C.I.R.P.É.E

Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **13-25**

The Effect of Education on Health : Cross-Country Evidence

Raquel Fonseca
Yuhui Zheng

Septembre/September 2013

Fonseca : Université du Québec à Montréal, RAND Corporation and CIRPÉE
Corresponding author: UQAM ESG, Département des sciences économiques, 315 rue Ste-Catherine Est, Montréal, QC, Canada H2X 3X2; Phone +1 514 987-4114 ext. 5850; Fax +1 514 987-8494
fonseca.raquel@uqam.ca
Zheng : Leonard D. Schaeffer Center, USC
yuhuizhe@healthpolicy.usc.edu

We would like to thank Zhuo Chen, Bernard Fortin, Mohammad Hajizadeh and participants at the Canadian Economics Association Conference 2013; the SCSE Conference 2011; and the ASHE Conference, Cornell 2010. As well, comments from seminar participants at the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University were most useful. This research was supported by the National Institute on Aging, under the grants 5P01AG08291 and P01 AG022481. Yuhui Zheng would like to acknowledge postdoctoral fellowship funding support from the RAND Bing Center for Health Economics and the Harvard Center for Population and Development Studies. We thank David Boisclair for his excellent editorial assistance. This paper is a revised version of the RAND WP 864. Errors are our own.

Abstract:
This paper uses comparable micro-data from over 15 OECD countries to study the causal relationship between education and health outcomes. We combine three surveys (SHARE, HRS and ELSA) that include nationally representative samples of people aged 50 and over in these countries. We use variation in the timing of educational reforms across these countries as an instrument for the effect of education on health. Using instrumental variables Probit models (IV-Probit), we find causal evidence that more years of education lead to better health for a limited number of health markers. We find lower probabilities of reporting poor health, of having limitations in functional status (ADLs and iADLs) and of having been diagnosed with diabetes. These effects are larger than those from a Probit that does not control for the endogeneity of education. We cannot find evidence of a causal effect of education on other health conditions. Interestingly, the relationship between education and cancer is positive in both Probit and IV-Probit models, which we interpret as evidence that education fosters early detection.

Keywords: Education, health, causality, compulsory schooling laws

JEL Classification: I1, I14, I2

1. INTRODUCTION

There is abundant evidence that health and education are linked irrespective of how the former is measured (i.e. mortality, morbidity, or health-related behaviors).[3] These associations exist across time and countries, even though their magnitude might differ (Banks et al., 2006; Andreyeva et al., 2007; Mackenbach et al., 2008; Avendano et al., 2009; Michaud et al., 2011). These associations are important for policy if it aims to reduce disparities. However, reducing disparities requires that we untangle the causal structure of the relationship. Does raising education lead to better health, or does better health lead to better educational outcomes? Are there other – perhaps unobserved – characteristics that might determine both? Unfortunately, finding an association does not provide an answer to these key questions, which limits the ability to effectively reduce disparities.

There are plenty of reasons why education may improve health. Education could improve efficiency in health production (productive efficiency – see Grossman, 1972); reallocate inputs in health production (allocative efficiency – see Grossman, 2005); change time preference (Fuchs, 1982); modify behavior-related factors such as smoking, obesity or preventive care (Huisman et al., 2005; Mackenbach et al., 2008); and finally, education could generate more resources, e.g. through higher income or occupational status, or through better housing, food, quality of care, and living environment (Case and Deaton, 2005; Cutler et al., 2008).

Conversely, early health endowments could affect both educational outcomes and health in later life. Early-life health problems can lead to lowered school attendance and affect education during childhood, which can subsequently lead to poor health in later life (Grossman and Kaestner, 1997; Cutler and Lleras-Muney, 2010).

[3] For a review, see Grossman (2005), Cutler and Lleras-Muney (2006), and Cutler et al. (2008) among others.

Finally, common factors, like preferences, genetic elements or family background, could affect both education and health outcomes. Changing preferences, such as tastes, discount rates or risk aversion can affect both education decisions and health behaviors (Fuchs, 1982; Becker and Mulligan, 1997; Cutler and Lleras-Muney, 2010). Genetic factors may also be a determinant of both education and health; differences in biomedical history and genetic codes have been used in sibling populations to link health and education outcomes (e.g. Fletcher and Lehrer, 2009). Family background can also affect both education and health; for instance, in the U.S. Adams (2002) used family background as an instrument for education in analyzing the first wave of the Health and Retirement Study (HRS). The study found a positive effect of education on self-reported health status and functional status in both OLS and instrumental variable (IV) estimates.

To the extent that these common factors are heterogeneously distributed in the population, they could very well generate the observed association between education and health.

Recent studies have attempted to disentangle these mechanisms by using plausible sources of exogenous variation in education. Lleras-Muney (2005) explored the use of state variation in compulsory education laws from 1915 to 1939 as instruments for education in the U.S. She found that each additional year of education lowered the probability of dying in the subsequent 10 years by as much as 3.6 percentage points.[4] This result was much larger than that obtained using OLS methods.

Studies exploiting education reforms in other countries also look at other health markers. Using nationwide compulsory schooling law changes in the UK as instruments for education, Oreopoulos (2006) found a statistically significant relationship between education and self-

[4] Mazumder (2008) extended and performs Lleras-Muney's analysis by using data from the Survey of Income and Program Participation rather than from the U.S. Census. In contrast with Lleras-Muney's results, he found larger effects of education on health using compulsory schooling laws and including state-specific time trends.

reported health status. Again for the UK using compulsory schooling law changes, Silles (2009) found increased schooling leads to higher self-reported health and lower probabilities of long-term illness. Powdthavee (2010) found that education in UK lowers the prevalence of hypertension by 7-10 percentage points, when using the first change in the law (in 1947), but found no significant evidence when using the second law change (in 1973). Jürges et al. (2013) used two nationwide changes in minimum school leaving age in the UK as exogenous variation in education. The health outcomes considered included self-reported health and two biomarkers: blood fibrinogen and blood C-reactive protein. No causal effects were found between education and the two biomarkers. Further, the effect of education on self-reported health was positively significant only among the cohorts of older women, whereas it was negative among cohorts younger of women – and insignificant among men, regardless of age. Exploiting both the 1947 and 1973 changes to British compulsory schooling laws and using regression discontinuity methods, Clark and Royer (2010) found little evidence that additional schooling improved health outcomes or changed health behaviors.

Using a French longitudinal dataset, Albouy and Lequien (2009) used two increases in minimum school age in France as instruments for education. They failed to find a statistically significant causal effect of education on mortality. Employing Danish school reforms as an instrumental variable for education in a Danish panel dataset, Arendt (2005) found that the IV estimates of the effect of education on self-reported health and body mass index (BMI) were statistically insignificant, and not statistically different from those estimated using OLS. Finally, Kemptner et al. (2011) applied state variations in the timing of introducing a 9th grade as instruments for years of schooling. Using microcensus data from West Germany, they found that more years of schooling had a negative causal effect on long-term illness, work disability and obesity among men but not among women. Smoking behavior was not causally affected by education for either gender.[5]

Our research design is closest to Lleras-Muney (2005) and Kemptner et al. (2011). We use variation in the timing of education reform in more than 15 countries, which potentially generates considerable variation in education. Results tend to vary considerably from one health measure to the other. We look at a much larger set of health outcomes than previous studies, which includes diagnosed conditions, self-reported health and functional status. We use three data sets across different countries: the Health and Retirement Study (HRS) for the U.S., the English Longitudinal Study of Ageing (ELSA), and the Study of Health, Ageing and Retirement in Europe (SHARE). The survey questionnaires are highly similar, which yields datasets that include a wide range of health measures and a comparable education measure.[6]

We first show that the reforms experienced by the countries we look at have lead to higher education outcomes with varying levels of effectiveness. In some countries very little change in education was observed, while in others the reforms implemented were largely successful in raising education. We then use this variation in education compulsory laws to estimate the effect of education on health. Our results for self-reported health and functional status are largely consistent with existing studies. We find substantial beneficial effects of education on health outcomes. In terms of health conditions, we could find an effect of education on the prevalence of diabetes but no effects were found for other conditions (heart disease, hypertension, lung disease, stroke, arthritis, and psychiatric illnesses). Better education may lead to higher diagnosis rates, which could blur the relationship. In the case of diabetes, this implies

[5] More discussion and an extended survey of the use of compulsory schooling laws as instruments to study the causal effect of education on health can be found in Eide and Showalter (2011).

[6] Brunello et al. (2011) use compulsory schooling laws as instruments for 6 countries in SHARE as well as England, focusing on health behaviors.

that the effect of education on health dominates the diagnosis effect. We find the opposite for cancer, with a positive effect of education on prevalence. In this case, it is plausible that early detection leads to higher prevalence among those with greater education.

The paper is structured as follows. In section 2, we describe the data. In section 3, we present our estimation strategy, and in section 4 we present the estimation results. Finally, we conclude in section 5.

2. DATA AND DESCRIPTIVE ANALYSIS

2.1. HRS, SHARE and ELSA

We focus on respondents aged 50 and over in 15 countries: the United States, England, and thirteen continental European countries.[7] Our main data sources are three longitudinal surveys on aging: the Health and Retirement Study (HRS) in the U.S., the English Longitudinal Study of Ageing (ELSA) in England, and the Study of Health, Ageing and Retirement in Europe (SHARE). These surveys were specifically designed to be comparable with one another and all targeted people living in the community are aged 50 and over. Follow-up surveys were conducted biennially. We use data from wave 10 of HRS (2006), wave 3 of ELSA (2006), and wave 2 of SHARE (2006), all of which have been collected over the period 2006 to 2007[8]. The surveys provide harmonized data on health and socio-demographic variables relevant for our analysis. The age 50+ group is well suited for the analysis since at that age education level is

[7] SHARE countries included in our study are: Austria, Sweden, the Netherlands, Italy, France, Denmark, Greece, Switzerland, Belgium, Germany, Czech Republic, Poland and Spain.

[8] Currently, twelve waves of HRS (1992 - 2010), five waves of ELSA (2002 – 2010) and four waves of SHARE (2004 – 2010) are available. Wave 3 SHARELIFE (2008) is a retrospective survey. In order to keep comparability between the three surveys and maximize the number of countries to study, we do our analyses for wave 10 of HRS, wave 3 of ELSA and wave 2 of SHARE.

likely to be a permanent characteristic, and the potential effects of education on health have had time to occur.

All three surveys contain a large set of health status measures, most of which are comparable across surveys. From the data we constructed a set of subjective and objective measures of health outcomes. Subjective measures include overall self-rated health status, self-reported difficulties with activities of daily living (ADLs) as well as instrumental ADLs (iADLs). Self-rated health status was measured by asking respondents to rate their health on a five-point scale: excellent, very good, good, fair, poor. We defined a binary variable for "poor health", which takes the value of 1 if self-rated health is "fair" or "poor" and 0 otherwise. For limitations in ADLs, questions were asked in all surveys about difficulties in five basic activities: bathing, dressing, eating, getting in and out of bed, and walking across a room. Individuals were classified as having any ADL limitation if they reported limitations with one or more of the five activities. Limitations in iADLs were measured by questions about difficulties in the following five activities: preparing meals, shopping, making phone calls, taking medications and managing money. Those who reported having some difficulty with any of the five activities were classified as having an iADL limitation.

Objective measures included in all surveys were the same set of doctor-diagnosed disease questions on cancer, diabetes, hypertension, heart disease, stroke, lung disease, arthritis, and psychiatric illness.[9] We created a binary variable for "any illness" that takes a value of 1 when individuals report having any of these conditions. We also analyze the disease variables individually.

[9] The measure of "psychiatric illness" in SHARE is different from those in HRS and ELSA. HRS and ELSA ask the following question: "Have you ever had or has a doctor ever told you that you had any emotional, nervous, or psychiatric problems?" In SHARE the closest measure is taken from the following question: "Has there been a time or times in your life when you suffered from symptoms of depression which lasted at least two weeks?"

Our main independent variable is "years of education". In HRS, respondents were asked about the highest grade of school or year of college completed. In ELSA, the question relates to the age at which the individual finished full-time education. We converted the values into years of education by subtracting the usual school starting age of five from the age at which the respondent left school. In SHARE, respondents were asked directly about years of full-time education.

Other demographic variables include gender and age. To check the robustness of our results, we also control in some specifications for employment status (working versus non working), marital status (married/with partner, divorced/separated, widowed or never married), and household size.

2.1.1. Descriptive Statistics: Table 1 presents summary statistics of the data in more detail. We report the number of observations, the mean responses and standard deviations, and the minimum and maximum values. Thirty-two percent of the sample reported poor health, and 72% had one or more diagnosed conditions. The prevalence for specific health conditions ranges from 5% for stroke to 42% for hypertension. Regarding functional status, 14% reported having one or more ADL limitations, while 11% reported having one or more iADL limitations. Our key independent variable, "*years of education*", ranges from 0 to 25 years of education, with a mean of 11.13 years and a standard deviation of 3.91. The average age of the sample is 66 years old, ranging from 50 years to 89 years, and 46% of respondents are male.

Insert Table 1

2.1.2. Health and Education: Table 2 shows the unadjusted prevalence of health outcomes by education and by country. For ease of presentation, we recode years of education into three categories - tertiary, secondary, and primary or less - based on the educational system in each country. "Tertiary" is the category with the highest level of education, while "primary" indicates the lowest level of education. In the first column, we list the percentage reporting "*poor health*". In all countries, there is a clear gradient for the relationship between education and poor health, with those in the highest level of education reporting better health than those in the middle category, and these report better health than those in the lowest category. The second column shows the percentage of people with any of the health conditions mentioned. Americans report higher levels of disease than Europeans. England, Germany, Spain, Italy, France, Denmark, Poland, Czech Republic and Belgium report higher levels of disease than the other European countries. Countries with larger gradients in reporting conditions by education are the U.S., England, Greece, Italy, Austria, Spain, Poland, Czech Republic and Germany.

In columns 5 to 12, we show the prevalence of specific health conditions. The three most prevalent conditions are hypertension, arthritis and heart disease. Americans, English and Danish report higher percentages of cancer, arthritis and lung disease than other countries in each education category. Americans, Polish and Czech report higher percentages of diabetes in each educational category than the rest of the countries. For hypertension, arthritis, heart disease and lung disease, the low-educated report higher percentages than the high-educated in every country, the only exceptions being Poland and Switzerland for the latter condition (for which prevalence is the same). For diabetes, prevalence is also higher among the low-educated relative to the high-educated, here in 14 countries out of 15, the exception being Switzerland again. For cancer, prevalence is higher among the high-educated, relative to that among the low-educated, in four countries; it is similar between the two groups in six countries. For stroke, prevalence is higher in low-educated relative to that in high-educated in all countries with the exception of Poland and

Czech Republic. Finally, the patterns of self-reported psychiatric illness by education differ by country. In England, Sweden, France, Denmark, Czech Republic and Switzerland, the higher-educated individuals report more prevalent psychiatric illness; in the U.S., Austria, Spain, Italy and Greece, prevalence is higher among lower-educated individuals; and in Germany, the Netherlands, Belgium and Poland, prevalence is similar between the two groups.

Columns 3 and 4 show the percentages of people with difficulties for ADLs and iADLs. All countries present very high proportions of low-educated individuals with both ADL and iADL difficulties compared to the middle-educated and high-educated groups; with few exceptions (notably the Netherlands on both measures), there is a rather clear education gradient. As well, differences appear to be larger in England, the U.S., Poland, Austria, Germany and Sweden, and in every country as they relate to iADLs. Of course, self-reported health could be subject to various measurement errors and biases (see Jürges, 2007); our empirical analysis will deal with this by controlling for country specific effects, and clustering the standard errors by birth year and country.

Insert Table 2

Table 3 shows the correlations between health status and years of education without adjustment for other variables. The results are in line with the literature: there is a negative correlation between health and years of education. All the correlation coefficients are statistically significant at the 1% level. The more educated are less likely to report poor health, any health condition, or any ADL or iADL limitations. One exception is the positive correlation between having cancer and the level of education. At least one previous study also found the positive association between education and cancer (Cutler et al., 2008). Possible explanations are that more educated people are likelier to visit the doctor and are diagnosed earlier; survive longer; or have specific risk factors related to years of education, such as late childbearing among women. The correlation between self-reported health and years of education is stronger than the relationships between education and other health outcomes.

Insert Table 3

2.2. Compulsory Schooling Laws and Education

To examine the causal relationship between education and health, we use the cross-country variation in compulsory schooling laws over time as an instrument for years of education. Different compulsory schooling laws can affect education differently across birth cohorts and across countries in an exogenous way, given that the laws can differ through time and/or by country. Since individuals in our sample are aged 50 years and older, we consider compulsory schooling laws that would impact individuals born between 1905 and 1955. In 9 of the 15 countries in our analysis, there was a nationwide change in compulsory schooling law for cohorts born between 1905 and 1955. For the other countries there was either no such law change, or the change varied geographically within the country. We obtain information about law changes from different data sources. In England, the 1944 Education Act raised the minimum school leaving age from 14 to 15 years old, for cohorts born in April 1933 or later (Oreopoulos, 2006; Jürges et al., 2013; Silles, 2009); in France, minimum school leaving age was raised from 13 to 14 for those born after 1923 (Albouy and Lequien, 2009). Information on compulsory schooling and reforms for Austria, Greece, Italy, the Netherlands and Sweden was obtained from Murtin and Viarengo (2008). In Denmark, the schooling reform of 1958 reduced barriers to education and improved education attainment beyond seven years, especially for children from less educated background or the countryside. Although compulsory attendance remained at 7 years and was not raised to 9 years until the 1975 reform, there have been studies documenting that the

1958 reform raised education achievement in Denmark (Arendt, 2005). We therefore assume that the reform in 1958 is equivalent to a one-year increase in compulsory attendance. Finally, in the Czech Republic compulsory schooling was at 8 years from 1869. With the education reform in 1948, compulsory schooling age increased to 9 years (Filer et al., 1999). The six other countries are assumed to have had no compulsory schooling law changes that would affect the 1905 to 1955 birth cohorts. For Switzerland, Belgium and Spain, the documented compulsory education reforms took place in 1970 or later and did not affect the cohorts in our sample. The compulsory education reform in Poland was implemented in the 1960s and a small number of individuals were affected in our sample. For Germany and the United States, education reforms varied across geographic areas within the country and we are not able to define a beginning date for a nationwide reform in compulsory schooling law.

Table 4 reports average years of education by country, for those aged 50 and over in our sample. The table also shows for each country the years of compulsory attendance required before and after compulsory schooling law changes, as well as the first birth cohort that was subject to a modified compulsory schooling law. The average years of education for individuals aged 50 and over are lowest in Spain (7.38 years), and highest in Denmark (13.02 years).

Insert Table 4

Figure 1 shows the potential effect of compulsory schooling law changes on education levels, for each of the nine countries where law changes took place between year 1905 and 1955. We show the average number of years of education for birth cohorts born 10 years before and up to 10 years after the first cohort affected by the reforms. For England, Denmark, the Netherlands and Sweden, the reforms appear to have shifted the average education level upwards by approximately half a year. For the other five countries the effect is less obvious. Specifically, for

Greece and Italy, only three birth cohorts in our sample were affected by law changes so the sample sizes are limited. For Czech Republic, it seems that the schooling reforms did not have increase effect. However, as a whole there is evidence that the schooling reforms increase years of education. We will confirm this relationship below using econometric models.

Insert Figure 1

In Figure 2, we draw the reduced-form relationship between compulsory schooling law changes and one of the health outcomes, i.e., "poor health". We pool the data from the nine countries with law changes and calculate the proportion reporting poor health by birth cohort, for individuals born 5 years before and 5 years after the first cohort affected by law changes, adjusting for gender, cohort and country. There is a sharp reduction in the proportion reporting poor health for the cohorts affected by the reforms, and the downward shift persists after the cohort initially affected. This raw evidence would suggest an important effect of education on health.

Insert Figure 2

3. EMPIRICAL STRATEGY

We first model the effect of education on different health outcomes using a Probit model. $H_{j,i}$ indicates health outcome j, a binary measure, for an individual i; it takes the value of 1 if the underlying latent variable, $H^*_{j,i}$, is positive and a value of 0 otherwise. Ed_i represents education for the individual, measured as years of education obtained. X_i contains a set of demographic variables: gender, birth cohort dummies for nine age groups, and country. Pooling the three data sets (from HRS, ELSA and SHARE), we estimate models for the latent variable $H^*_{j,i}$ for all health outcomes. For example, we estimate the probability that an individual is in "poor health" or the probability that an individual has any disease using the following model:

$$H^{*}_{j,i} = \alpha_j + Ed_i \beta_j + X_i \gamma_j + \varepsilon_j$$

$$H_{j,i} = I(H^{*}_{j,i} \geq 0) \qquad (1)$$

where ε_j is a random error that is normally distributed. We add country fixed effects to the specification.

However, education can be endogenous. As mentioned in the introduction, different factors can drive this endogeneity, such as reverse causality or unobserved heterogeneity. This potential endogeneity can be addressed with an instrumental variable Probit model. The relationship between health and education can be estimated in a model that takes into account the possibility that education might be endogenous. The model is then both, the set of equations (1) identical to the previous model and the equation (2) below. In equation (2), we model education as a function of the demographic variables, X_i, as well as of Z_i, which is the minimum number of years of education required for a given individual and which varies by country and birth cohort.

$$Ed_i = \theta + X_i \lambda + Z_i \phi + \nu_j \qquad (2)$$

ε_j, ν_j are random errors that are normally distributed. The key coefficient of interest is β_j. If education is exogenous, the Probit estimation of equation (1) generates an unbiased estimation of β_j. However, education might be endogenous. In our instrumental variable approach to estimate β_j, the variable Z_i, the minimum number of years of education required by law, is the instrument. Since we control for country and birth cohort in both stages of the model, the effect of Z_i on Ed_i is estimated "net" of the country- and cohort-specific effects.

For our instrument to be valid, it should be positively correlated with years of education – this is indeed suggested by figure 1. In addition, Z_i should not affect health outcomes other than through its effect on years of education. This cannot be directly tested as we only have one instrument[10]. However, it seems a reasonable assumption that there were no co-occurring factors that affected both compulsory schooling laws and health; the use of compulsory schooling law changes from multiple countries reduces the possibility of such co-occurrence. Our empirical strategy is very similar to Lleras-Muney's (2005): we use the compulsory schooling minimum age law changes across different countries and we analyze the effects of these changes on education and different health outcomes. We study the different variability in countries where education reforms were implemented compared to countries where no reform was done, the latter being used as a control group.

So we first estimate equation (1) for each health outcome using a Probit. For the instrumental variable approach, we jointly estimate (2) and (3) using maximum likelihood and assuming ε_j, ν_j are multivariate normal with a correlation coefficient ρ_j. We then test whether ρ_j is statistically different from zero. If it is, we may reject the null hypothesis that the education variable is exogenous. If ρ_j is not significantly different from zero, then we cannot reject the null hypothesis, and the multivariate Probit model estimates based on (1) are appropriate and might yield smaller standard errors. All analysis is conducted using Stata Statistical Software, release 11.0, Special Edition. Probit and IV-Probit models are estimated using the "probit" and "ivprobit" commands in Stata. The "ivprobit" command estimates " atanh" of ρ_j instead of ρ_j:

[10] With multiple instruments at hand one can apply an over-identification test.

$a\tanh(\rho_j) = \ln((1+\rho_j)/(1-\rho_j))$. While the range for ρ_j is (-1, 1), $a\tanh(\rho_j)$ varies from negative infinity to positive infinity, which eliminates the need for adding a constraint to the estimation. To test the exogeneity of the education variable, we test whether $a\tanh(\rho_j)$ is statistically different from zero.

4. RESULTS

4.1. Health and Education across countries

As we have shown in the previous sections there is much evidence in the literature about the strong correlation between health and education. We first estimate model (1) above to replicate this evidence across countries and health outcomes, using HRS, ELSA and SHARE data sets. The main results are reported in Table 5, which shows the marginal effects of years of education on health. The coefficients are all negative and significant at the 1 % level, meaning that more education is associated with a lower probability of having health problems. The only exception is cancer, for which the coefficient is positive and significant. For self-reported health, each additional year of schooling is associated with a 2.8 percentage points reduction in the probability of reporting poor health. Each additional year of schooling is also associated with a 1 percentage point reduction in the probability of having ADL or iADL limitations. As for health conditions, the marginal effects range from a 1.2 percentage points reduction for arthritis to a 0.8 percentage point reduction for diabetes, a 0.2 percentage point reduction for stroke and a 0.2 percentage point increase for cancer. All models control for gender and birth cohort. In addition, country dummies are included in all specifications to account for institutions and cultural differences. The complete results are available from the authors upon request.

Insert Table 5

4.2. Causal Relationship between Health and Education

We next turn to instrumental variable estimation to examine the effects of education on health. Table 6 shows the first- and second-stage estimates for each of the health outcomes. Marginal effects and robust standard errors are displayed.

The first-stage estimation is a linear regression of the individual's years of education against the minimum number of years of education required by compulsory schooling laws, controlling for gender, birth cohort, and country. The estimate is statistically significant at the 1% level. The F-statistic is about 27. Raising the minimum number of years of education required by one year increased the average length of education by 0.35 years (around 4 months).

Insert Table 6

The second-stage estimation is a Probit model of each health outcome against years of education, gender, birth cohort, and country. The results are mixed for the second-stage estimates. For 4 of the 12 health outcomes, i.e., poor health, ADL and iADL limitations and diabetes, the effect of education remains negative and statistically significant, for poor health, ADL and iADL at the 1% level and for diabetes at the 5% level. For cancer, the effect of education is positive and statistically significant at the 1% level. The magnitudes of point estimates are much larger using IV estimation. For example, the marginal effect of education on the probability of reporting poor health increases from 2.8 to 6.3 percentage points, and that on functional status – both ADL and iADL limitations – increases from 1.0 to 3.2 and 4.2 percentage points, respectively, non-negligible effects. The changes in estimates for cancer and diabetes are more dramatic. It is not uncommon that IV estimates are larger, probably due to heterogeneous treatment effects or measurement errors in reported years of education (Card, 2001). Exogeneity test statistics for self-reported poor health, ADLs and iADLs and cancer are all statistically significant; for poor health at the 1% level, for iADLs and cancer at the 5% level and for ADLs at the 10% level, meaning that for these health outcomes we can reject the null hypothesis that education is an exogenous variable.

For the other seven outcomes, the effects are no longer significant. However, exogeneity test statistics for those seven outcomes and diabetes are insignificant, meaning there is not sufficient evidence to treat education as an endogenous variable in the analysis. Therefore, in those cases the Probit estimates might be consistent and more efficient.

4.3. Robustness

We have replicated our analysis with different specifications for age, using age and age quadratic terms instead of birth cohort dummies. Our results were robust. We also controlled for additional socioeconomic variables as described in section 2.1, including employment status, marital status, and household size. The coefficients were a bit smaller in magnitude but qualitatively similar. In another set of regressions, we controlled for whether parents were alive at interview time, in both the Probit and IV-Probit models. The rationale is that parents' survival reflects family background and genetic factors, which could be correlated with both education and health. The results for the Probit and IV-Probit models were qualitatively unchanged. The results were also robust when including two instruments, adding the minimum compulsory schooling age as a quadratic term. Finally, we also ran regressions with Linear Probability models, with compulsory schooling laws as the instrument. In this case we found the effect of education on diabetes was no longer statistically significant, while other estimates were qualitatively unchanged.

5. CONCLUSION

This paper studies the causal relationship between education and health outcomes. We combine three surveys that include nationally representative samples of individuals aged 50 and over from 15 OECD countries. We use differences in educational reforms across these countries as an instrument for the number of years of education achieved by individuals. A number of recent papers studied this problematic, and our results are in line with other papers in the literature lead by Lleras-Muney (2005): we find a causal effect of education on some health measures. In particular, we find that more years of education lead to a lower probability of reporting poor health and functional status (ADL and iADL limitations) and to a lower prevalence for diabetes. These effects are larger than the simple Probit estimates. The causal relationship between education and several other health conditions, i.e., heart disease, hypertension, lung disease, stroke, arthritis, and psychiatric illness, is statistically insignificant but not different in magnitude from Probit estimates. Both Probit and IV estimates indicate that more education leads to higher reported rates of cancer diagnosis. The relationships between education and self-reported poor health and functional status are very robust and most likely to be causal. The relationship between education and specific diagnosed health conditions is more uncertain and requires further investigation. Our analysis focuses on health outcomes but an extension will be to examine whether compulsory schooling laws can affect life expectations through education. Estimating the causal impact of education on health is of important policy interest, since education might be a cost effective way to improve health.

6. REFERENCES

Adams, S. J. 2002. Educational Attainment and Health: Evidence from a Sample of Older Adults. *Education Economics* 10(1): 97-109.

Albouy, V. and L. Lequien. 2009. Does compulsory education lower mortality? *Journal of Health Economics* 28(1): 155-168.

Andreyeva, T., P.C. Michaud and A. van Soest. 2007. Obesity and Health in Europeans aged 50 years and older. *Public Health* 121, pp. 497-509.

Arendt, J. N. 2005. "Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification." *Economics of Education Review* 24(2): 149-160.

Banks, J., M. Marmot et al. 2006. Disease and disadvantage in the United States and in Britain. *JAMA* 295(17): 2037-2045.

Brunello, G., M. Fort, N. Schneewei and R. Winter-Edmer. 2011. The Causal Effect of Education on Health. What is the role of Health Behaviors? CEPR Discussion Paper No. DP8707.

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

Case, D. and A. Deaton. 2005. Broken Down by Work and Sex: How our Health Declines, in D.A. Wise (ed.), *Analysis in the Economic of Aging,* The University of Chicago Press, pp. 185-212.

Clark, D. and H. Royer. 2010. The Effect of Education on Adult Health and Mortality: Evidence from Britain. NBER Working Paper No. 16013.

Cutler, D.M. and A. Lleras-Muney. 2006. Education and Health: Evaluating Theories and Evidence. National Bureau of Economic Research, Inc, NBER Working Papers: 12352.

Cutler, D.M., A. Lleras-Muney and T. Vogl. 2008. Socioeconomic Status and Health: Dimensions and Mechanisms. National Bureau of Economic Research, Inc, NBER Working Papers: 14333.

Cutler, D.M. and A. Lleras-Muney. 2010. Understanding differences in health behaviors by education. *Journal of Health Economics* 29(1): 1-28.

Eide, E. R. and M.H. Showalter. 2011. Estimating the relation between health and education: What do we know and what do we need to know? *Economics of Education Review* 30, 778-791.

Filer, R.K., S. Jurajda and J. Planovsky. 1999. Education and Wages in the Czech and Slovak Republics during transition. *Labour Economics* 6 (1999): 581-593.

Fletcher, J. M. and S. F. Lehrer. 2009. Using genetic lotteries within families to examine the causal impact of poor health on academic achievement. National Bureau of Economic Research, Inc, NBER Working Papers: 15148.

Fuchs, V. R. 1982. Time Preference and Health: An Exploratory Study, National Bureau of Economic Research, Inc, NBER Working Papers: 0539.

Grossman, M. 1972. On the Concept of Health Capital and the Demand for Health. *The Journal of Political Economy*, Vol.80, No. 2, pp. 223-255.

Grossman, M. 2005. Education and Nonmarket Outcomes. National Bureau of Economic Research, Inc, NBER Working Papers: 11582.

Grossman, M. and R. Kaestner. 1997. Effects of Education on Health, in *The Social Benefits of Education*, ed. by J. R. Behrman and N. Stacey, pp. 69-124. University of Michigan Press, Ann Arbor, MI.

Huisman, M., A.E., Kunst and J.P.Mackenbach. 2005. Educational Inequalities in Smoking Among Men and Women Aged 16 Years and Older in 11 European Countries. *Tobacco Control* 14:106-13.

Jürges, H. 2007. True health vs response styles: exploring cross-country differences in self-reported health. *Health Economic* 16 (2): 163-178.

Jürges, H., E. Kruk and S. Reinhold. 2013. The effect of compulsory schooling on health – evidence from biomarkers. *Journal of Population Economics* 26(2): 645-672.

Kemptner, D., H. Jürges and S. Reinhold. 2011. Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *Journal of Health Economics* 30(2): 340-354.

Lleras-Muney, A. 2005. The Relationship between Education and Adult Mortality in the United States. *Review of Economic Studies* 72(1): 189-221.

Mackenbach, J. P., I. Stirbu et al. 2008. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 358(23): 2468-2481.

Mazumder, B. (2008). Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives* 33(2): 2–16.

Michaud, P.C., D. Goldman, D. Lakdawalla, A. Gailey and Y. Zheng. 2011. Differences in Health between Americans and Western Europeans: Effects on Longevity and Public Finance. *Social Science and Medicine* 73(2): 254-263.

Murtin, F. and M. Viarengo. 2008. The convergence process of compulsory schooling in Western Europe: 1950-2000. London School of Economics, Centre for the Economics of Education Discussion Paper 95.

Oreopoulos, P. 2006. Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. *American Economic Review* 96(1): 152-175.

Powdthavee, N. 2010. Does Education Reduce the Risk of Hypertension? Estimating the Biomarker Effect of Compulsory Schooling in England. *Journal of Human Capital*, vol. 4, no. 2.

Silles, M. A. 2009. The Causal Effect of Education on Health: Evidence from the United Kingdom. *Economics of Education Review* 28(1): 122-128.

Figures

Figure 1. Average years of education by year of birth and by country

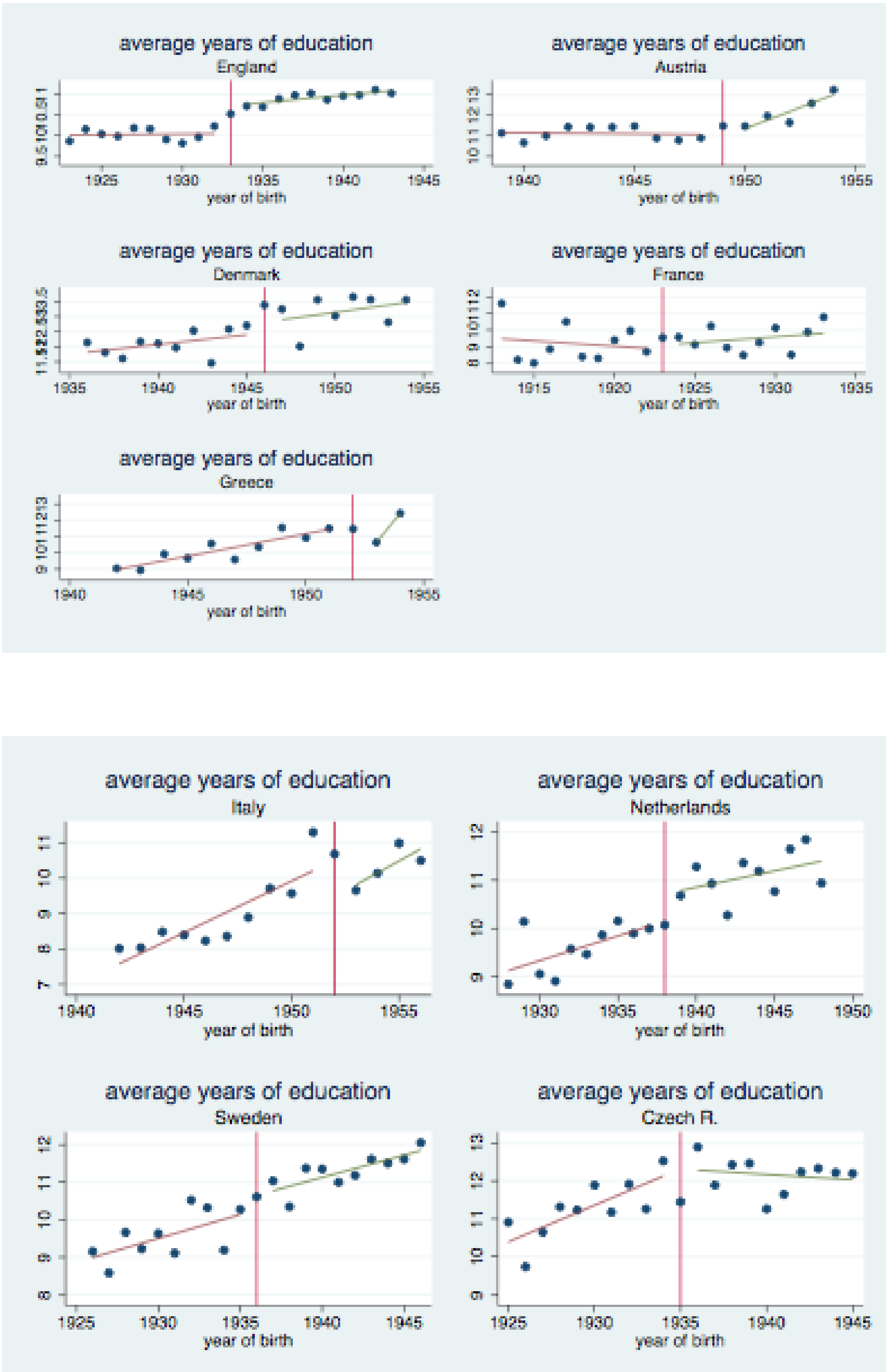
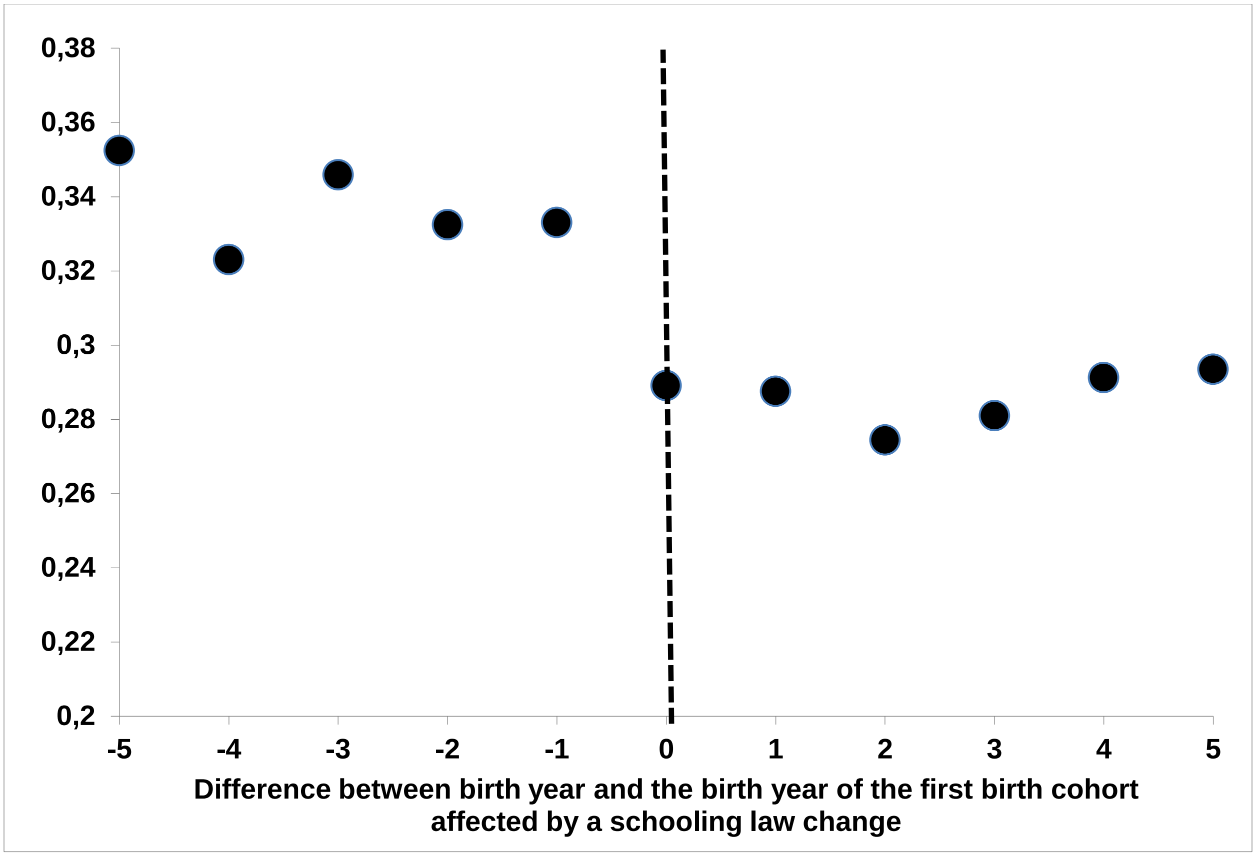


Figure 2. Adjusted proportion reporting poor health by birth cohort



Average proportion of individuals reporting poor health is based on a Probit regression of self-reported poor health against gender, birth cohort dummies, and country dummies. The sample includes individuals born up to five years before and after the birth cohort first affected by a schooling law change.

Data source: HRS wave 10, ELSA wave 3, and SHARE wave 2

Data are weighted by sampling weight (normalized by country)

Tables

Table 1. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Health variables					
Poor health	56800	0,32	0,46	0	1
1+ Disease	56942	0,72	0,45	0	1
1+ ADLs	56815	0,14	0,34	0	1
1+ IADLs	56814	0,11	0,32	0	1
Cancer	56768	0,08	0,27	0	1
Diabetes	56821	0,13	0,34	0	1
Heart disease	56820	0,18	0,38	0	1
Hypertension	56836	0,42	0,49	0	1
Arthritis	56830	0,36	0,48	0	1
Lung disease	56823	0,07	0,26	0	1
Stroke	56832	0,05	0,22	0	1
Psychiatric illness	56688	0,16	0,37	0	1
SES variables					
Years of education	55381	11,13	3,91	0	25
Age	56942	66,01	10,50	50	89
Cohort	56942	3,76	1.96	1	8
Male	56942	0,46	0,50	0	1
Marital status	55294	1,55	0,92	1	4
Employment status	56570	0,31	0,46	0	1
Household size	56942	2,18	1,05	1	14
Family background					
Mother alive	55340	0,08	0,27	0	1
Father alive	55203	0,20	0,40	0	1

Data source: HRS wave 10, ELSA wave 3, and SHARE wave 2

Data are weighted by sampling weight (normalized by country)

Table 2. Health outcomes by level of education and by country

		Poor health	1+ Disease	1+ ADLs	1+ IADLs	Cancer	Diabetes	Heart disease	Hypertension	Arthritis	Lung disease	Stroke	Psychiatric illness
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
England	*High*	16%	65%	8%	4%	8%	7%	17%	36%	25%	3%	2%	12%
	Medium	25%	69%	14%	8%	7%	8%	20%	40%	33%	5%	4%	11%
	Low	43%	81%	26%	18%	9%	12%	26%	51%	46%	10%	7%	10%
United States	*High*	16%	80%	10%	8%	13%	15%	20%	48%	51%	7%	6%	16%
	Medium	28%	88%	15%	13%	13%	19%	25%	57%	62%	12%	7%	18%
	Low	52%	91%	27%	25%	16%	28%	31%	65%	68%	15%	12%	24%
Austria	*High*	20%	45%	6%	4%	2%	8%	8%	27%	7%	2%	3%	6%
	Medium	30%	57%	9%	6%	3%	10%	10%	36%	15%	4%	2%	10%
	Low	48%	66%	18%	20%	2%	15%	16%	39%	19%	7%	4%	14%
Germany	*High*	28%	57%	7%	3%	5%	8%	11%	35%	10%	5%	3%	15%
	Medium	41%	63%	11%	8%	4%	13%	12%	38%	14%	6%	4%	15%
	Low	60%	70%	20%	20%	5%	20%	14%	44%	18%	6%	5%	16%
Sweden	*High*	17%	51%	3%	1%	6%	4%	11%	23%	10%	1%	3%	18%
	Medium	29%	55%	6%	5%	5%	9%	12%	30%	10%	4%	4%	13%
	Low	36%	64%	13%	10%	6%	11%	20%	38%	12%	4%	5%	13%
Netherlands	*High*	20%	49%	5%	5%	5%	6%	7%	21%	7%	4%	2%	17%
	Medium	25%	55%	4%	3%	4%	7%	10%	23%	12%	5%	3%	20%
	Low	37%	62%	10%	9%	4%	13%	11%	31%	12%	8%	5%	18%
Spain	*High*	22%	54%	2%	1%	3%	7%	3%	30%	17%	4%	0%	17%
	Medium	22%	52%	4%	3%	1%	10%	3%	23%	15%	5%	1%	17%
	Low	52%	70%	14%	11%	2%	16%	10%	36%	32%	7%	3%	26%
Italy	*High*	17%	48%	5%	5%	5%	1%	8%	27%	10%	3%	1%	12%
	Medium	30%	58%	5%	4%	4%	8%	9%	32%	21%	3%	2%	16%
	Low	53%	75%	14%	13%	3%	14%	13%	45%	40%	9%	4%	20%

To be continued on next page

Table 2. Health outcomes by level of education and by country (continued)

		Poor health	1+ Disease	1+ ADLs	1+ IADLs	Cancer	Diabetes	Heart disease	Hypertension	Arthritis	Lung disease	Stroke	Psychiatric illness
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
France	*High*	19%	59%	4%	2%	5%	4%	9%	19%	16%	2%	2%	28%
	Medium	31%	65%	8%	5%	5%	9%	8%	28%	24%	4%	1%	23%
	Low	45%	69%	14%	14%	5%	11%	15%	31%	30%	6%	4%	21%
Denmark	*High*	16%	61%	5%	5%	6%	4%	9%	28%	24%	4%	3%	21%
	Medium	25%	66%	7%	6%	7%	9%	11%	35%	27%	7%	5%	18%
	Low	39%	75%	15%	19%	7%	10%	17%	39%	36%	11%	9%	19%
Greece	*High*	9%	40%	3%	3%	2%	7%	6%	22%	6%	1%	2%	7%
	Medium	16%	49%	3%	2%	2%	8%	7%	28%	11%	2%	1%	10%
	Low	34%	66%	9%	8%	2%	13%	14%	41%	23%	5%	3%	11%
Switzerland	*High*	12%	46%	5%	2%	3%	8%	7%	22%	8%	3%	2%	19%
	Medium	14%	52%	4%	2%	5%	5%	4%	26%	11%	3%	2%	16%
	Low	26%	56%	10%	7%	4%	7%	9%	33%	13%	3%	4%	12%
Belgium	*High*	20%	60%	8%	4%	4%	8%	10%	32%	21%	4%	2%	15%
	Medium	25%	61%	8%	4%	3%	7%	12%	32%	21%	4%	2%	14%
	Low	37%	67%	17%	13%	4%	11%	15%	37%	26%	7%	4%	16%
Czech R.	*High*	29%	74%	6%	1%	4%	11%	10%	40%	10%	2%	6%	41%
	Medium	37%	70%	7%	4%	4%	13%	11%	43%	14%	4%	4%	34%
	Low	52%	77%	9%	9%	5%	16%	17%	45%	19%	6%	5%	33%
Poland	*High*	42%	66%	11%	7%	4%	10%	20%	42%	20%	6%	9%	20%
	Medium	55%	69%	16%	9%	3%	9%	16%	38%	29%	4%	4%	21%
	Low	74%	79%	32%	26%	2%	14%	25%	48%	41%	6%	7%	20%

Data source: HRS wave 10, SHARE wave 2, and ELSA wave 3

Data are weighted by sampling weight (normalized by country)

Table 3 Correlation coefficients between years of education and health outcomes

Poor health	1+ Disease	1+ ADLs	1+ IADLs	Cancer	Diabetes
-0.249	-0.0479	-0.1186	-0.1255	0.0595	-0.06
Heart disease	Hypertension	Arthritis	Lung disease	Stroke	Psychiatric illness
-0.0338	-0.0459	-0.0176	-0.05	-0.0218	-0.0169

All Spearman correlations are significant at the 1% level

Data source: HRS wave 10, ELSA wave 3, and SHARE wave 2

Data weighted by sampling weight (normalized by country)

Table 4. Average years of education and minimum years of education required before and after compulsory schooling law changes, by country

Country	Average Years of Education	Minimum years of education required before compulsory schooling law change	First birth cohort affected by compulsory schooling law change	Minimum years of education required after compulsory schooling law change
Austria	8.57	8	1949	9
England	10.8	9	1933	10
Sweden	11.19	8	1936	9
Netherlands	11.05	6	1938	8
Italy	7.92	5	1952	8
France	11.18	7	1923	8
Denmark	13.07	7	1946	8
Greece	8.36	6	1952	9
Czech Republic	12.21	8	1935	9
Poland	9.08	7	No nationwide compulsory schooling law change	
Switzerland	11.2	8	No nationwide compulsory schooling law change	
Belgium	11.67	8	No nationwide compulsory schooling law change	
Germany	12.29	8	No nationwide compulsory schooling law change	
Spain	7.41	7	No nationwide compulsory schooling law change	
United States	12.85	8.42	No nationwide compulsory schooling law change	

Average years of education are from HRS, ELSA and SHARE weighted data (normalized by country). Information on minimum years of education required and compulsory schooling law changes is mainly obtained from Murtin and Viarengo (2008), with the following exceptions: for Britain, the compulsory schooling law and changes have been described in several papers including Jürges (2009); for Denmark, the information comes from Arendt (2005) and Murtin and Viarengo (2008); for France, the compulsory schooling law reform was described in Albouy and Lequien (2009); for Czech Republic, the information comes from Filer et al. (1999); for the United States, compulsory schooling laws varied by state and changed at different points in time from 1915 to 1939, as described in Lleras-Muney (2005). We calculated the average minimum years of education in all countries during this period, based on the "Compulsory Attendance and Child Labor State Laws" dataset, provided by Adriana Lleras-Muney on her website: http://www.econ.ucla.edu/alleras/research/data.html

Table 5. Probit models: effect of years of education on health outcomes (Coefficients reported as marginal effects)

	Poor health	1+ Disease	1+ ADLs	1+ IADLs	Cancer	Diabetes	Heart disease	Hypertension	Arthritis	Lung disease	Stroke	Psychiatric illness
Years of education	-0.0278***	-0.008***	-0.010***	-0.0097***	0.002***	-0.008***	-0.0045***	-0.0098***	-0.012***	-0.0054***	-0.0016***	-0.0037***
	(0.00128)	(0.00071)	(0.00061)	(0.00047)	(0.00036)	(0.00052)	(0.0005)	(0.00080)	(0.00085)	(0.00038)	(0.00027)	(0.00071)
Number of observations	55 259	55 381	55 295	55 294	55 252	55 303	55 302	55 318	55 313	55 306	55 315	55 186

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Robust standard errors in parentheses, clustered at birth year-country level

All models control for gender, cohort (dummies), and country (dummies)

Data source: HRS wave 10, SHARE wave 2, and ELSA wave 3

Data are weighted by sampling weight (normalized by country)

Table 6. IV-Probit models: effect of years of education on health outcomes (Coefficients reported as marginal effects)

	Poor health	**1+ Disease**	**1+ ADLs**	**1+ IADLs**	**Cancer**	**Diabetes**
First stage: dependent variable is years of education						
Minimum years of schooling required	0.356***	0.355***	0.354***	0.354***	0.354***	0.355***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
Second stage						
Years of education	-0.0626***	-0.0059	-0.0318***	-0.0415***	0.025***	-0.0256**
	(0.0067)	(0.038)	(0.009)	(0.0112)	(0.0115)	(0.040)
Exogeneity test						
α tanh (ρj)	0.458***	-0.017	0.346*	0.551**	-0.488**	0.274
	(0.115)	(0.125)	(0.148)	(0.186)	(0.183)	(0.145)
N.obs.	55381	55295	55294	55294	55252	55303
	Heart disease	**Hypertension**	**Arthritis**	**Lung disease**	**Stroke**	**Psychiatric illness**
First stage: dependent variable is years of education						
Minimum years of schooling required	0.355***	0.355***	0.355***	0.355***	0.355***	0.352***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.047)
Second stage						
Years of education	-0.0071	-0.0122	0.012	-0.0073	0.004	-0.0017
	(0.011)	(0.0129)	(0.0143)	(0.0076)	(0.0082)	(0.0076)
Exogeneity test						
α tanh (ρj)	0.036	0.033	-0.241	0.04	-0.199	-0.026
	(0.156)	(0.117)	(0.156)	(0.185)	(0.252)	(0.172)
N.obs.	55302	55318	55313	55306	55315	55186

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at birth year-country level. All models control for gender, cohort (dummies), and country (dummies). Data source: HRS wave 10, SHARE wave 2, and ELSA wave 3. Data are weighted by sampling weight (normalized by country).