childhood up to age 16, which is available in the surveys; that they were between six (or seven) and thirteen (or fourteen) years of age during the time when the new school system was gradually implemented (which is already assumed under the previous information); and, the year when the municipality where they lived during their childhood implemented the new reformed schooling. It is assumed that people did not move between municipalities during their childhood, or if they did, that migration was not substantial[13].

In summary the reform indicator will be used to predict years of formal schooling completed. The predicted value of schooling will then be employed as a regressor in health outcome equations (weighted symptoms index of bad health, body mass index and body mass index in healthy range).

[12] It is assumed that people born between 1945 and 1955 were affected by the reform if they lived in the municipality that introduced the reformed compulsory school.

Chapter IV

Data Description and Variable Specifications

The Swedish Level of Living Survey (SLLS) is analyzed in this study. The SLLS is conducted in several years starting with 1968, then in 1974, 1981 and 1991[14]. This study is primarily concentrated on the 1991 SLLS survey. However the availability of health symptoms in both 1981 and 1991 prompted the analysis to be extended to a joint 1981 and 1991 set. Each survey represents a random sample of 1/1000 (or 0.1 percent) of the Swedish population between 15 and 75 years of age[15].

The SLLS surveys cover several different dimensions of a person's level of living. These are socio-economic background and individual histories such as conditions while growing up, family structure and social relations, living conditions (i.e. housing), education and schooling opportunities (i.e. highest level of education), health and access to health care, economic resources (i.e. income and wealth), employment, working biography and working conditions, recreation and culture (i.e. leisure time activities) and political participation.

---

[13] Similar assumption is also made in the Costas Meghir and Marten Palme (1999) study. People who moved followed the school system they would have followed if they stayed in the municipality where they lived while growing up.

[14] The SLLS 2000 has been recently conducted.

[15] In 1968 there were 7.9 million people living in Sweden. In the years of the SLLS survey there were respectively 8.2 million (1974), 8.3 million (1981) and 8.6 million (1991) people living in Sweden. In 2002 there are about 9 million people living in Sweden. On December 31, 2001 about 6.6 million people living in Sweden were between ages 15 and 75. (Statistics Sweden, http://www.scb.se).

The key information provided by the SLLS surveys pertains to the respondent's health. End result measures of health or health outputs include different health symptoms or conditions people suffered in the previous year, self-rated health status, and body mass index[16]. Self-rated health status, asked for the first time in 1991, has three response categories: good, bad, and in between. This health outcome is further employed in the construction of index of bad health, which represents the main health outcome of interest in this study due to health symptoms availability in both the 1981 and the 1991 survey.

The health symptoms measure, which was asked in all four surveys, pertains to the presence of a variety of conditions (illnesses and ailments) in the previous year. Information on about fifty different symptoms is obtained. They are diverse and range from simple and minor illnesses such as a cough or cold, or chronic illnesses such as back pain or fatigue, to major health conditions such as heart attack, high blood pressure, diabetes or cancer. A continuous health symptoms measure can be created by using the separate conditions as predictors in an ordered probit specification of self-rated health. The estimated coefficients in this specification serve as weights for the construction of a linear latent health index.

The linear latent health index is standardized with a mean of 0 and a standard deviation of 1. Since the standardized weighted symptoms index of health is based on whether people suffered from any of the conditions in the previous year, it actually represents the index of bad health. Therefore decreases in the index of bad health are associated with better health outcomes. A comparison of different health weights

---

[16] Six major conditions variable is another health outcome created during preliminary analysis. This is a binary variable that incorporates information if people have suffered from the following 6 conditions in the previous year: heart attack, high blood pressure,

estimations in the ordered probit specification of self-rated health on separate health conditions is shown in Table A1 in the Appendix [17].

The index of bad health is the main outcome of interest in the analysis of a joint 1981 and 1991 set. Respondents are interviewed about any of the health symptoms they suffered in the previous year in 1981 and in 1991. Even though self-rated health status question is only asked in 1991, it is possible to create a continuous linear health index in the total set. Estimated health weights from the 1991 set and symptoms questions from the 1981 survey and the 1991 survey are used to linearly predict the index of bad health. It is assumed that health symptom weights did not change significantly between the two surveys.

Measures of functional ability to perform activities such as walking for 100 meters, running for 100 meters, walking up and down stairs and turning the tap on and off are also available in 1991. These functional ability indicators can be included in the index of bad health together with the other conditions and illnesses. Functional ability variables denote that a person has difficulty performing any of the functional activities (walking,

---

diabetes, kidney problems or disease, cancer and if they were overweight. The analysis employing this variable as another health outcome is shown in the Appendix, Table A6.

[17] Mikael Lindahl (2002) uses a standardized index of bad health as the dependent variable in his analysis of the effect of income on health using the SLLS surveys. His health weights estimates are used for comparative reason and presented in the first column in Table A1 in the Appendix. I estimated my own health weights and used them to construct my own index of bad health. Estimations of my health weights in the ordered probit equation of self-rated health on separate health symptoms included a female dummy since it was done on the sample of both males and females. Overall, my own health weights are very similar to Lindahl's except in a few cases when they are not of the same sign; however, they are small in magnitude.

running, walking the stairs and turning the tap). This means that lower values of functional ability denote better functional ability.

Respondents are asked for the first time in the 1991 survey to self-report their height and weight. This information is used to create two health outcomes body mass index and body mass index in healthy range. Body mass index (BMI) is defined as weight in kilograms divided by height in meters squared. Persons with BMI equal to or greater than 25 are classified as overweight, while persons with BMI equal to or greater than 30 are classified as obese. BMI in the healthy range is defined as greater than or equal to 18.5 and lower than 25.

A vector of explanatory variables contains indicators of a person's individual characteristics, socio-economic status and family background. Individual characteristics refer to a person's age, gender and marital status. Socio-economic status includes information on a person's education (self-reported education in the surveys), employment status, hours worked and earnings potential (hourly, monthly and annual individual income in the surveys)[18]. Family background refers to people's living conditions while they were growing up. These include parents' education (father's and mother's educational attainment), if they lived with parents or not, siblings' information, characteristics of residential area (rural, town, major town area), a proxy for family wealth situation (family economic difficulties during childhood) and measures of their past health (if a family member suffered from any health problems).

[18] Annual individual income is derived from the tax registers. Real individual income is employed in regressions deflated by the CPI (Statistics Sweden, http://www.scb.se).

Chapter V

Empirical Analysis and Results

1. Data and Summary Statistics

People between 15 and 75 years of age are interviewed in the SLLS surveys. The cohort of people most affected by the compulsory schooling reform is the cohort born between 1945 and 1955. This cohort is between 36 and 46 years of age in the 1991 survey. Summary statistics pertains to the entire sample (N=6773) and to a subsample of people born between 1945 and 1955 (N=1522) as well as across males and females for both samples (Table 1A). It is interesting to note that mean years of education for the entire sample is 10.9 while it is 11.9 for the sample born between 1945 and 1955. The cohort exposed to compulsory school reform, people born between 1945 and 1955 have one more year of education than the entire sample on average. This holds across sexes as well, and males on average have slightly higher (about 3 months higher) education. This finding is going to be investigated further having in mind that the effect of educational reform on health is in the center of the analysis.

Summary statistics of different measures of health shows that on average the cohort of people born between 1945 and 1955 is of better general health. This cohort has lower body mass index, it is slightly less obese, and it has better general functional ability compared to the overall sample (Table 1B). As to having at least one of the major six health conditions, about 22% of all respondents suffer at least from one of them while only 17% of the cohort of people born between 1945 and 1955. It is also important to note that about 87% of the sample born between 1945 and 1955 were exposed to the school reform according to the predicted (administrative) reform instrument and only 54% according to self-assessed reform instrument (Table 1C).

Simple correlations between different measures of health and self-reported education show that more schooling is negatively associated with index of bad health, BMI, smoking habits and having functional difficulties. Schooling is positively associated with self-reported health status (Table 1D). Distributions of various health measures across educational categories show that people with more years of education generally smoke and drink less, are of better self-reported health and less obese, and walk and run with fewer problems (Table 1E)[19].

2. The Effect of the Compulsory Schooling Reform on Education - The First Stage Results

The Rationale for Instrumental Variables

The time preference hypothesis is based on the premise that there are variables commonly assumed to be unobserved, such as an orientation towards the future, which affect both health and education to rise. If health and education are affected by a common important variable missing from the health equation then estimating the effect of education on health would not account for the effect of that missing variable on health. Our estimates of the effect of education on health will be not only biased, but more

[19] There are five educational categories shown as: seven and less than seven years of elementary education, 8 and 9 years of new-reformed school, 12 years and less of high school, 16 years and less of college education and more than 16 years of schooling. Each measure of health takes a value of 1 if the characteristic is present and 0 if it is not. Similar results hold for the sample of people born between 1945 and 1955.

importantly, inconsistent, ending up in estimating the submerged effect of that missing variable on health under the total effect of education on health. In other words, the total effect of education on health is going to reflect not only the direct effect of education on health but also the effect of the missing variable that is not accounted for. If this is the case, a prori it is believed that the OLS estimate of the effect of education on health is inconsistent. A consistent estimate of this effect is obtained employing the IV econometric technique.

The idea is to separate the effect of education on health from the effect of the omitted variable that affects both education and health on health. Additionally, it should be noted that time preference might depend on schooling. Becker and Mulligan (1997) develop a model with an endogenous determination of time preference. It may also be true that a person that goes to school longer could become more patient and future oriented. It is rather obvious that time preference, education and health are tightly interdependent. However, a practical difficulty in disentangling the effect of education on health is the measurement of time preference, since conventional available data sets usually lack this information. If the time preference variable is available then we could include time preference in the health equation and estimate the health equation with OLS.

The inability to include time preference in the health structural equation point towards the IV method as the adequate methodology. The IV method will ensure that the estimate of the direct effect of education on health is consistent. We need a valid instrumental variable highly correlated with education, but not correlated with health, or with unobserved variables that are correlated with both health and education[20].

[20] An instrument should be highly correlated with the variable suspected of endogeneity once other exogenous variables of the model are included. Further, an instrument should

The First Stage Results – Instrument's Analysis

The SLLS is a source of two possible instrumental variables. Both instruments are constructed based on a person's exposure to the Swedish comprehensive schooling reform. Predicted reform status or administrative instrument is based on whether a person during their childhood lived in a municipality when it implemented the new system of schooling. This instrument predicts a person's schooling reform status, therefore predicted reform status.

Self-assessed reform status instrument is based on self-reported information given by a person about his or her own school attendance and is constructed out of survey questions, which directly ask people if they went to the old pre-reform or the new post-reform school. Self-assessed reform status instrument may be subject to measurement errors since people could be biased in determining which type of school they actually went to. Measurement errors may also arise because the school reform was gradually implemented over longer period of time from 1949 to 1962.

The first stage is the estimation of a reduced form equation for the endogenous explanatory variable education. In this regression self-reported education serves as dependent variable and the instrumental variable as an explanatory variable along with other exogenous variables that are added one at a time in order to test for the robustness of results. Other exogenous variables are cohort and regional dummies and father's education dummies. Cohort dummies are basically year-of-birth fixed effects. Regional dummies are county fixed effects. Father education dummies denote if a person's father

---

be uncorrelated with the error term in the structural equation for health since it serves to replace the endogenous variable suspected of correlation with the error term.

has elementary, vocational, secondary or university education[21]. Father's elementary education is the omitted category. The unit of observation is a sample of men born between 1945 and 1955.

The first stage analysis shows that administrative instrument performs better in predicting years of formal schooling completed than self-assessed reform status (Table 2A). Administrative instrument is highly significant with no other covariates included (t-statistic is 3.66 and P>|t|=0.000), and stays significant even with different covariates added. The effect of the administrative instrument on education is positive, significant and robust to the inclusion of cohort and regional dummies and father's education (t-statistic is 1.95 and significant at 5% level).

The effect is also strong in magnitude. The coefficient of 1.769 implies that people who lived in a municipality that enacted the new schooling system completed about 1.8 years more schooling this compared to completed 0.6 more years of schooling if people self-assessed their reform status in the model with no covariates. The significance of self-assessed reform status decreases in regressions with other covariates. It is noticeable that reform status instruments are correlated with cohort dummies. This correlation, however, is unusual, and especially is obvious in the case of self-assessed reform status. The coefficient on the reform instrument gains in significance (t-statistic increases from 1.91 to 2.18) when cohort dummies are added to the regression.

Additional test of validity of instruments suggested by literature are F-statistic and partial R squared from the first stage estimation (Table 2B). The number of observations available for the administrative instrument is N=437. The self-assessed instrument

---

[21] Father's education dummies serve as a proxy for family background characteristics.

generated a bigger sample (N=548). If a sample is not constrained to equal size then self-assessed reform status has higher partial F-statistic and R^2 compared to administrative instrument. However, when the sample is constrained to be of equal size, that is when N is equal to 437, the partial F-statistic in self-assessed reform status regression significantly falls.

Further if both instruments are included at the same time in the first stage equation, then the administrative coefficient is significant and robust to inclusion of cohort, regional, and father's education dummies (Table 2C)[22]. For example, in the regression with all covariates, the t-statistic on administrative instrument is 1.92 (significant at 5% one-tail test). The size of the coefficient on the administrative instrument does not change that much (nor does its t-statistic) if included with the self-assessed instrument in the reduced form equation for education, compared to regressions where administrative instrument is the only instrument included. In other words, if two instruments are entered together in the equation for education, the self-assessed instrument does not bring any additional significance towards explaining education.

The joint effect of two instruments together may exhibit some weakness since the joint partial F-statistic on the two instruments falls with the addition of other covariates. Even though the administrative instrument performs better in predicting education in the reduced form equation, it loses some of its significance when covariates are added. It should be noted, however, that when cohort and regional dummies are added to the first stage equation, a part of the explanatory instrument's power for education may go away

[22] The correlation of the self-reported education to the administrative instrument is 0.37 and to the self-assessed instrument 0.35. Administrative and self-assessed instruments are correlated 0.80.

since the instrument itself incorporates the same type of information[23]. Table A2 in the Appendix displays an additional sensitivity check with the introduction of a third instrument that is generated as an interaction of administrative and self-assessed reform instrument.

In summary, the administrative instrument outperforms the self-assessed reform status instrument. The former is singled out by diagnostics checks as a better instrument compared to the alternatives.

3. The Effect of Education on Health Outcomes - The Second Stage Results

The end result of the first stage estimation is predicted education which now replaces actual education in the second stage equation. The second stage estimates the structural equation for health. It accounts for possible correlation between the endogenous explanatory variable education and the error term. It is assumed that predicted education is free from any potential correlation with error term. Two stage least squares (TSLS) estimation accounts for endogeneity of education. The result of this estimation is a consistent causal estimate of the effect of education on health outcome. Before turning to TSLS estimation it is important to overview results obtained from the estimation of the health reduced form equation and a set of results from the OLS estimation of health structural equation assuming exogeneity of education.

---

[23] The correlation of the administrative instrument with any given cohort is about 0.10 (there are 10 cohorts), except for the cohort born in 1946 and 1947 with correlation of 0.03. The correlation of the administrative instrument with any given region is between 0.01 to 0.06 (there are 25 regions).

Reduced Form Equation of Health[24]

Essentially, association between the exogenous reform instrument and endogenous health outcome should exist. Estimation of the health reduced form equation has a couple of advantages. It is convenient to use the OLS method of estimation since it is the best method to estimate a linear model. The interpretation of the results obtained by OLS is straightforward. Finally, estimation of the health reduced form equation refers to a model where health outcome is predicted with exogenous reform status variable (and other exogenous variables) that is assumed uncorrelated to omitted third variables.

A reduced form equation of health indicates how well the instrument explains health. In this regression health serves as a dependent variable and the instrumental variable is one of the explanatory variables. The estimation of this equation should indicate if reform status causes better health. Other exogenous variables are added one at the time together with reform status instrument such as cohort and regional dummies, father's and mother's education dummies, if a person lived with parents during childhood and a person's siblings situation. Generally, adding other exogenous variables and employing different measures of health as health outcomes usually represents a test of sensitivity of the results.

Index of bad health regression generates the strongest results (Table 3). Reform status instrument is significant at 5% level one-tailed test with inclusion of other exogenous variables. Administrative instrument is insignificant only in the regression with no covariates and in the regression with cohort dummies. Finally, the administrative

[24] A reduced form equation (a single equation linear model) implies that an endogenous variable is linearly projected onto all exogenous variables of the model. A reduced form equation of health implies that a health outcome is linearly projected onto all exogenous

instrument is significant at 5% level one-tailed test using BMI in healthy range with cohort and regional effects, mother's and father's education, living with parents and siblings' situation included.

Estimation of the Health Structural Equation by OLS Method

In order to obtain a consistent estimate of the direct effect of education on health by OLS, it is assumed that all variables that affect health are accounted for. Therefore all other effects on health, including variables that are not explicitly included in the equation, are jointly assumed under the error term component only if the error term is uncorrelated to variables that are already in the equation. In this case, there is no endogeneity of education and the OLS method can be employed.

It is discussed elsewhere that future orientation is a variable that may still be missing from the health equation. This may potentially rise some doubts about the consistency of the OLS estimate of educational effect on health. In this section, the health structural model is estimated by OLS, assuming no endogeneity of education and taking for instance that family background variables (such as father's and mother's educational attainment, living conditions and economic resources while a person was growing up) could be potential proxies for a person's future orientation. While a child, one learns from one's parents and environment. It may be believed that a person's future orientation is influenced by one's parents and childhood. This may not be entirely true if time preference also depends on schooling.

---

variables and the reform status instrument. This equation is an exactly identified model, with one endogenous variable (education) and one instrument (reform status).

The OLS results will be compared with those obtained by TSLS. Index of bad health, BMI and BMI in healthy range are dependent variables. In all regressions the effect of education on health is positive and significant at 1% level of significance and it is robust to inclusion of other covariates (Table 4). A negative education coefficient in the index of bad health regression means that more education implies better health while a positive education coefficient in the BMI in healthy range regression means that more education increases probability of having a BMI in healthy range and therefore better health[25].

Estimation of the Health Structural Equation by TSLS Method

Two stage least squares (TSLS) estimation accounts for the endogeneity of education. The end result of the health structural equation estimation by this method is a consistent causal estimate of the effect of education on health. Index of bad health, BMI and BMI in healthy range are the same dependent variables as in OLS regressions (Table 5). Educational coefficients obtained by TSLS are larger compared to their OLS counterparts. TSLS coefficients are sometimes also much larger than the counterpart OLS coefficients (3 to 6 times larger), especially with inclusion of other covariates such as cohort and regional effects and father's education[26].

---

[25] Six major conditions is another measure of health that is analyzed (Appendix, Table A6). It is probably the case that this variable does not vary a lot across individuals to reflect their current health status because they are young. (The cohort of people born between 1945 and 1955 is between 36 and 46 in the 1991 survey.)

[26] This is not an unusual result. IV estimates larger than OLS are present in the wage equation estimation literature. The interpretation of this result is that the effect of the extra school year is greater for a marginal person than for an average person (LATE). The marginal person would have stopped going to school if it weren't for the compulsory

The main result is that in the index of bad health regression with the administrative instrument in the first stage, the coefficient on education is significant. The educational coefficient from the second stage with cohort and regional effects is –0.141 (t-statistic is -1.65, is significant at 5% level one-tailed test). In the BMI in healthy range regression, educational coefficient is significant at the 10% level with cohort and regional effects.

It is important to note that the positive effect of education on health remains if either OLS or TSLS is employed. All results indicate that more education implies better health. For instance, negative and significant educational coefficients in the index of bad health or BMI regressions denote that higher education implies lower values for the index, i.e., positive impact on health. Assuming a priori that education is endogenous, then TSLS indicates that a positive significant educational coefficient in the BMI in healthy range regression implies that the effect of education is direct and causal.

The significant educational effect on health in the TSLS regressions remains if other family background variables are included as well (Table 6). There is some evidence, though limited, of the importance of some of the family background characteristics for person's health (Father's and mother's education are significant, however, of a wrong sign. Siblings' coefficient is significant and negative.)

schooling law that imposed longer schooling for that person.

Diagnostics check: Hausman Test of Exogeneity of Education

Based on theory, there are reasons to believe that education is an endogenous variable in the structural model for health. The purpose of the test of exogeneity of education is to determine which estimation method (OLS or IV) is an appropriate estimation model.

The Hausman test is a test of a null hypothesis of exogeneity. The test consists of several steps. First, a reduced form equation for education is estimated where education is the dependent variable and all exogenous variables and instrumental variable are explanatory variables. The administrative reform status is the only instrument considered. Predicted education, the end result of the reduced form first stage estimation, serves as an explanatory variable in the health equation together with actual education and exogenous variables except for instruments.

In order to reject a null of exogeneity, the t-statistic on predicted education should be significant. If this is the case, then there is some part of education in addition to actual education that explains health, since both predicted and actual education are included in the health equation. There thus exists a variable missing from the model affecting both education and health at the same time.

The health structural model coefficients estimated by the OLS method are systematically different from coefficients that are estimated by the IV method if exogeneity is rejected. IV estimation is the method that generates consistent estimates of the effect of education on health if education is found to be endogenous.

The main finding is that overall it is not possible to reject uniformly a null hypothesis of exogeneity of education. This result is sensitive to gender. The Hausman test is conducted separately for men and women. The null of exogeneity is rejected and education found endogenous in index of bad health regression for men, but not for women, with covariates included (Appendix, Table A3). Exogeneity of education is rejected at a p-value of 0.090 in the index of bad health regression for men with cohort and regional effects and family background characteristics (Table 6).

Opposing results are generated with BMI as health outcome. Exogeneity of education is rejected for women but not for men. This indicates that education is endogenous for women and not for men. Education is also endogenous for the sample of men and women together for BMI and it could be a result driven by characteristics of females. Results for BMI in healthy range as health outcome are mixed. Exogeneity of education is rejected for both men and women only for regressions with no covariates.

Conclusions about endogeneity of education are mixed overall. It is not possible to reject exogeneity of education in all regressions. For example, education is endogenous for men in index of bad health regression with exogenous covariates added but not in the BMI in healthy range regression. Education is endogenous in BMI regressions for women with and without covariates but this is not the case for male only regressions.

The finding that exogeneity of education is not rejected for some of the regressions may render some confidence in OLS estimation[27]. A priori for some health outcomes, the OLS method was inconsistent because of suspicion about endogeneity of education. However, ex-post coefficients estimated by the OLS method are not worse than IV estimated coefficients in some of the regressions.

4. The Analysis of the Effect of Education on Health Outcomes - Comments

Diagnostics checks demonstrate that it is not possible to overall reject exogeneity of education. It turns out that for some health outcomes (with respect to gender), coefficients estimated by OLS ex-post are not any worse than coefficients estimated by IV. Hence, it is not surprising that outcomes for women are statistically different from outcomes for men (Appendix, Tables A4 and A5). Exogeneity of education for index of bad health, to recall, is not rejected for women but rejected for men. Separate OLS regressions for men and women show that the effect of education on health is greater for females in index of bad health and BMI regressions. One additional year of education has stronger effect on female health status than on male health status. To instrument education with the reform status in female TSLS regressions is still possible even though the null of exogeneity is not rejected for females in index of bad health regressions. Education is significant at 10% level in health index regression for men but it is insignificant for women (Appendix, Table A5).

[27] Table A4 in the Appendix displays the OLS results for male and female regressions. The overview of the TSLS results by gender is also shown in the Appendix, Table A5.

The Effect of Education on Health Accounting for Income

Previous analysis did not include a person's income as a covariate in the OLS or IV regressions. It is known that a person's current income affects a person's current health status. A person's current income may also affect his or her intention to add to the current stock of knowledge by going to school longer. Therefore, it is necessary to include individual income as a covariate in the equation for health.

The next step is to include individual real income together with family background and other personal information in the original OLS and TSLS regressions (Tables 7, 8 and 9)[28]. Index of bad health, BMI and BMI in healthy range are health outcomes. The significant and positive effect of education on health is not altered if individual real income is added as another covariate. If the OLS regressions are considered first, the effect of education on health is significant at the 5% level in index of bad health regression, and at the 1% level in BMI and BMI in healthy range regressions (Table 7). An additional year of education decreases the number of conditions people may suffer from and increases the probability of having a BMI in healthy range. The income coefficient in the index of bad health regression is significant at the 1% level. Higher income improves health status as measured by number of conditions; however, income does not matter as much for a person to have a BMI in healthy range. It is interesting to mention that cohort effects are significant in the BMI regressions while

[28] Family background are parents schooling, lived with parents, number of siblings, if a family experienced any economic hardships while a person was growing up (a proxy for family economic resources or wealth), family health and rural, town or major town residence during childhood. Individual characteristics include income (a proxy for own wealth) and marital status.

regional effects are significant in the BMI in healthy range regressions. County residence during childhood is important and may indicate healthier adult BMI.

It is already mentioned that if education is endogenous, the OLS method produces inconsistent estimates. Hence, the health equation adding income in both stages of estimation is estimated by TSLS (Table 8). The effect of education on health is not changed when adding income. The educational coefficient from the second stage is negative and significant at the 10% level in the index of bad health regression. It is positive and significant at the 10% level in the BMI in healthy range regression.

Including persons' current income in the first stage may not be adequate since it may not reflect a contemporaneous effect. The change in education analyzed in this study is induced by the school reform, assuming that a person's current health is also a function of a person's current income. Therefore, income is excluded from the first stage and included in the second stage in the TSLS regressions with administrative reform status as an instrument for education (Table 9).

The TSLS regressions with income only in the second stage should be estimated manually because the set of exogenous variables included in the first stage differs from the set of exogenous variables included in the second stage[29]. Personal income and being married which are included in the second stage are not included in the first stage regression. The educational coefficient (-0.173) in the index of bad health regression is significant (10% level) and indicates better health with more education. Similarly, the educational coefficient (0.112) in the BMI in healthy range regression is also significant at 10% and indicates a higher probability of having BMI in healthy range with more

[29] Standard errors from the second stage are based on predicted schooling rather than on actual schooling.

education. Some family background characteristics matter for health (for instance, additional siblings). Higher income significantly decreases the index of bad health (5% level) indicating that current income matters for current health status. It is worth mentioning that there is still a significant and direct effect of education on adult health even after controlling for the contemporaneous effect of income on health.

An interesting question is about how much of the total effect of education on health is a direct effect of education on health and how much is an indirect effect of education on health that works through income in the regression that includes income. The difference in educational coefficients comparing the regression without income (the total effect) to the regression with income (direct effect) is -0.007 (the OLS result) and 0.004 (the IV result). This result indicates that about one-quarter of the total effect of education on health is the effect of education that works through income (indirect effect)[30].

The OLS results suggest that one year increase in schooling is equivalent to about a $1,000 increase in income in terms of health. The IV results suggest that one year increase in schooling is equivalent to about $10,000 increase in income in terms of health. The IV result suggests that lower additional schooling will achieve better health outcomes, since the IV result is up to ten times larger than the OLS result.

---

[30] **Educational coefficient: in the regression without income is (-0.028) and with income is (–0.021). Their difference is (–0.007), which is 0.25 of the total effect.**

The Effect of Education on Health with joint 1981 and 1991 data set

A bigger sample size is generated by pooling 1981 and 1991 observations (Appendix, Table A7). The advantage of pooling observations is that the number of total observations is being doubled. The main health outcome of interest, index of bad health, is also available in both years. For example, for men born between 1945 and 1955 there are 426 observations in 1991 and 900 in the pooled total set if only health index, education and reform status are considered. The pooled set contains observations from 1981 and 1991 that may come from the same individual. It is necessary to cluster around an individual in order to correct standard errors since there will be more than one observation per individual[31].

The first stage results are similar to the 1991 set's first stage results. In the regression with no covariates, a person gains 1.775 additional years of education if exposed to compulsory school reform compared to the reform status coefficient of 1.769 in the previous analysis (Table 10). The reform status t-statistic from the first stage is slightly higher (3.87) and is significant at 1% level. The 1991 dummy variable is included in the regressions to control for time trends[32].

---

[31] Some clusters will have two observations per individual (observations from both years) while other clusters will contain only one observation.

[32] The 1991 dummy variable is negative and insignificant. This result is due to misreported education. There are people who reported education completed in 1991 lower then in 1981. A couple of strategies are employed to deal with this problem (Appendix, Table A8). The change in the specification of education did not affect conclusions about the results. The 1991 dummy variable turns positive and significant in changed specifications of education denoting a positive trend in education over time. People get more education over time, for example, people who lived in 1970s and 1980s will get on average less education than people who lived in 1990s.

TSLS results that account for possible endogeneity of education are similar to results obtained with the 1991 set (Table 11). More education has positive and significant effect on adult health. The educational coefficient is -0.12, denoting a lower value for the index of bad health with additional education, and is significant (10% level) in the regression with cohort and regional effects and family background characteristics. The result from the same model with the 1991 set is -0.17, which is also significant at 10% level.

Generally, educational coefficients are lower for the total set for both the TSLS and the OLS regressions. IV coefficients are up to 10 times larger than the equivalent OLS coefficients obtained with the 1991 set. Further, the importance of adding income as a covariate is stressed elsewhere. If the OLS regressions with income are considered first, the educational coefficient is -0.01, significant at 10%, and half the size of the corresponding 1991 OLS coefficient (-0.02). Income coefficients of the same model are about the same size (-0.0016 for the total set and -0.0015 for the 1991 set) and both significant at 1% level. The negative signs on educational and income coefficients indicate that both higher education and higher income decrease the index of bad health, the result that implies better health status (Table 12).

It should be noted that other explanatory variables in the total set regression matter for health, which differs somewhat from the results obtained from the 1991 set. Mother's secondary education may improve person's adult health. If a person lived in a major town during childhood, this may have had the greatest positive effect on a person's adult health. It may make a difference for adult health if there is an ill person in the family, increasing the probability of having health problems later in life. Finally, marriage significantly improves adult health status, which is an established empirical finding (Table 13).

In all three TSLS models with and without income and with income only in the second stage there is a positive causal impact of education on health (Table 14). The educational coefficient is -0.110 (significant at 10%) in the regression with income only in the second stage. More years of schooling decrease the index of bad health implying better health status. The income coefficient is -0.0015 (significant at 1%) denoting better health with higher income. Finally, in the total set regression without income, exogeneity of education is rejected for men only, which is not the case in the regression with income only in the second stage (Tables 14 and 15).

Regarding the relationship between educational and income effects holding health constant, the OLS results suggest that 1.4 years increase in schooling is equivalent to about $1,000 increase in income in terms of health. The IV results suggest that 1.4 years increase in schooling is equivalent to about $10,000 increase in income in terms of health.

Chapter VI

Summary and Conclusion

Using the 1981 and 1991 Swedish Level of Living Surveys, there is empirical evidence that the effect of education on health is direct and causal. This result holds true for a number of health measures that are employed as health outcomes and across different samples and years.

The effect of education on health is first analyzed with the 1991 set. Index of bad health, BMI and BMI in healthy range are three health outcomes considered. The results are gender-sensitive; hence gender differences were addressed in this analysis. The educational coefficient in the index of bad health regression, controlling for individual income, is -0.173 and is significant at the 10% level for men born between 1945 and 1955. This is the cohort most affected by the school reform. Additional year in schooling enacted by the basic school reform translates in better health status given by the lower index of bad health.

Similar results are obtained when observations from the 1981 SLLS survey are added to observations from the 1991 SLLS survey. Joining observations from the two years nearly doubled the original sample. The educational coefficient in the index of bad health regression, controlling for individual income, is -0.110 and is significant at the 10% level for men born between 1945 and 1955. Similarity of results generated by these two sets renders confidence in the results obtained in this study.

A positive direct and causal effect of education on health remains even after controlling for income in the health structural equation. Additional schooling generated by the compulsory school reform in Sweden produces better adult health outcomes controlling for cohort and regional effects, family background and childhood characteristics, and individual income.

Table 1A.
Summary Statistics of Individual Variables

Variables	1991: Entire sample			1991: Born between 1945-1955		
	All	Females	Males	All	Females	Males
Female	0.49			0.47		
	(0.50)			(0.50)		
Age	44.74	45.43	44.06	41.33	41.36	41.30
	(16.40)	(16.59)	(16.19)	(3.10)	(3.03)	(3.16)
Educ self-reported	10.90	10.73	11.06	11.93	11.76	12.09
	(3.38)	(3.25)	(3.50)	(3.40)	(3.13)	(3.62)
Income, ann. (real, in 000)	62.38	49.96	74.54	75.49	59.41	89.90
	(40.79)	(29.85)	(46.07)	(46.05)	(31.14)	(52.07)
Zero Income Dummy	0.016	0.02	0.01	0.01	0.01	0.02
	(0.125)	(0.14)	(0.11)	(0.11)	(0.10)	(0.10)
Live w/ parents	0.84	0.84	0.85	0.87	0.88	0.86
	(0.36)	(0.37)	(0.36)	(0.34)	(0.32)	(0.35)
Siblings	0.90	0.90	0.89	0.88	0.89	0.86
	(0.30)	(0.29)	(0.31)	(0.33)	(0.31)	(0.34)
Fathelem	0.68	0.68	0.67	0.69	0.70	0.68
	(0.47)	(0.46)	(0.47)	(0.46)	(0.46)	(0.46)
Fathvoc	0.11	0.11	0.11	0.10	0.11	0.10
	(0.32)	(0.32)	(0.32)	(0.30)	(0.31)	(0.30)
Fathsec	0.14	0.13	0.14	0.16	0.14	0.18
	(0.34)	(0.34)	(0.35)	(0.37)	(0.35)	(0.38)
Fathuniv	0.07	0.07	0.07	0.04	0.04	0.04
	(0.26)	(0.26)	(0.25)	(0.19)	(0.19)	(0.20)
Rural	0.51	0.51	0.51	0.46	0.47	0.45
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Town	0.26	0.27	0.25	0.27	0.28	0.25
	(0.44)	(0.44)	(0.43)	(0.44)	(0.45)	(0.43)
Majtown	0.13	0.13	0.14	0.16	0.14	0.18
	(0.34)	(0.33)	(0.35)	(0.37)	(0.34)	(0.39)
Married	0.67	0.67	0.66	0.78	0.80	0.77
	(0.47)	(0.47)	(0.47)	(0.41)	(0.40)	(0.42)
Single	0.23	0.19	0.27	0.14	0.10	0.18
	(0.42)	(0.39)	(0.44)	(0.35)	(0.30)	(0.38)
N	6773	3350	3423	1522	716	806

Table 1B.
Summary Statistics of Different Measures of Health

Variables	1991: Entire sample			1991: Born between 1945-1955		
	All	Females	Males	All	Females	Males
Self-reported Health	0.77	0.76	0.79	0.85	0.85	0.85
	(0.42)	(0.43)	(0.41)	(0.35)	(0.36)	(0.35)
BMI	24.07	23.54	24.60	23.98	23.23	24.68
	(3.57)	(3.83)	(3.20)	(3.42)	(3.53)	(3.15)
Obesity	0.06	0.06	0.06	0.05	0.04	0.06
	(0.23)	(0.24)	(0.23)	(0.22)	(0.21)	(0.24)
Functional	0.19	0.23	0.15	0.09	0.12	0.07
Ability	(0.39)	(0.42)	(0.35)	(0.29)	(0.32)	(0.26)
Difficulty	0.06	0.07	0.05	0.03	0.03	0.03
walking	(0.23)	(0.25)	(0.22)	(0.17)	(0.17)	(0.17)
Difficulty	0.18	0.21	0.14	0.08	0.10	0.06
running	(0.38)	(0.41)	(0.35)	(0.27)	(0.30)	(0.25)
Smoking	0.30	0.29	0.30	0.35	0.33	0.37
	(0.46)	(0.45)	(0.46)	(0.48)	(0.47)	(0.48)
Drinking	0.87	0.84	0.90	0.91	0.90	0.92
	(0.34)	(0.36)	(0.30)	(0.28)	(0.30)	(0.27)
Heart	0.008	0.002	0.014	0.003	0	0.006
	(0.09)	(0.05)	(0.12)	(0.06)	(0)	(0.08)
High Blood	0.094	0.100	0.088	0.047	0.053	0.042
Pressure	(0.29)	(0.30)	(0.28)	(0.21)	(0.22)	(0.20)
Cancer	0.009	0.009	0.008	0.002	0.003	0.002
	(0.09)	(0.10)	(0.09)	(0.05)	(0.06)	(0.04)
Diabetes	0.024	0.022	0.025	0.011	0.007	0.016
	(0.15)	(0.15)	(0.16)	(0.11)	(0.08)	(0.12)
Overweight	0.129	0.156	0.103	0.118	0.125	0.112
	(0.33)	(0.36)	(0.30)	(0.32)	(0.33)	(0.31)
Kidney	0.020	0.017	0.022	0.024	0.017	0.030
problem	(0.14)	(0.13)	(0.15)	(0.15)	(0.13)	(0.17)
Conditions	0.222	0.245	0.200	0.170	0.168	0.173
Variable	(0.42)	(0.43)	(0.40)	(0.38)	(0.37)	(0.38)

Note: Health index of bad health, another measure of health included in this study, is standardized with mean 0 and standard deviation of 1.

Table 1C.
Summary Statistics of the Reform Status

Variables	1991: Entire sample			1991: Born between 1945-1955		
	All	Females	Males	All	Females	Males
Predicted Reform	0.64 (0.48)	0.63 (0.48)	0.65 (0.48)	0.87 (0.34)	0.88 (0.33)	0 86 (0.34)
Self-assessed Reform	0.47 (0.50)	0.46 (0.50)	0.47 (0.50)	0.48 (0.50)	0.47 (0.50)	0.49 (0.50)
N Predicted	3882	1936	1946	853	415	438
N Self-assessed	4821	2385	2436	1067	517	550
Constrained to equal N						
Predicted Reform				0.87 (0.34)	0.88 (0.33)	0.86 (0.34)
Self-assessed Reform				0.54 (0.50)	0.55 (0.50)	0.54 (0.50)
N				853	414	438

Table 1D.
Simple Correlations between Different Measures of Health and Education

Variables	1991: Entire sample			1991: Born between 1945-1955		
	All	Females	Males	All	Females	Males
(I) Index of Bad Health with Education	-0.32	-0.36	-0.29	-0.17	-0.22	-0.13
(II) BMI with Education	-0.24	-0.31	-0.18	-0.17	-0.23	-0.13
BMI with Index of Bad Health	0.24	0.29	0.22	0.18	0.25	0.14
(III) Smoking with Education	-0.03	-0.04	-0.03	-0.12	-0.15	-0.10
Smoking with Index of Bad Health	0.04	0.04	0.04	0.11	0.15	0.07
(IV) Functional Ability with Education	-0.31	0.34	-0.26	-0.13	-0.15	-0.11
Functional Ability with Index of Bad Health	0.72	0.75	0.73	0.66	0.67	0.65
(V) Self-rep. health with Education	0.28	0.30	0.24	0.13	0.19	0.08
Self-rep. health with Index of Bad Health	-0.65	-0.66	-0.65	-0.57	-0.59	-0.56

Table 1E.
Health Outcomes by Educational Categories

Educational Categories	Health Outcomes	
	0	1
Smoking		
7 and less of basic school	74	26
8 and 9 years (new school)	58	42
12 years and less (high school)	70	30
16 years and less (college)	73	27
more than 16 years	77	23
Heavy Drinking		
7 and less of basic school	74	26
8 and 9 years (new school)	57	43
12 years and less (high school)	48	52
16 years and less (college)	53	47
more than 16 years	58	42
Self-reported health		
7 and less of basic school	48	52
8 and 9 years (new school)	28	72
12 years and less (high school)	16	84
16 years and less (college)	13	87
more than 16 years	11	89
Obesity		
7 and less of basic school	88	12
8 and 9 years (new school)	92	8
12 years and less (high school)	96	4
16 years and less (college)	98	2
more than 16 years	96	4
Walking with no problems		
7 and less of basic school	84	16
8 and 9 years (new school)	94	6
12 years and less (high school)	97	3
16 years and less (college)	98	2
more than 16 years	99	1
Running with no problems		
7 and less of basic school	55	45
8 and 9 years (new school)	79	21
12 years and less (high school)	90	10
16 years and less (college)	92	8
more than 16 years	94	6

First Stage Results - Instrument's Analysis[1]

Dependent Variable – Self-reported Years of Formal Education
SLLS 1991: Sample of Men Born between 1945 and 1955

Table 2A.
The Analysis of the Two Instruments:
Predicted (Administrative) Reform Status and Self-assessed Reform Status – regressed one at the time

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Regional Effects	Cohort, Reg. Eff. & Father's Educ. Dum.
(I)				
Self-assessed	0.645	0.907	0.479	0.269
Reform Status	(1.91)	(2.18)	(1.09)	(0.65)
P>|t|	0.057	0.030	0.277	0.516
(II)				
Predicted	1.769	2.13	1.66	1.25
Reform Status	(3.66)	(3.50)	(2.44)	(1.95)
P>|t|	0.000	0.001	0.015	0.052

Table 2B.
F-statistic and R^2 from the First Stage Estimation

		No Covariates	Cohort Effects	Cohort and Regional Effects	Cohort, Reg. Eff. & Father's Educ. Dum.
Self-assessed	F-stat	16.58***	18.30***	7.27***	4.14**
Reform Status	R^2	0.029	0.054	0.115	0.218
N=548					
Self-assessed	F-stat	3.65*	4.76**	1.18	0.42
Reform Status	R^2	0.008	0.048	0.109	0.225
N=437					
Predicted	F-stat	13.41***	12.23***	5.97***	3.79**
Reform Status	R^2	0.029	0.065	0.119	0.232
N=437					

[1] Stars next to numbers denote level of significance. Respectively * star denotes 10% level, ** stars denote 5% level and *** stars denote 1% level of significance.

Table 2C.
The Analysis of the Two Instruments, Revisited:
Predicted (Administrative) Reform Status and Self-assessed Reform Status – both instruments regressed at the same time

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Regional Effects	Cohort, Reg. Eff. & Father's Educ. Dummies
Self-assessed Reform Status	0.195	0.689	0.425	0.231
	(0.54)	(1.65)	(0.97)	(0.56)
P>|t|	0.590	0.099	0.334	0.577
Predicted Reform Status	1.658	1.967	1.630	1.234
	(3.15)	(3.19)	(2.39)	(1.92)
P>|t|	0.002	0.002	0.017	0.056
F-statistic (joint)	6.84	7.51	3.45	2.05
Prob>F	0.0012	0.0006	0.0325	0.1302
R^2	0.030	0.071	0.122	0.232

Note that F-statistic is a joint statistic on both instruments.

Table 3.
Health Reduced Form Regressions[2]
OLS Regressions: Different Health Outcomes on Instruments - one at the time
SLLS 1991: Sample of Men Born between 1945 and 1955

Index of Bad Health as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Self-assessed	-0.093	-0.105	-0.165	-0.152	-0.161	-0.154
Reform status	(-1.48)*	(-1.33)*	(-1.93)**	(-1.78)**	(-1.88)**	(-1.79)**
Predicted	-0.071	-0.116	-0.250	-0.246	-0.247	-0.252
Reform status	(-0.78)	(-0.98)	(-1.88)**	(-1.84)**	(-1.84)**	(-1.88)**

Body Mass Index (BMI) as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Self-assessed	-0.765	-0.536	-0.608	-0.537	-0.496	-0.526
Reform status	(-2.52)***	(-1.41)	(-1.51)	(-1.34)	(-1.23)	(-1.30)
Predicted	-0.666	-0.483	-0.926	-0.739	-0.895	-0.879
Reform status	(-1.51)	(-0.86)	(-1.49)	(-1.19)	(-1.44)	(-1.41)

Body Mass Index in Healthy Range as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Self-assessed	0.122	0.084	0.092	0.088	0.079	0.091
Reform status	(2.61)***	(1.43)*	(1.49)*	(1.42)*	(1.28)*	(1.46)*
Predicted	0.170	0.108	0.150	0.134	0.163	0.161
Reform status	(2.50)***	(1.25)*	(1.58)*	(1.40)*	(1.69)**	(1.67)**

[2] Relevant test is a one-tail test for the index of bad health and the BMI in healthy range regressions. Negative reform status coefficient in the index of bad health regression represents better health.

Table 4.

OLS Regressions: Different Health Outcomes on Education
SLLS 1991: Sample of Men Born between 1945 and 1955

Index of Bad Health as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Education*	-0.026 (-3.35)	-0.026 (-3.30)	-0.032 (-3.96)	-0.032 (-3.66)	-0.032 (-3.68)	-0.029 (-3.29)

* all coefficients are significant at 1%

Body Mass Index (BMI) as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Education*	-0.117 (-3.38)	-0.109 (-3.14)	-0.115 (-3.23)	-0.100 (-2.60)	-0.096 (-2.49)	-0.113 (-2.87)

* all coefficients are significant at 1%

Body Mass Index in Healthy Range as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Education*	0.020 (3.79)	0.019 (3.52)	0.020 (3.66)	0.021 (3.54)	0.021 (3.44)	0.021 (3.37)

* all coefficients are significant at 1%

Table 5.
Two Stage Least Squares Regressions[3,4]:
Different Health Outcomes on Education
Instrumented with the Reform Status
SLLS 1991: Sample of Men Born between 1945 and 1955

Index of Bad Health as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
(I) if Predicted reform status	-0.042 (-0.83)	-0.050 (-0.98)	-0.141 (-1.65)**	-0.183 (-1.49)*	-0.165 (-1.58)*
(II) if Self-assessed reform	-0.153 (-1.29)*	-0.107 (-1.23)	-0.321 (-1.07)	-0.520 (-0.69)	-1.015 (-0.36)

Body Mass Index (BMI) as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
(I) if Predicted reform status	-0.369 (-1.44)	-0.225 (-0.87)	-0.542 (-1.36)	-0.573 (-1.07)	-0.574 (-1.27)
(II) if Self-assessed reform	-1.101 (-1.66)*	-0.557 (-1.29)	-1.128 (-1.02)	-1.569 (-0.73)	-2.620 (-0.47)

[3] The educational coefficient from the second stage is shown.

[4] Relevant test is a one-tail test for the index of bad health and the BMI in healthy range regressions. Negative reform status coefficient in the index of bad health regression represents better health.

Table 5, continued

Body Mass Index in Healthy Range as Health Outcome

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
(I) if Predicted reform status	0.095 (2.17)***	0.050 (1.22)	0.088 (1.41)*	0.105 (1.21)	0.105 (1.43)*
(II) if Self-assessed reform	0.179 (1.68)**	0.087 (1.28)*	0.171 (1.00)	.258 (0.74)	0.455 (0.47)

Table 6.
Two Stage Least Squares Regressions, Revisited
SLLS 1991: Sample of Men Born between 1945 and 1955

t-statistic in brackets	Index of Bad Health	Body Mass Index (BMI)	BMI in healthy range
Education	-0.169	-0.623	0.111
	(-1.55)*	(-1.32)	(1.44)*
Father's Educ. Vocat.	0.237	-0.112	-0.059
	(1.16)	(-0.12)	(-0.39)
Father's Educ. Secon.	0.294	-0.302	-0.135
	(1.32)*	(-0.31)	(-0.84)
Father's Educ. Univ.	0.654	1.630	-0.465
	(1.45)*	(0.84)	(-1.48)*
Mother's Educ. Vocat.	0.062	0.821	-0.200
	(0.36)	(1.06)	(-1.57)*
Mother's Educ. Secon.	0.023	1.647	-0.145
	(0.12)	(1.81)*	(-0.98)
Mother's Educ. Univ.	0.075	-0.181	0.215
	(0.20)	(-0.11)	(0.80)
Live with parents	0.131	-0.219	0.050
	(0.89)	(-0.33)	(0.46)
Siblings	-0.228	-0.336	0.015
	(-2.02)**	(-0.70)	(0.19)
Econom. Resources	-0.033	-0.521	0.035
	(-0.17)	(-0.62)	(0.26)
Family Health	0.005	0.299	-0.041
	(0.04)	(0.57)	(-0.47)
Rural area	-0.130	-	-
	(-0.78)	-	-
Town area	-	0.755	-0.082
	-	(1.03)	(-0.68)
Major town area	0.002	0.189	0.005
	(0.01)	(0.22)	(0.04)
Intercept	1.810	33.359	-0.947
	(1.34)*	(6.75)***	(-1.17)
Hausman test	-1.70*	-1.19	1.50
--- p-value	(0.090)	(0.237)	(0.134)
Cohort Effects	Yes	Yes	Yes
Region Effects	Yes	Yes	Yes
N	425	434	434

Table 7.
The Effect of Education and Income on Health: Adding Income
OLS Regressions: Different Health Outcomes with Different Covariates
SLLS 1991: Sample of Men Born between 1945 and 1955

Index of Bad Health as Health Outcome

t-statistic in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Fam. Backgr.	Coh., Reg. Eff., Fam. Backgr. & Ind. Character.
Education	-0.026	-0.032	-0.028	-0.021
	(-3.35)***	(-3.96)***	(-3.08)***	(-2.23)**
Family Background[5]				
Mother's Educ. Secon.			-0.189	-0.184
			(-1.85)**	(-1.82)**
Siblings			-0.153	-0.137
			(-1.77)**	(-1.58)*
Rural area			-0.250	-
			(-1.46)*	-
Town area			-0.290	-0.250
			(-1.66)**	(-1.42)*
Major town area			-0.303	-0.266
			(-2.41)***	(-2.06)**
Individual Character.[6]				
Married				-0.135
				(-1.87)**
Income				-0.0015
				(-2.40)***
Zero Income Dummy				-0.111
				(-0.40)
Cohort Effects		Fc=0.61	Fc=0.55	Fc=0.71
Region Effects[7]		Fr=1.09	Fr=0.69	Fr=0.71

[5] Family background variables include father's and mother's schooling, if lived with parents during childhood, number of siblings, economic resources, family health and rural, town or major town residence are during childhood.

[6] Individual characteristics are the 1990 individual real income and being married.

[7] This is a joint F-test that is calculated separately on the set of cohort dummies (Fc) and on the set of regional dummies (Fr).

Table 7, continued

Body Mass Index as Health Outcome

t-statistic in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Fam. Backgr.	Coh., Reg. Eff., Fam. Backgr. & Ind. Character.
Education	-0.117	-0.116	-0.116	-0.123
	(-3.38)***	(-3.23)***	(-2.90)***	(-2.93)***
Family Background				
Father's Educ. Secon.			-0.852	-0.850
			(-2.14)**	(-2.12)**
Mother's Educ. Secon.			0.975	0.967
			(2.19)**	(2.17)**
Individual Character.				
Married				-0.157
				(-0.49)
Income				0.00087
				(0.32)
Zero Income Dummy				-0.917
				(-0.73)
Cohort Effects		Fc=1.60*	Fc=1.50*	Fc=1.49*
Region Effects		Fr=1.13	Fr=1.03	Fr=0.99

Body Mass Index in Healthy Range as Health Outcome

t-statistic in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Fam. Backgr.	Coh., Reg. Eff., Fam. Backgr. & Ind. Character.
Education	0.020	0.020	0.020	0.020
	(3.79)***	(3.66)***	(3.14)***	(3.12)***
Family Background				
Live with parents			0.079	
			(1.30)*	
Family Health				-0.069
				(-1.28)*
Individual Character.				
Married				0.043
				(0.87)
Income				-0.0002
				(-0.56)
Zero Income Dummy				-0.060
				(-0.31)
Cohort Effects		1.27	1.25	1.25
Region Effects		1.79**	1.52*	1.45*

Table 8.

Two Stage Least Squares Regressions with Income: Different Health Outcomes[8]
SLLS 1991: Sample of Men Born between 1945 and 1955

Index of Bad Health as Health Outcome

t-statistics in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Family Background	Cohort, Region Eff., Fam. Back. & Ind. Character.
Education	-0.042	-0.141	-0.169	-0.212
	(-0.83)	(-1.65)**	(-1.55)*	(-1.43)*
Income				0.0015
				(0.67)
Zero Income				-0.806
Dummy				(-1.38)*

Body Mass Index (BMI) as Health Outcome

t-statistics in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Family Background	Cohort, Region Eff., Fam. Back. & Ind. Character.
Education	-0.369	-0.542	-0.623	-0.690
	(-1.44)	(-1.36)	(-1.32)	(-1.15)
Income				0.007
				(0.87)
Zero Income				1.513
Dummy				(0.65)

Body Mass Index in Healthy Range as Health Outcome

t-statistics in brackets	No Covariates	Cohort and Region Effects	Cohort, Region Eff. & Family Background	Cohort, Region Eff., Fam. Back. & Ind. Character.
Education	0.095	0.088	0.111	0.137
	(2.17)**	(1.41)*	(1.44)*	(1.34)*
Income				-0.002
				(-1.30)*
Zero Income				-0.060
Dummy				(-0.15)

[8] Real income is included in the first stage and in the second stage regressions (last column). Family background variables and individual characteristics are the same as before. Administrative reform status is the instrument in the first stage regression.

Table 9.
Two Stage Least Squares Regressions: Income in the Second Stage Only[9]
SLLS 1991: Sample of Men Born between 1945 and 1955

corrected st. err. in [.] t-stats in (.)	Index of Bad Health	Body Mass Index (BMI)	BMI in healthy range
Education	-0.173	-0.564	0.112
	[0.11201]*	[0.47016]	[0.07851]*
Father's Educ. Vocat.	0.234	-0.198	-0.058
	[0.20938]	[0.90160]	[0.15056]
Father's Educ. Second.	0.313	-0.379	-0.136
	[0.22740]*	[0.97391]	[0.16264]
Father's Educ. Univ.	0.754	1.509	-0.465
	[0.46271]*	[1.91687]	[0.32011]*
Mother's Educ. Vocat.	0.074	0.734	-0.199
	[0.17840]	[0.77101]	[0.12875]*
Mother's Educ. Second.	0.034	1.565	-0.144
	[0.20595]	[0.89565]*	[0.14957]
Mother's Educ. Univ.	0.0003	-0.230	0.200
	[0.38943]	[1.63322]	[0.27274]
Live with parents	0.150	-0.266	0.049
	[0.15077]	[0.66327]	[0.11076]
Siblings	-0.214	-0.332	0.019
	[0.11541]**	[0.47583]	[0.07946]
Econom. Resources	-0.071	-0.488	0.036
	[0.20255]	[0.82817]	[0.13830]
Family Health	-0.0016	0.344	-0.039
	[0.11842]	[0.52064]	[0.08694]
Rural area	-0.138	-	-
	[0.19384]	-	-
Town area	0.010	0.686	-0.082
	[0.18303]	[0.72938]	[0.12180]
Major town area	-	0.098	0.007
	-	[0.83558]	[0.13954]
Married	-0.113	0.006	0.018
	[0.10287]	[0.41988]	[0.07012]
Income	-0.0013	-0.0012	-0.00019
	[0.000815]**	[0.00329]	[0.00055]
Zero Income Dummy	-0.641	2.056	-0.169
	[0.52127]	[2.15745]	[0.36028]

[9] Real income is included in the second stage regression only. Being married is also included in the second stage regression only. Administrative reform status is the instrument for education in the first stage regression.

Table 9, continued

Two Stage Least Squares Regressions: Income in the Second Stage Only
SLLS 1991: Sample of Men Born between 1945 and 1955

corrected st. err. in [.] t-stats in (.)	Index of Bad Health	Body Mass Index (BMI)	BMI in healthy range
Intercept	2.079	32.939	-0.951
	[1.32536]*	[4.92865]***	[0.82306]
Hausman test	-1.81*	-1.09	1.52
--- p-value	0.071	0.278	0.130
Administrative reform	1.51	1.48	1.48
(1st stage)	(2.36)**	(2.35)**	(2.35)**
Cohort Effects	Yes	Yes	Yes
Regional Effects	Yes	Yes	Yes
N	425	434	434

The Set of Tables Showing the Analysis of the Total 1981 and 1991 Data Set
All Regressions Refer to Men Only Born between 1945 and 1955

Table 10.
First Stage Results Comparison: 1991 set vs. 1981 and 1991 Joint Set
Coefficient on Administrative Reform Status Instrument is shown in Reduced Form Equation for Education

	No Covariates	Cohort Effects	Cohort & Region Effects	Cohort, Region Effects & Father's Education	Cohort, Region Effects & Family Background	Cohort* Region Interactions
1991						
Coefficient	1.769	2.13	1.66	1.25	1.42	1.856
t-statistic	(3.66)	(3.50)	(2.44)	(1.95)	(2.25)	(1.71)
P>|t|	0.000	0.001	0.015	0.052	0.025	0.089
R^2	0.030	0.065	0.120	0.232	0.294	0.488
$F_{\text{on reform}}$	13.41	12.23	5.97	3.79	5.07	2.92
N=437						
1981&1991 with 1991 dummy						
Coefficient	1.775	2.13	1.773	1.35	1.44	1.626
t-statistic	(5.10)	(4.90)	(3.71)	(3.00)	(3.25)	(2.35)
P>|t|	0.000	0.000	0.000	0.003	0.001	0.019
R^2	0.028	0.053	0.105	0.217	0.259	0.408
$F_{\text{on reform}}$	26.06	24.01	13.74	9.01	10.56	5.50
N=910						
-if clustered						
Coefficient	1.775	2.13	1.773	1.35	1.44	1.626
t-statistic	(3.87)	(3.34)	(2.54)	(1.99)	(2.22)	(1.62)
P>|t|	0.000	0.001	0.011	0.047	0.027	0.105
R^2	0.028	0.053	0.105	0.217	0.259	0.408
$F_{\text{on reform}}$	14.97	11.16	6.44	3.97	4.94	2.63

Note: Adding the 1991 dummy variable does not change the reform coefficient in the first stage compared to the regressions that do not include the 1991 dummy. The 1991 dummy is not significant in the first stage (though is negative) when included in the regressions.

Table 11.
Overview of the Second Stage Results: Index of Bad Health as Health Outcome
Cohort and Region Effects are included in Both Stages of Estimation

t-statistics in brackets	Cohort Effects and Region Effects	Cohort Effects, Region Effects and Other covariates
(I) 1991	**-0.141**	**-0.169**
	(-1.65)**	**(-1.55)***
(II) 1981 and 1991	**-0.091**	**-0.121**
	(-1.76)**	**(-1.77)****
(III) 1981 and 1991 clustered on id	**-0.091**	**-0.121**
	(-1.50)*	**(-1.47)***
(IV) 1981 and 1991 with 1991 dummy	**-0.090**	**-0.119**
	(-1.73)**	**(-1.75)****
---1991 dummy	0.078	0.080
	(1.71)**	(1.66)**
(V) 1981 and 1991 clustered with 1991 dummy	**-0.090**	**-0.119**
	(-1.49)*	**(-1.46)***
---1991 dummy	0.078	0.080
	(2.07)**	(2.04)**

Note: Other covariates included are father's and mother's education, lived with parents, number of siblings, economic resources (a wealth proxy) and family health during childhood and rural, town or major town residence during childhood.

Table 12.
OLS Regressions: Sensitivity Analysis with Different Covariates
Index of Bad Health as Health Outcome

t-statistics in brackets (s.e. clustered)	(I) No Covariates	(II) Cohort Effects	(III) Cohort and Region Effects
Education	-0.021	-0.020	-0.023
	(-2.93)***	(-2.89)***	(-3.16)***
1991 dummy	0.065	0.064	0.070
	(1.99)**	(1.96)**	(2.15)**
	(IV) Cohort, Region Effects and Family Background Vars.	(V) Cohort, Region Effects, Family Backgr. & Married	(VI) Cohort, Region Eff., Fam. Backgr., Marr. & Income and Zero Inc. dummy
Education	-0.016	-0.015	-0.011
	(-2.13)**	(-2.05)**	(-1.39)*
1991 dummy	0.074	0.078	0.114
	(2.26)**	(2.37)***	(3.16)***
Married		-0.102	
		(-1.88)**	
Income			-0.0016
			(-2.88)***
Zero Income dummy			-0.119
			(-0.54)

Table 13.
OLS Regressions with and without Real Income and Health Index as Health Outcome

t-statistics in brackets	(I) Without Income	(II) With Income
Education	**-0.015**	**-0.011**
	(-2.05)**	**(-1.39)***
Father's Educ. Vocat.	0.082	0.082
	(0.92)	(0.93)
Father's Educ. Second.	-0.024	-0.029
	(-0.36)	(-0.42)
Father's Educ. Univers.	0.044	0.058
	(0.34)	(0.45)
Mother's Educ. Vocat.	-0.070	-0.073
	(-0.94)	(-1.00)
Mother's Educ. Second.	-0.104	-0.101
	(-1.55)*	(-1.50)*
Mother's Educ. Univers.	-0.062	-0.074
	(-0.46)	(-0.52)
Live with parents	-0.114	-0.114
	(-1.20)	(-1.20)
Siblings	-0.006	0.005
	(-0.11)	(0.08)
Econ. Resources	0.104	0.094
	(1.08)	(0.97)
Family Health	0.107	0.106
	(1.55)*	(1.54)*
Rural area	-0.237	-0.214
	(-1.67)**	(-1.47)*
Town area	-0.247	-0.218
	(-1.67)**	(-1.44)*
Major town area	-0.273	-0.240
	(-2.19)***	(-1.86)**
Married	-0.102	-0.078
	(-1.88)**	(-1.48)*
Income	-	-0.0016
		(-2.88)***
Zero Income Dummy	-	-0.119
		(-0.54)
Dummy 1991	0.078	0.114
	(2.37)***	(3.16)***
Intercept	0.456	0.479
	(2.26)***	(2.38)***
Cohort Effects	Yes	Yes
Region Effects	Yes	Yes
N	1294	1293

Table 14.
TSLS Regressions with and without Real Income (N=898)[10]

t-statistics in (.) brackets corrected s.e. in [.] brackets	(I) w/o Income	(II) w/Income in both stages	(III) w/Inc. in 2nd stage only
Education	-0.124	-0.127	-0.110
	(-1.50)*	(-1.32)*	[0.08086]
Father's Educ. Vocat.	0.241	0.246	(-1.36)*
	(1.49)*	(1.39)*	
Father's Educ. Second.	0.162	0.167	
	(0.98)	(0.92)	
Father's Educ. Univers.	0.547	0.545	
	(1.57)*	(1.60)*	
Mother's Educ. Vocat.	0.056	0.059	
	(0.46)	(0.46)	
Mother's Educ. Second.	0.051	0.056	
	(0.32)	(0.32)	
Mother's Educ. Univers.	0.015	0.020	
	(0.08)	(0.10)	
Live with parents	0.036	0.037	
	(0.30)	(0.30)	
Siblings	-0.059	-0.062	
	(-0.66)	(-0.66)	
Econ. Resources	-0.002	-0.002	
	(-0.01)	(-0.01)	
Family Health	0.054	0.053	
	(0.56)	(0.54)	
Rural area	-0.095	-0.096	
	(-1.07)	(-1.05)	
Town area	-	-	
	-	-	
Major town area	-0.007	-0.006	
	(-0.07)	(-0.06)	
Married	-0.059	-0.065	
	(-0.87)	(-0.95)	
Income	-	0.00034	-0.0015
		(0.22)	[0.00065]
Zero Income Dummy	-	-0.002	(-2.33)***
		(-0.01)	
Dummy 1991	0.082	0.073	
	(2.06)**	(1.40)*	
Intercept	1.327	1.350	
	(1.33)*	(1.26)	
Cohort Effects	Yes	Yes	Yes
Region Effects	Yes	Yes	Yes
Hausman test	-1.70 (p=0.090)*		-1.55 (p=0.121)

[10] Hausman test is a t-statistic on predicted education in the second stage. The same t-statistic with the opposite sign is obtained on predicted residuals in the second stage.

Table 15.
Hausman Test: Index of Bad Health as Health Outcome and Different Covariates
All Regressions include 1991 Dummy and Clustered Standard Errors

Males				
t-statistic in brackets	No covariates	Cohort Effects	Cohort Effects, Regional Effects	Cohort Effects, Regional Effects, Fam. Backgr.
Predicted Education	-0.005 (-0.14)	-0.015 (-0.39)	-0.075 (-1.40)	-0.110 (-1.63)*
P>|t|			0.161	0.104
N	900	900	900	898
Females				
t-statistic in brackets	No covariates	Cohort Effects	Cohort Effects, Regional Effects	Cohort Effects, Regional Effects, Fam. Backgr.
Predicted Education	0.008 (0.04)	-0.019 (-0.09)	0.097 (0.47)	0.216 (0.31)
N	845	845	845	842
Males and Females				
t-statistic in brackets	No covariates	Cohort Effects	Cohort Effects, Regional Effects	Cohort Effects, Regional Effects, Fam. Backgr.
Predicted Education	0.004 (0.09)	0.011 (0.16)	-0.107 (-0.92)	-0.113 (-1.07)
N	1745	1745	1745	1740

Appendix - Table A1.

Estimated Weights---Comparing Different Estimation Methods---SLLS 1991

HEALTH SYMPTOMS	A	B	C	D	E
Ability Indicators					
Walk 100m	0.47	0.44	0.92	-	-
Run 100m	0.69	0.76	0.78	-	-
Stairs walk up and down	0.39	0.39	0.43	-	-
Pains					
Pain in shoulders (light)	0.21	0.22	0.31	0.23	0.2
(severe)	0.58	0.59	1.02	0.57	0.87
Pain in back (light)	0.06	0.1	-0.03	0.13	0.05
(some)	0.45	0.45	0.44	0.56	0.65
Pain in joints (light)	0.15	0.16	0.18	0.21	0.23
(severe)	0.27	0.28	0.52	0.52	0.74
Cardiovascular					
Pain in chest (light)	0.08	0.1	-0.11	0.18	0.1
(severe)	0.26	0.31	0.78	0.25	0.71
Weak heart	0.25	0.3	-0.4	0.51	-0.18
High blood pressure (light)	0.24	0.33	0.04	0.41	0.12
(severe)	0.26	0.34	-0.4	0.47	-0.14
Varicose veins/ulcer (light)	0.25	0.25	0.39	0.28	0.3
(severe)	0.37	0.34	-0.66	0.36	-0.73
Swollen legs (light)	0.01	-0.01	-0.3	0.13	-0.23
(severe)	-0.35	-0.32	1.07	-0.001	1.28
Breathing difficulty (light)	0.15	0.11	0.25	0.32	0.41
(severe)	-0.17	-0.29	-0.001	0.17	0.7
Dizziness (light)	0.07	0.14	0.7	0.19	0.79
(severe)	0.1	0.07	0.38	0.15	0.4
Mental					
General tiredness (light)	0.35	0.31	0.28	0.31	0.28
(severe)	0.71	0.52	0.003	0.57	0.21
Insomnia (light)	0.19	0.17	0.11	0.22	0.15
(severe)	0.31	0.29	0.55	0.31	0.32
Nervousness/anxiety (light)	0.23	0.3	0.45	0.3	0.23
(severe)	0.46	0.38	0.87	0.43	1.01
Depression (light)	0.25	0.2	0.37	0.22	0.45
(severe)	0.24	0.27	0.38	0.2	0.27
Mental illness	0.6	0.69	0.13	0.79	0.16
Stomach/Intestinal					
Stomach pains (light)	-0.03	-0.04	0.13	-0.07	0.16
(severe)	0.27	0.23	0.29	0.18	0.26
Gall stones	0.07	0.07	0.2	0.11	0.39

Appendix – Table A1, continued

HEALTH SYMPTOMS	A	B	C	D	E
Unwell/feel sick (light)	-0.1	-0.14	-0.42	-0.16	-0.67
(severe)	-0.09	-0.04	-0.32	-0.06	-0.11
Vomiting (light)	-0.13	-0.04	0.49	-0.06	0.54
(severe)	0.28	0.18	0.46	0.13	0.41
Diarrhea (light)	-0.16	-0.12	-0.3	-0.11	-0.2
(severe)	-0.36	-0.39	-0.61	-0.39	-0.54
Gastric ulcer	0.33	0.33	0.23	0.25	0.3
Constipation (light)	0.01	0.02	-0.04	0.1	-0.05
(severe)	0.13	0.29	0.008	0.28	-0.01
Haemorrhoids (light)	0.08	0.12	0.34	0.06	0.28
(severe)	0.3	0.21	0.56	0.14	0.51
Other					
Genital discomfort (light)	-0.21	-0.29	-0.4	-0.29	-0.39
(severe)	0.09	0.16	-0.78	0.28	-0.3
Menstrual discomfort (light)	-0.03	-0.12	-0.19	-0.21	-0.17
(severe)	0.09	-0.05	0.31	-0.12	0.36
Headache (light)	0.02	-0.06	-0.17	-0.14	-0.13
(severe)	0.09	0.1	-0.13	-0.02	-0.15
Cold/Influenza (light)	-0.11	-0.16	-0.11	-0.21	-0.13
(severe)	-0.06	-0.11	-0.19	-0.21	-0.22
Poor vision (light)	0.33	0.21	0.74	0.24	0.77
(severe)	0.51	0.46	-0.66	0.6	-0.5
Impaired hearing (light)	0.19	0.27	0.24	0.32	0.18
(severe)	0.23	0.3	0.51	0.43	0.47
Chronic bronchitis/asthma (ligh	0.21	0.31	1.3	0.37	1.12
(severe)	0.7	0.67	2.36	0.82	2.57
Thyroid enlargement	-0.23	-0.25	-0.35	-0.18	0.07
Heart attack	0.24	0.18	-0.73	0.29	-0.37
Kidney stones (light)	0.11	0.01	0.56	-0.14	0.45
(severe)	-0.27	-0.49	0.003	-0.46	0.06
Urination problems (light)	-0.06	0.02	0.31	0.13	0.31
(severe)	0.25	0.27	0.27	0.3	0.26
Inguinal hernia	0.17	0.12	0.44	0.17	0.39

Appendix – Table A1, continued

HEALTH SYMPTOMS	A	B	C	D	E
Hot flushes (light)	-0.25	-0.14	-1.05	-0.21	-0.95
(severe)	-0.31	-0.17	0.72	-0.17	0.31
Cough (light)	0.02	-0.003	-0.1	0.001	-0.17
(severe)	-0.12	-0.02	-0.52	-0.04	-0.67
Weight loss	0.19	0.17	-1.08	0.19	-1.01
Overexertion (light)	-0.08	-0.14	0.1	-0.25	-0.02
(severe)	-0.13	-0.06	-0.41	-0.16	-0.42
Rashes/eczema (light)	0.06	0.03	0.24	0.01	0.22
(severe)	0.12	0.22	-0.23	0.19	-0.37
Cancer	0.86	0.78	2.24	0.78	1.91
Anemia	0.12	0.1	0.07	0.06	-0.24
Diabetes	0.48	0.45	1.32	0.57	1.21
Overweight (light)	-0.01	0.03	-0.06	0.11	0.002
(severe)	-0.12	0.01	0.37	0.2	0.61
Organic nervous disorder	0.32	0.34	0.05	0.62	0.48
Female		-0.16	-0.19	-0.06	-0.16
N	3551	5202	1179	5206	1180

Note:
Column A: Mikael Lindahl's weight estimates; People are between 34-76; Cubic in age and a gender dummy are included in regressions.
Columns B-E: Own weights for index of health with female dummy.
Column B: All individuals. Column C: People born between 1945-55.
Column D: Functional Ability variables not included; All individuals.
Column E: Functional ability not included; People born between 1945-55.

Appendix – Table A2.
First Stage Results – Instrument's Analysis
Dependent Variable – Self-reported Years of Formal Education
SLLS 1991: Sample of Men Born between 1945 and 1955

Table A2A.
Predicted Reform Status and Interaction Reform Status
(Predicted Reform Status and Self-assessed Reform Status Interacted) – regressed at the same time

t-statistic in brackets	No Covariates	Cohort Effect	Cohort and Regional Effects	Cohort, Reg., Eff. & Father's Educ. Dum.
Predicted	1.607	1.899	1.447	1.202
Reform Status	(3.01)	(3.06)	(2.47)	(1.86)
P>|t|	0.003	0.002	0.014	0.063
Interaction	0.263	0.789	0.542	0.347
Reform Status	(0.71)	(1.86)	(1.36)	(0.82)
P>|t|	0.476	0.064	0.175	0.411

Note: Sample size is N=437 (if father's education included sample size is N=436)

Table A2B.
Predicted Reform Status, Self-assessed Reform Status and Interaction Reform Status – regressed at the same time

t-statistic in brackets	No Covariates	Cohort Effect	Cohort and Regional Effects	Cohort, Reg., Eff. & Father's Educ. Dum.
Predicted	1.511	1.811	1.321	1.052
Reform Status	(2.78)	(2.87)	(2.22)	(1.60)
P>|t|	0.006	0.004	0.027	0.111
Self-assessed	-1.912	-1.626	-2.255	-2.295
Reform Status	(-0.93)	(-0.80)	(-1.18)	(-1.17)
Interaction	2.175	2.412	2.791	2.629
Reform Status	(1.04)	(1.16)	(1.43)	(1.32)

Note: Sample size is N=437 (if father's education included sample size is N=436)

Appendix - Table A3.
Hausman Test: Different Health Outcomes[11]
SLLS 1991: Population Born between 1945 and 1955

Index of Bad Health as Health Outcome

Men Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	-0.020 (-0.39)	-0.031 (-0.56)	-0.125 (-1.57)	-0.153 (-1.67)*

Women Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	0.106 (0.26)	-0.146 (-0.87)	-0.081 (-0.40)	-0.157 (-0.30)

Men and Women Together

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	0.00069 (0.01)	0.053 (0.44)	-0.095 (-0.53)	-0.086 (-0.58)

[11] t-statistic in brackets with stars denotes level of significance (* significant at 10%, ** significant at 5%, *** significant at 1%) as previously noted.

Appendix - Table A3, continued

Body Mass Index (BMI) as Health Outcome

Men Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	-0.247 (-0.99)	-0.094 (-0.36)	-0.405 (-1.09)	-0.467 (-1.09)

Women Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	-5.555 (-3.34)***	1.415 (2.01)**	1.494 (1.79)*	4.085 (1.86)*

Men and Women Together

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	-0.882 (2.92)***	-0.824 (-1.54)	-1.565 (1.98)**	-1.279 (-1.97)**

Appendix - Table A3, continued

BMI in Healthy Range as Health Outcome

Men Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	0.080 (2.06)**	0.034 (0.84)	0.072 (1.26)	0.093 (1.41)

Women Only

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	0.543 (2.48)**	-0.090 (-0.96)	-0.065 (-0.60)	-0.175 (-0.62)

Men and Women Together

t-statistic in brackets	No Covariates	Cohort Effects	Cohort and Region Effects	Cohort, Reg., Eff., Fath.& Moth. Educ. Live, Siblings
Predicted Education	0.137 (3.16)***	0.095 (1.24)	0.173 (1.54)	0.142 (1.53)

Appendix – Table A4.
OLS Regressions: Different Health Outcomes
SLLS 1991: Population born between 1945 and 1955

Index of Bad Health as Health Outcome

Men only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.026 (-3.35)	-0.026 (3.30)	-0.032 (-3.96)	-0.032 (-3.66)	-0.029 (-3.29)

Note: Sample size is N (607).

Women only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.056 (-5.28)	-0.056 (-5.20)	-0.055 (-5.02)	-0.049 (-4.14)	-0.049 (-3.98)

Note: Sample size is N (572).

Men and Women together

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.038 (-5.98)	-0.039 (-5.97)	-0.042 (-6.46)	-0.039 (-5.54)	-0.037 (-5.18)

Note: Sample size is N (1179).

Appendix – Table A4, continued

Body Mass Index (BMI) as Health Outcome

Men only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.117 (-3.38)	-0.109 (-3.14)	-0.115 (-3.23)	-0.100 (-2.60)	-0.113 (-2.87)

Note: Sample size is N (620).

Women only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.265 (-5.73)	-0.243 (-5.16)	-0.235 (-4.90)	-0.211 (-4.09)	-0.200 (-3.74)

Note: Sample size is N (579).

Men and Women together

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	-0.177 (-6.30)	-0.166 (-5.86)	-0.168 (-5.89)	-0.147 (-4.80)	-0.150 (-4.75)

Note: Sample size is N (1199).

Appendix – Table A4, continued

BMI in Healthy Range as Health Outcome

Men only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education*	0.020 (3.79)	0.019 (3.52)	0.020 (3.66)	0.021 (3.54)	0.021 (3.37)

Note: Sample size is N (620). Also, * on the educational coefficient denotes that all coefficients are significant at 1%.

Women only

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education	0.018 (3.04)***	0.015 (2.42)**	0.013 (2.24)**	0.017 (2.67)***	0.016 (2.45)**

Note: Sample size is N (579).

Men and Women together

t-statistic in brackets	No Covariates	Cohort effects	Cohort and Region Effects	Coh., Reg., Eff. & Father's Education	Coh., Reg., Eff., Fath.& Moth. Educ, Live Sibl.
Education*	0.019 (4.89)	0.018 (4.40)	0.018 (4.47)	0.020 (4.62)	0.019 (4.35)

Note: Sample size is N (1199). Also, * on the educational coefficient denotes that all coefficients are significant at 1%.

Appendix – Table A5.
TSLS Regressions: Different Health Outcomes[12]
SLLS 1991: Sample of Males, Females, and All Born between 1945 and 1955

t-statistic in brackets	Health Index	Health Index without functional ability	BMI	BMI in healthy range	Six major conditions
Males					
Education	-0.165	-0.154	-0.574	0.105	-0.096
	(-1.58)*	(-1.47)*	(-1.27)	(1.43)*	(-1.48)*
N	425	425	434	434	436
Females					
Education	-0.283	-0.141	6.326	-0.262	0.205
	(-0.26)	(-0.14)	(0.32)	(-0.28)	(0.53)
N	406	407	411	411	414
All					
Education	-0.106	-0.107	-1.334	0.147	-0.147
	(-0.77)	(-0.75)	(-1.52)	(1.32)*	(-1.23)
Female	0.116	0.073	-1.984	0.188	-0.061
	(1.74)*	(1.06)	(-5.10)***	(3.82)***	(-1.22)
N	831	832	845	845	850

[12] All regressions include cohort and regional effects, father's and mother's education, living with parents and number of siblings. Administrative reform status is the instrument for education in the first stage regression.

Appendix – Table A6.
Additional Health Outcome Considered: Six major conditions
SLLS 1991: Sample of Men Born between 1945 and 1955

Health Reduced Form Regressions

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Self-assessed Reform status	-0.031 (-0.83)	-0.027 (-0.59)	-0.035 (-0.70)	-0.035 (-0.70)	-0.025 (-0.51)	-0.018 (-0.37)
Predicted Reform status	-0.073 (-1.34)*	-0.049 (-0.71)	-0.114 (-1.48)*	-0.111 (-1.43)*	-0.139 (-1.79)**	-0.140 (-1.80)**

OLS Regressions: Six major conditions on Self-reported Schooling

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh., Reg., Fath.Ed., Live, Sibl	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
Education	-0.006 (-1.36)*	-0.005 (-1.09)	-0.006 (-1.39)*	-0.006 (-1.22)	-0.004 (-0.94)	-0.005 (-1.13)

TSLS Regressions: Six major conditions on Self-reported Schooling

t-statistic in brackets	No Covariates	Cohort Effects	Cohort, Region Effects	Coh. Reg., Father's Educat.	Coh.,Reg., Fath.Ed., Live, Sibl, Moth. Educ.
(I) if Predicted reform status	-0.041 (-1.29)*	-0.023 (-0.71)	-0.068 (-1.33)*	-0.088 (-1.21)	-0.096 (-1.48)*
(II) if Self-assessed reform	-0.047 (-0.77)	-0.030 (-0.58)	-0.073 (-0.62)	-0.130 (-0.50)	-0.153 (-0.24)

The Set of Tables in the Appendix Showing the Analysis of the Total 1981 and 1991 Data Set

All Regressions Refer to Men or Population Born between 1945 and 1955

Appendix – Table A7.
Sample Size Gain

N	Males		Females		Males and Females	
	1991	1981&1991	1991	1981&1991	1991	1981&1991
Health Index	609	1303	572	1224	1181	2527
Education	622	1446	586	1337	1208	2783
Income	792	1603	709	1431	1501	3034
Reform Status	438	913	415	856	853	1769

Appendix - Table A8.
First Stage Results Comparison, Revisited[13]:
Sensitivity of Results to Changes in Specification of Education

1981&1991 with 1991 dummy clustered	No Covariates	Cohort Effects	Cohort & Region Effects	Cohort, Region Effects & Father's Education	Cohort, Region Effects & Family Backgr. Vars	Cohort* Region Interactions
Original Education						
Coefficient	1.775	2.13	1.773	1.35	1.44	1.626
t-statistic	(3.87)	(3.34)	(2.54)	(1.99)	(2.22)	(1.62)
P>|t|	0.000	0.001	0.011	0.047	0.027	0.105
R^2	0.028	0.053	0.105	0.217	0.259	0.408
$F_{\text{on reform}}$	14.97	11.16	6.44	3.97	4.94	2.63
Educ1						
Coefficient	1.776	2.15	1.81	1.37	1.466	1.538
t-statistic	(3.83)	(3.23)	(2.48)	(1.94)	(2.16)	(1.52)
P>|t|	0.000	0.001	0.013	0.053	0.031	0.128
R^2	0.028	0.052	0.125	0.237	0.271	0.479
$F_{\text{on reform}}$	14.67	10.44	6.16	3.78	4.65	2.31
1991 dum.	0.42	0.38	0.42	0.44	0.47	0.42
t-statistic	(4.49)	(4.15)	(4.45)	(4.82)	(5.16)	(4.48)
Educ2						
Coefficient	1.377	1.77	1.60	1.37	1.53	2.18
t-statistic	(2.39)	(2.30)	(2.00)	(1.72)	(2.02)	(2.27)
P>|t|	0.018	0.022	0.046	0.087	0.044	0.024
R^2	0.026	0.059	0.120	0.238	0.293	0.562
$F_{\text{on reform}}$	5.69	5.31	4.00	2.95	4.09	5.15
1991 dum.	0.51	0.52	0.53	0.51	0.53	0.60
t-statistic	(6.11)	(6.19)	(6.42)	(6.07)	(6.33)	(7.00)

[13] About 15% of the sample reported lower education in 1991 than in 1981. Two auxiliary variables are created in order to test the sensitivity of results if the original education variable is altered to account for this misreporting. The first approach is to create a new variable (Educ1) with self-reported education in 1981 replacing self-reported education in 1991 for observations where the 1991 education is lower than the 1981 education. The second approach is to create another variable (Educ2) with missing replacing education in 1991 if education in 1991 is lower than education in 1981. The shortcoming of the second approach is that some of the observations will be lost because they are turned into missing. Also note that in these regressions the 1991 dummy is positive and significant at 1% level while in the regressions with the original education the 1991 dummy is negative and insignificant.

Appendix – Table A9.
Overview of the Second Stage Results:
Sensitivity of Results to Changes in Specification of Education
Index of Bad Health as Health Outcome
Cohort and Region Effects are included in Both Stages of Estimation
Total 1981 and 1991 Data Set with 1991 Dummy Clustered

t-statistics in brackets	Cohort Effects and Region Effects	Cohort Effects, Region Effects and Other covariates[14]
(I) Original Education	-0.090	-0.119
	(-1.49)*	(-1.46)*
---1991 dummy	0.078	0.080
	(2.07)**	(2.04)**
(II) Educ1	-0.087	-0.116
	(-1.47)*	(-1.43)*
---1991 dummy	0.120	0.135
	(2.52)***	(2.40)***
(III) Educ2	-0.050	-0.074
	(-0.71)	(-0.93)
---1991 dummy	0.125	0.137
	(2.04)**	(2.15)**

[14] Note, other covariates are father's and mother's education, lived with parents, number of siblings, economic resources (a wealth proxy), family health and rural, town or major town residence during childhood.

References

Acemoglu, Daron and Joshua Angrist, "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws", NBER Working Paper No. 7444, December 1999

Adams, Scott J., "Educational Attainment and Health: Evidence from a Sample of Older Adults", Education Economics, Vol. 10, No. 1, 2002

Angrist, Joshua D. and Alan. B. Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?", Quarterly Journal of Economics, 106 (4), November 1991, pp. 979-1014

Angrist, Joshua D., and Guido W. Imbens, "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity", Journal of the American Statistical Association, 90 (430), June 1995, pp. 431-442

Arai, Mahmood and Christian Kjellström, "Sweden", in "Education and Earnings in Europe: A Cross Country Analysis of the Returns to Education", edited by Colm Harmon, Ian Walker and Niels Westergaard-Nielsen, Edward Elgar, 2001, Chapter 14, pp. 265-284

Arendt, Jacob Nielsen, "Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification", Institute of Local Government Studies, July 2002

Arendt, Jacob Nielsen, "Education Effects on Health: A Panel Data Analysis Using School Reforms for Identification", Institute of Economics, University of Copenhagen, November 2001

Arendt, Jacob Nielsen, "Education Effects on Health: Causal or from Unobserved Components? A Panel Data Analysis with Endogenous Education," Institute of Economics, University of Copenhagen, September 2001

Arendt, Jacob N., "Education and Income Related Inequality in Self-Reported Health: Evidence from the CPS", AKF, Institute of Local Government Studies, 2001

Arendt, Jacob N., "Endogeneity and Heterogeneity in LDV Panel Data Models", Institute of Local Government Studies, November 2001

Arkes, Jeremy, "Does Schooling Improve Adult Health?", RAND, Paper presented at the AERA Conference, April 2001

Auster, Richard, Irving Leveson and Deborah Sarachek, "The Production of Health, An Exploratory Study", Journal of Human Resources, 4 (4), Fall 1969, pp. 411-436

Becker, Gary S. and Casey B. Mulligan, "The Endogenous Determination of Time Preference", Quarterly Journal of Economics, August 1997, pp. 729-758

Behrman, Jere R., "Family Background, Education Determination and Policy Implications: Some Selected Aspects from Various Countries", Conference on After the Bell: Education Solutions Outside the School, The Jerome Levy Economics Institute of Bard College Blithewood, Annandale-on-Hudson, New York, June 2001

Behrman, Jere R. and Barbara L. Wolfe, "Does More Schooling Make Women Better Nourished and Healthier? Adult Sibling Random and Fixed Effects Estimates for Nicaragua", Journal of Human Resources, Fall 1989, pp. 644-663

Berger, Marc C. and J. Paul Leigh, "Schooling, Self-Selection, and Health", Journal of Human Resources, 24 (3), Summer 1989, pp. 433-455

Blackburn, McKinley L. and David Neumark, "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look", Review of Economics and Statistics, 77 (2), May 1995, pp. 217-230

Bolin, Kristian, Bengt Liljas, and Björn Lindgren, "Investments in Education and Health as Simultaneous Decisions: An Extended Human Capital Model and Empirical Tests", Lund University, Department of Community Medicine and Department of Economics, Lund University Centre for Health Economics (LUCHE), 2002b

Bolin, Kristian, Lena Jacobson and Björn Lindgren, "The Demand for Health and Health Investments in Sweden 1980/81, 1988/89, and 1996/97", Lund University, Department of Community Medicine, Department of Economics and Lund University Centre for Health Economics (LUCHE), 2002a

Bound, John, David A. Jaeger, and Regina M. Baker, "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak", Journal of the American Statistical Association, 90 (430), June 1995, pp. 443-450

Card, David, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", Econometrica, 69 (5), September 2001, pp. 1127-1160

Card, David, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", NBER Working Paper No. 7769, June 2000

Card, David, "The Causal Effect of Education on Earnings", Handbook of Labor Economics, edited by Orley Ashenfelter and David Card, Elsevier Science, Volume 3, 1999, Chapter 30, pp. 1801-1863

Card, David, "Earnings, Schooling, and Ability Revisited", NBER Working Paper No. 4832, August 1994

Christenson, Bruce A., and Nan E. Johnson, "Educational Inequality in Adult Mortality: An Assessment with Death Certificate Data from Michigan", Demography, 32 (2), May 1995, pp. 215-229

Currie, Janet and Enrico Moretti, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data", NBER Working Paper No. 9360, December 2002

Deaton, Angus and Christina Paxson, "Mortality, Education, Income, and Inequality Among American Cohorts", NBER Working Paper No. 7140, May 1999

Elo, Irma T. and Samuel H. Preston, "Educational Differential in Mortality: United States, 1979-85", Social Science of Medicine, 42 (1), 1996, pp. 47-57

Farrell, Phillip and Victor R. Fuchs, "Schooling and Health, The Cigarette Connection", Journal of Health Economics, 1, 1982, pp. 217-230

Gerdtham, Ulf G. and Magnus Johannesson, "New estimates of the demand for health: results based on a categorical health measure and Swedish micro data", Social Science & Medicine, 49, 1999b, pp. 1325-1332

Gerdtham, Ulf G., M. Johannesson, L. Lundberg and D. Isacson, "The demand for health: results from new measures of health capital", European Journal of Political Economy, Vol. 15, 1999a, pp. 501-521

Gerdtham, Ulf G. and Gun Sundberg, "Measuring Income Related Health Inequalities in Sweden", Stockholm School of Economics, Scandinavian Working Papers in Economics No. 120, August 1996

Gilleskie, Donna B., and Amy L. Harrison, "The Effect of Endogenous Health Inputs on the Relationship Between Health and Education", Economics of Education Review, 17 (3), June 1998, pp. 279-297

Goldman, Dana and Darius Lakdawalla, "Understanding Health Disparities Across Education Groups", NBER Working Paper No. 8328, June 2001

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems", Econometrica, 45 (1), January 1977, pp. 1-22

Grossman, Michael, "Household Production and Health", NBER, paper presented at a conference in honor of Jacob Mincer's 80th Birthday at Columbia University on July 15th, 2002, June 2002

Grossman, Michael, "The Human Capital Model", Handbook of Health Economics, edited by Anthony J. Culyer and Joseph P. Newhouse, Amsterdam: North-Holland, Elsevier Science, Volume 1A, 2000, Chapter 7, pp. 348-408

Grossman, Michael, "The Human Capital Model of the Demand for Health", NBER Working Paper No. 7078, April 1999

Grossman, Michael, "The Correlation Between Health and Schooling", in "Household Production and Consumption", edited by N. E. Terleckyj, Studies in Income and Wealth, Volume 40, Conference on Research in Income and Wealth, Columbia University Press for the National Bureau of Economic Research, New York, 1975, pp. 147-211

Grossman, Michael, "On the Concept of Health Capital and the Demand for Health", Journal of Political Economy, 80 (2), March-April 1972a, pp. 223-255

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

Grossman, Michael and Robert, Kaestner, "Effects of Education on Health", in "The Social Benefits of Education", edited by Jere R. Behrman and Nevzer Stacey, Ann Arbor: University of Michigan Press, 1997, Chapter 4, pp. 69-123

Harmon, Colm, Ian Walker and Niels Westergaard-Nielsen, "Introduction" in "Education and Earnings in Europe: A Cross Country Analysis of the Returns to Education", edited by Colm Harmon, Ian Walker and Niels Westergaard-Nielsen, Edward Elgar, 2001, Chapter 1, pp. 1-37

Harmon, Colm and Ian Walker, "Estimates of the Economic Return to Schooling for the United Kingdom", American Economic Review, 85 (5), December 1995, pp. 1278-1286

Hartog, Joop and Hessel Oosterbeek, "Health, Wealth and Happiness: Why Pursue a Higher Education?", Economics of Education Review, 17 (3), June 1998, pp. 245-256

Ippolito, Richard A., "Education and Investment Behavior in the Family: A Study of Health Human Capital in Middle Age", George Mason University School of Law, April 2002

Kemna, Harrie J.M.I., "Working Conditions and the Relationship Between Schooling and Health" Journal of Health Economics, 6, 1987, pp. 189-210

Kenkel, Donald S., "Prevention", Handbook of Health Economics, edited by Anthony J. Culyer and Joseph P. Newhouse, Amsterdam: North-Holland, Elsevier Science, Volume 1B, 2000, pp. 1675-1720

Kenkel, Donald S., "Health Behavior, Health Knowledge, and Schooling", Journal of Political Economy, 99 (2), April 1991, pp. 287-305

Kling, Jeffrey R., "Interpreting Instrumental Variables Estimates of the Returns to Schooling", NBER Working Paper No. 7989, October 2000

Lang, Kevin and David Kropp, "Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws", Quarterly Journal of Economics, 101 (3), August 1986, pp. 609-624

Leigh, Paul J., "The Social Benefits of Education: A Review Article", Economics of Education Review, 17 (3), June 1998, pp. 363-368

Leigh, Paul J., "Schooling and Seat Belt Use", Southern Economic Journal, 57 (1), 1990, pp. 195-207

Leigh, Paul J., "Direct and Indirect Effects of Education on Health", Social Science of Medicine, 17 (4), 1983, pp. 227-234

Lindahl, Mikael, " Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as Exogenous Source of Variation in Income", University of Amsterdam and IZA, Bonn, Discussion Paper No. 442, February 2002

Lleras-Muney, Adriana, "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939", NBER Working Paper No. 8563. October 2001

Lleras-Muney, Adriana, "The Relationship Between Education and Adult Mortality in the U.S.", NBER Working Paper No. 8986, June 2002

Meghir, Costas and Marten Palme, "The Effect of a Social Experiment in Education", Stockholm School of Economics, Scandinavian Working Papers in Economics No. 451, May 2001

Meghir, Costas and Marten Palme, "Assessing the Effect of Schooling on Earnings Using a Social Experiment", Stockholm School of Economics, Scandinavian Working Papers in Economics No.313, March 1999

Munasinghe, Lalith and Nachum Sicherman, "Why Do Dancers Smoke? Time Preference, Occupational Choice, and Wage Growth", NBER Working Paper No. 7542, February 2000

Sander, William, "Schooling and Quitting Smoking", Review of Economics and Statistics, 77 (1), February 1995, pp. 191-199

Statistics Sweden, http://www.scb.se

Staiger, Douglas and James H. Stock, " Instrumental Variables Regression with Weak Instruments", Econometrica, Vol. 65, No. 3, May 1997, pp. 557-586

Ross, Catherine E., and Chia-ling Wu, "The Links Between Education and Health", American Sociological Review, 60 (5), October 1995, pp. 719-745

Wagstaff, Adam, "The Demand for Health, Some New Empirical Evidence", Journal of Health Economics, 5, 1986, pp. 195-233

Walque, de Damien, "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda", Ph.D. Dissertation, University of Chicago, November 2002

Willis, Robert J., and Sherwin Rosen, "Education and Self-Selection", Journal of Political Economy, 87 (5), Part 2, October 1979, pp. S7-S36

Wooldridge, Jeffrey M., "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model", Economic Letters, 56, 1997, pp. 129-133

Wooldridge, Jeffrey M., Econometric Analysis of Cross Section and Panel Data, Cambridge: The MIT Press, 2002