The Relationship Between Education and Adult Mortality in the U. S.

Adriana Lleras-Muney*

May 2, 2001

Abstract

Prior research has uncovered a large and positive correlation between education and health, but there are difficulties in determining whether this relationship is causal. In this paper I reexamine whether education has a causal impact on health. I follow synthetic cohorts using successive U.S. censuses to estimate the impact of educational attainment on mortality rates. I then use compulsory education laws from 1915 to 1939 as instruments to obtain a consistent causal estimate of this effect. While least squares estimates suggest that an additional year of education lowers the probability of dying in the next 10 years by approximately 1.3 percentage points, results from the IV estimation show that the effect is in fact much larger, at least 3.6 percentage points. Overall, one more year of education increases life expectancy at age 35 by 1.2 years. These results provide evidence of a causal effect from education to health and suggest that the income returns to education substantially underestimate the overall returns to education.

JEL: I12, I20, J10, J18, N32, N42

*Email: al277@columbia.edu. Department of Economics, Columbia University, New York, NY 10027. I would like to thank Ann Bartel, Francisco Ciocchini, Ana Corbacho, Rajeev Dehejia, Phoebus Dhrymes, William Gentry, Kenneth Leonard, Manuel Lobato, Alexander Peterhansl, Nachum Sicherman and the seminar participants at the Federal Reserve Board, Michigan State University, University of Montreal, UNC-Greensboro, Princeton University, UC Berkeley, UC Davis, UC Irvine, University of Illinois at Urbana Champaign for their comments and suggestions. I am especially grateful to my advisor Sherry Glied who fully supported me throughout this project. All errors are mine. This research was partially funded by Columbia University's Public Policy Consortium, and the Bradley Foundation.

1. Introduction

Access to health care insurance,[1] expenditures on health care,[2] and even income levels[3] have been shown to have little effect on health. On the other hand, there is a large and positive correlation between education and health (Grossman and Kaestner 1997). This correlation is strong and significant even after controlling for different measures of socio-economic status, such as income and race, and regardless of how health is measured (morbidity rates, self-reported health status or other measures of health). Given that the measured effects of education are large, investments in education might prove to be a cost-effective means of achieving better health,[4] if education indeed helps us to be healthier. But prior research has not ascertained whether the relationship between education and health is causal.

The purpose of this paper is to determine whether education has a causal effect on health, in particular on mortality. Recent studies[5] suggest that the value of a healthy life is very large. The negative relationship between education and mortality, the most basic measure of health, has become well established since the famous Kitagawa and Hauser (1973) study, which found significant differences in mortality rates across educational categories for both sexes. More recent studies (e.g. Feldman et al., 1989, Pappas et al., 1993, Christenson and Johnson, 1995, Deaton and Paxson, 1999) confirm these findings. Elo and Preston (1996) control for a variety of other mortality factors such as income, race, marital status, region of residence, and region of birth. Rogers et al. (2000) further control for access to health care, insurance, smoking, exercise, occupation, and other factors. Figures 1 and 2 document this relationship using consecutive census data for the US: in all cohorts, those who survive have higher education than those who do not.

The existing literature has explained this correlation in three ways. One controversial hypothesis is that education increases health, either because education makes people better decision-makers (Grossman 1975) and/or because more educated people have better information about health (Kenkel 1991, Rosenzweig and Schultz 1981).[6] Another possibility is that poor health results in little education (Perri 1984, Curry and Hyson 1999). Finally, this correlation could be caused by a third unobserved variable that affects both education and health, for example genetic characteristics or parental background. Many studies have attempted to include these factors.[7] How-

[1] See Newhouse (1993).

[2] For example see Filmer and Prichett (1997).

[3] Grossman (1975) shows that income levels beyond a certain minimum did not have an impact on health outcomes.

[4] That education might be a more cost-effective means to increase health than increasing medical care expenditures was first suggested by Auster et al (1969). Their findings suggest that the elasticity of mortality rates with respect to education is twice as large as that of medical expenditures.

[5] For example Nordhaus (1999).

[6] Kenkel (1991) and Gilleskie and Harrison (1998) find evidence to support both Grossman's and Kenkel's hypotheses.

[7] Some studies suggest that adding controls will eliminate the observed relationship. Behrman et al (1998) find that using a random effects model to control for unobserved heterogeneity, the effect of education on mortality disappears. Other examples include Wolfe and Behrman (1987), Duleep (1986) and Menchik (1993).

ever, Fuchs (1982) argued that discount rates (which no study controls for) would also explain the correlation: people who are impatient invest little in education and health, while people who are patient invest a lot in both.[8] Of course, these theories are not necessarily mutually exclusive.

In this paper I address this issue using a unique quasi-natural experiment: between 1915 and 1939, at least 30 states changed their compulsory schooling laws and child labor laws. If compulsory schooling laws forced people to get more schooling than they would have chosen otherwise, and if education increases health, then individuals who spent their teens in states that required them to go to school for more years should be relatively healthier and live longer.

This natural experiment allows me to use these laws as instruments for education and thus to identify the causal relationship between education and mortality. These laws are valid as a natural experiment for various reasons. The 1914-1939 period saw the largest historical increases in the number of students graduating from high school. There were also a significant number of changes in the laws at that time. Historians suggest that the laws were enforced during this time, and therefore were likely to have affected many individuals. These laws were external to individuals;and there were no systematic differences between children in the states where the laws required (for example) 8 years of education and children in the states which required 9 years.[9] Finally, the laws are very likely to be uncorrelated with health. Nonetheless, I control for state-of-birth, cohort year, state expenditures on education in state-of-birth, and a host of other state factors that might be correlated with the laws and health outcomes.

The intuition that compulsory education laws provide a natural experiment was put forward first by Angrist and Krueger (1991). They argued that because compulsory education laws forced individuals to stay in school until a certain age, those born in later quarters would stay in school longer. Although they were criticized for their choice of quarter of birth as an instrument,[10] the underlying principle is appealing and implementable. Other researchers have successfully used compulsory education laws as instruments in the context of the returns to education for other countries.[11]

No other papers have used natural experiments to measure the effect of education on mortality. A few studies (Berger and Leigh 1988, Sander 1995, and Leigh and Dhir 1997) have used instrumental variable (IV) estimation with other measures of health, such as blood pressure, smoking or exercise.[12] But these studies are inconclusive

[8] Fuchs (1982) tried to measure discount rates (through a telephone survey), and use them to predict education and health. His results were mixed. Farrel and Fuchs (1982) examined the issue again, but the evidence they provide is indirect. Munasinghe and Sicherman (2000) however provide evidence which suggests that time preference plays an important role in the determination of smoking and earnings growth.

[9] See Lleras-Muney (2001) for a more detailed analysis of the effect of these laws on educational attainment.

[10] See Bound, Jaeger and Baker (1995) and Bound and Jaeger (1996).

[11] Harmon and Walker (1995) look at the effects of the laws in the UK. Meghir and Palme (1999) used Swedish data. Acemoglu and Angrist (1999) used compulsory education laws in the US as instruments for average education at the state level to determine the size of the social returns to education.

[12] Berger and Leigh (1988) estimate the effect of education on blood pressure using the NHANES I.

because each paper's choice of instrument is questionable. For example, all of these studies use parents' background/education as instruments, but we know these are correlated with children's health,[13] and furthermore, we know that health shocks during childhood or gestation have persistent health effects into adulthood.[14] Income and education expenditures in state-of-birth could serve as instruments (Berger and Leigh 1988), but again they might be correlated with state expenditures on health, state industrial composition and other state characteristics that affect health.

Using the 1960, 1970 and 1980 Censuses of the US, I select those individuals who were 14 years of age between 1915 and 1939. I then construct synthetic cohorts and follow them over time to calculate their mortality rates. I then match cohorts to the compulsory attendance and child labor laws that were in place in their state-of-birth when they were 14 years old. The census data have not been used to calculate mortality rates before in economic analyses,[15] but this methodology has many advantages. Because census data are very extensive and go back well into the 19th century, this method could be used to analyze mortality experiences in periods where no other data are available.

Several IV estimations are presented, including an original two-stage procedure for grouped data that can be applied when the first stage can be estimated at the individual level but the second stage can only be estimated at the aggregate level. This procedure, inspired by the traditional two-stage least squares (2SLS) method, can easily be applied to other cases as well. Comparison of these results with efficient Wald estimates and standard 2SLS estimates confirms that the procedure is valid.

The results provide evidence that suggests there is a causal effect of education on mortality and that this effect is larger than the previous literature suggests. While GLS estimates suggest that an additional year of education lowers the probability of dying in the next 10 years by approximately 1.3 percentage points, my results from the IV estimation show that the effect is in fact much larger: at least 3.6 percentage points.

This paper is organized as follows. Section 2 describes the data used in this project, including a description of how the census is used to obtain mortality rates.

They use state-of-birth, income and education expenditures per capita from year-of-birth to age 6 in state-of-birth, and dummies for ancestry as instruments for education. They also estimate the effect of education on disability with NLS data, using IQ and family background measures as instruments. In both cases schooling is significant. Using a sample of older persons from the 1986 PSID, Leigh and Dhir (1997) use parental education, background, and state-of-residence at age 16 to instrument for education in regressions for disability and exercise. Alternatively, they include direct measures of time preferences and risk aversion. Education was not always significant. Finally Sander (1995) examines the effect of schooling on the odds of quitting smoking using the General Social Survey. He uses parental schooling as an instrument for schooling and finds that the effect of schooling is quite large for whites.

[13] Many development studies show that family background affects children's health. For a thorough review of these studies see Strauss and Thomas (1995). IQ measures (Berger and Leigh, 1988) suffer from the same problem.

[14] For examples see studies that looked at the consequences of the Dutch famine on the health of adults conceived during the famine, such as Hoek, Brown and Susser (1998) or Roseboom et al (2000).

[15] However this methodology is used in epidemiology. For example see the work by Haines and Preston (1996).

Section 3 shows that compulsory attendance and child labor laws had an impact on the educational attainment of individuals, and presents evidence that these laws are good instruments. Section 4 presents the general econometric framework used in the mortality analysis. The framework includes least squares as well as two IV estimations. The results are presented and discussed in Section 5, and conclusions are given in section 6.

2. Data

I use the U.S. censuses of 1960, 1970 and 1980, which are one percent random samples of the population.[16] The census provides information on age, sex, race, education, urban/rural residence, marital status, state of residence and state of birth. My samples include all white persons born in the 48 states,[17] that were 14 years of age between 1914 and 1939, with no missing values for completed years of education.[18]

I use the censuses to follow "synthetic cohorts." Although I do not observe the same individuals over time (so I cannot observe individual deaths), I do observe the same groups over time, which allows me to estimate group death rates. I aggregate the censuses into groups defined according to their gender/cohort and state-of-birth (descriptive statistics in Table 1). I follow 25 cohorts, born in 48 states. Using the 1960, 1970 and 1980 censuses, I can calculate two 10-year death rates for each group: one for 1960-1970, and another for 1970-1980. For example, the 1960-1970 death rate for a group is the number of people alive in 1960 (N_{60}) minus the number of people alive in 1970 (N_{70}) divided by the population in 1960 (N_{60}):

$$\frac{N_{60} - N_{70}}{N_{60}}$$

One issue that arises in estimating death rates by groups is measurement error. As Figure 3 shows, because of random sampling the number of deaths will be overestimated about half the time and underestimated half the time for all cohorts. As a result, some estimated death rates are negative. In the data, we observe more negative death rates for younger cohorts and fewer negative death rates for older cohorts (see Figure 4a); this is a pattern we should expect. As we can see in Figure 3B, with a zero death rate (no change in the population), two successive samplings of the same population result in a negative death rates half the time. When the death rate increases (as the population ages), the likelihood that the second sample will contain more observations than the first falls, resulting in fewer negative death rates. We also observe fewer negative death rates for states with large population (Figure 4b), which is also to be expected since the sampling error is smaller for larger populations.

[16] The data come from the IPUMS 1960 general sample, the 1970 Form 2 State sample (originally 15% state sample), and the 1980 1% Metro sample (originally B sample). These data sets were downloaded from the following web site: http://www.ipums.umn.edu

[17] Hawaii and Alaska were not then part of the Union.

[18] For consistency across censuses, I recoded completed years of education to be a maximum of 18 years instead of 20 in 1980.

The negative death rates are not a source of concern for two reasons. First, the estimated death rates will result in consistent estimates of the true death rates.[19] Second, average cohort death rates from the censuses are very similar to those obtained from individual data from the NHEFS described below (see Figure 4c).Note that the graph suggests there is evidence of age heaping: for ages that are multiples of 10, the death rates fall, because individuals tend to over-report their age and chose a multiple of ten when doing so.

I also used the National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study, 1992 (hereafter NHEFS). This survey followed 14,407 individuals who were between 25 and 74 years of age when interviewed for the first National Health and Nutrition Examination Survey (NHANES I) between 1971 and 1974. The NHEFS collected information on individuals in four subsequent waves (1982-84, 1986, 1987 and 1992). The sample is composed of whites[20] who were born in the 48 states between 1901 and 1925 and who were followed successfully, with no missing observations for years of completed education. The sample is further restricted to those who were alive in 1975 (N=4554). The NHEFS followed individuals and recorded whether they had died by 1985. Table 2 shows the summary statistics for this data.

The data on compulsory attendance and child labor laws come from a number of sources. There are eight years of state-level data (1915, 1918, 1921, 1924, 1929, 1930, 1935 and 1939) on these laws,[21] and some additional information for other years. I imputed missing observations by using the older values. The information was not recorded consistently by a single agency; in cases of conflicting pieces of data, I used the newer information to correct the data. (See the Appendix for tabulations and trends of the laws) I also collected data on state-level factors that contributed to the growth of secondary education from 1915 to 1939[22] or that could affect mortality. These include state expenditures on education, number of school buildings per acre, percent of the population that was living in urban areas, percent of the white population that was foreign born, percent of the population that was black, percent of the population employed in manufacturing, average annual wages in manufacturing per worker, average value of farm property per acre, and number of doctors per capita (See Lleras-Muney 2001 for information on data sources).

Each individual is matched to the laws and state characteristics that were in place in their state-of-birth when they were 14 years old. I choose this age because it is the lowest common drop-out age across states.[23] This procedure assumes that individuals

[19] Also note that IV estimates are only consistent, not unbiased, estimates of structural parameters. A consistent estimate of the dependent variable is sufficient for the IV estimators to be consistent.

[20] Other researchers have suggested that blacks had significantly different school experiences during the begining of the century. See Card and Krueger (1992). Also preliminary work on my part suggests that compulsory schooling laws and child labor laws did not affect blacks. The laws are never significant. It is unclear why. See Lleras-Muney (2001).

[21] Acemoglu and Angrist (1999) have gathered similar data. The data for this project was collected independently.

[22] The state-level variables were suggested by the work of Goldin (1994) and Goldin and Katz (1997).

[23] Schmidt (1996) tested this assumption and found that the effect of the laws was larger when matching at this age. Also, because grandfather clauses are common, it is reasonable to think that the laws in place at age 14 were the laws that would be binding for individuals even when they were

went to school in their state-of-birth. Inevitably some individuals were mismatched. However, Card and Krueger (1992) show that mobility was low during this period and that this assumption results in a small error, roughly 10 percent. Furthermore, if such an error exists, it likely will be uncorrelated with laws on compulsory attendance and with child labor laws, because these laws were probably not the reason why individuals moved across states. Indeed the data suggests these laws cannot explain mobility once we control for education.[24]

3. Did Compulsory Attendance and Child Labor Laws affect schooling? First Stage

The validity of the methodology proposed in this paper rests on the crucial assumption that compulsory attendance laws and child labor laws can be used as instruments. This section estimates the first stage, showing that the laws are good predictors of educational attainment both at the individual and aggregate level. These results will then be used in the two-stage (IV) estimations in Sections 4 and 5. I also provide additional evidence here that the laws are good instruments.

3.1. What do we know about Compulsory Attendance and Child Labor Laws?

Since their inception in Massachusetts in 1852, compulsory attendance laws have been complex. They specify a minimum and a maximum age between which school attendance is required; a minimum period of attendance; penalties for non-compliance; and the conditions under which individuals could be exempted from attending school, including achievement of a certain level of education (for example the completion of eighth grade), mental or physical disability, distance from school, and so on. The most common exemption was for work. Work permits were available even for young children, generally even younger than the minimum dropout age specified by compulsory education laws. Child labor laws, which extensively regulated the employment of minors, also included several conditions for the granting of such permits and for exemptions.

Child labor laws and compulsory attendance laws often were not coordinated. Each stipulated different requirements for leaving school. For example, in 1924 in Pennsylvania, the ages for compulsory attendance were 8 to 16, but the child labor laws allowed 14 year-olds to get work permits and leave school.[25] Continuation school

15 or 16 years old.

[24] I regressed mobility between state-of-birth and state-of-residence in1960 as a function of education, compulsory education laws and all other covariates used in this paper. The F statistic of joint significance of the laws has a value of 1.17 (p value of 0.3151), suggesting the laws cannot explain mobility. Also Lleras-Muney (2001) shows that restricting the sample to those that are still living in their state-of-birth yields estimates of the effect of the laws that are statistically identical to those presented here.

[25] Work permits evolved over time and today they can be obtained for part-time employment which does not involve dropping out of school. During this early period however, work permits effectively allowed children to leave school. See Woltz (1955).

laws, which forced children at work to continue their education on a part-time basis, were the only laws that attempted to bridge this gap. Compulsory attendance laws and child labor laws were in place in all states by 1918, and were modified frequently thereafter.

There is little agreement regarding the effectiveness of these laws. Landes and Solomon (1972) analyze the impact of the laws on attendance from 1880 to 1910. They find that compulsory education laws did not contribute to the increase in enrollments during this period. They further suggest that states with higher enrollments were more likely to pass more restrictive laws than other states. Eisenberg (1988) shows that attendance levels and expenditures per school-aged child were important factors in explaining the passage of the compulsory attendance laws from 1870 to 1915. Stigler (1950) and Edwards (1978) look at the impact of the laws on enrollments from 1940 to 1960 and conclude that they were not effective.

There are a number of studies that support a different conclusion. Margo and Finnegan (1996) use the 1900 census and find that the compulsory education laws did have an impact, but only when including measures of child labor laws as well. Schmidt (1996) finds large effects of compulsory education laws on the probability of high school completion between 1920 and 1934. Lang and Kropp (1986), using data from 1908 to 1970, show that compulsory education laws affect enrollments, even for groups not targeted by the laws. Angrist and Krueger (1991) test whether the laws affected enrollments in 1960, 1970 and 1980 with a difference-in-difference estimator (by state, quarter of birth and law) and find significant effects. Lastly, Acemoglu and Angrist (2000) find that the effects of the laws on educational attainment are positive and significant, but note that the effect of child labor laws is larger.[26]

Previous studies (including my own[27]) suggest that only three of the many aspects of these laws had an impact on individual educational attainment: the age at which a child had to enter school (*enter age*), the age at which the child could get a work permit and leave school (*work age*), and whether or not the state required children with work permits to attend school on a part-time basis (*contsch*). Following Acemoglu and Angrist (1999), I combine the age at which a child had to enter school and the age required for work permit into a single variable, *childcom*, defined as:

$$childcom = work\ age - enter\ age$$

This variable is the implicit number of years that a child had to attend school, given that the entering age and the work permit age were enforced. This variable takes the values of 0, 4, 5, 6, 7, 8, 9, or 10.[28] The other variable, *contsch*, takes the value of 1 if continuation school laws were in place. National trends and tabulations describing these laws throughout the period of study are shown in Appendix E.

The period from 1915-1939 is when compulsory education laws (hereafter I refer to both compulsory attendance laws and child labor laws as "compulsory education

[26] For a detailed review of these studies see Lleras-Muney (2001).

[27] See Lleras-Muney (2001), Angrist and Acemoglu (1999) and Schmidt (1996).

[28] Note that there are only a few cases when childcom was 0 (9 states, sometime before 1920). This occurs only at the beginning of the period if the state had no law that defined either entering age or work permit age.

laws") are more likely to have affected many individuals.[29] Secondary schooling was experiencing remarkable growth, especially in the first 40 years of this century. Goldin and Katz (1997) show that the percentage of young adults with high school degrees increased from 9 percent in 1910 to more than 50 percent in 1940. Also, it has been suggested by other social sciences that, in the previous period (up to 1915), these laws were perceived as ineffective; most studies seem to confirm that view.[30] But social scientists agree that the laws were enforced by the 1920s[31] and Schmidt's work (1995)—the only study to concentrate on this period—confirms it. Also, note that Goldin and Katz (1997) show that high school graduation rates were unusually low in the Second World War years due to the high wages that inexperienced workers could command. Stigler (1950), Edwards (1978), and partially Angrist and Krueger (1991)[32] suggest that the laws declined in importance after the 1940s. So the first part of the 20th century provides the perfect window of opportunity for using the laws as instruments. Finally, from a technical point of view, this period is interesting because states were constantly changing their compulsory education and child labor laws, and there is a sufficient amount of variation over time.

3.2. The effect of Compulsory Attendance and Child Labor Laws on educational attainment

As preliminary evidence of the effect of these laws on education, I graph the average education by *childcom* for the entire sample (Figure 5) and by cohort, for every 5th cohort in the data (Figure 6). Both graphs show that average education is higher for those in states where more education was compulsory. In order to add further controls, I turn to regression analysis.

Pooling individual data from the 1960 and 1970 census, I estimate the following model:

$$E_{\mathsf{ics}} = b + CL_{\mathsf{cs}}\pi + X_{\mathsf{ics}}\beta + W_{\mathsf{cs}}\delta + \gamma_{\mathsf{c}} + \alpha_{\mathsf{s}} + \varepsilon_{\mathsf{ics}}$$

The dependent variable is years of completed education for individual i of cohort c born in state s. CL is a set of dummies for compulsory education laws in place in state s when the individual was 14, X_{ics} are individual characteristics such as gender and place of current residence, W_{cs} is a set of characteristics of individual i's state-of-birth at age 14 (such as manufacturing wages, expenditures in education, per capita doctors, etc.), γ_{c} are cohort dummies, α_{s} are state-of-birth dummies. The regression also includes interactions between region-of-birth and cohort, an intercept (b) and a

[29] Schmidt (1996) confirms this intuition.

[30] Katz (1976) and Ensign (1921) suggest that in the 19$^{\mathsf{th}}$ century and early 20$^{\mathsf{th}}$ century laws were created but not enforced. Many state laws did not even provide enforcement mechanisms, and if they did, there were often insufficient means to enforce them, especially in rural areas.

[31] Truant officers had become commonplace. They were in charge of making sure that students of age were effectively in school, and they could penalize the parents in cases of non-compliance. Also, expenditures on education increased. Although it would be hard to argue that this increase was solely the result of the passage of compulsory education laws, it is certainly true that the laws required that the states provide public schools and pay for enforcement agencies. See Tyack (1974), Katz(1976).

[32] They find that the impact of the laws is about 4% in 1960, 2% in 1970 and 0.5 in 1980.

dummy for 1970. I also estimate the model by aggregating the data at the state-of-birth/cohort and gender level.[33] Both estimations will be used in Section 4 (first stage).

Table 3 shows the results. The first column estimates the relationship including only state effects, cohort effects, a female dummy, and a set of dummies for the laws. The coefficients are fairly robust to the addition of other controls (see column 2).[34] The last column shows the results from estimating the equation using the data aggregated at the state-of-birth, cohort and gender level. The estimations show that the laws increased the educational attainment of individuals. As expected, all dummies for the laws are positive and significant and they generally increase as the number of compulsory years increases. Overall, the implied increase in educational attainment due to *childcom* is around 4.8 percent.[35] The effect is identical if the sample is restricted to include only those observations for which *childcom* is not 0.[36] This estimate is similar to those reported by Acemoglu and Angrist (2000), who report an increase between 1 and 6 percentage points; by Eisenberg (1988), who finds an effect of about 2 percent; and by Angrist and Krueger (1991), who find that the impact of the laws was about 4 percent in 1960.[37] Also, the continuation school dummy is positive.[38]

Before turning to the effect of education on mortality, I present evidence that the laws are good instruments. At the bottom the Table 3, I report the F-test of joint significance of the laws; it shows that the laws are always jointly significant at the 5% level for both specifications. Additionally the F-statistic is greater than (or very close to) 5, which suggests that the instruments are strong. I also report the partial R-squared coefficient, another measure of the instruments' strength.[39] It has a value of 0.0001 or higher, which compares favorably to those reported by Bound, Jaeger, and Baker (1995).

It is also worth pointing out that the changes in the laws that took place during

[33] The model estimated would be:

$$\overline{E_{gcs}} = b + CL_{cs}\pi + \overline{X_{gcs}}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \overline{\varepsilon_{gcs}}$$

where $\overline{E_{gcs}}$ is the average education in a given state, cohort and gender; and $\overline{X_{gcs}}$ are the average characteristics of that group. The number of individuals in each group are used as weights.

[34] Inclusion of other variables, such as income, immigrant status of parents, and so on, has no impact on them. Also, identical regressions by region-of-birth or by gender yield similar results. See Lleras-Muney (2001) for these results.

[35] This was calculated by replacing the set of dummies by the continuous variable childcom. See Lleras-Muney (2001).

[36] See Lleras-Muney (2001).

[37] Schmidt (1996) finds much larger effects, about 20% for her analysis of New York State.

[38] Continuation school is not significant in this sample, but previous work (see Lleras-Muney, 2001) showed that this law affected white males and individuals born in the north and south of the U.S. Therefore I include it.

[39] There is a large literature on the problem of weak instruments. Bound, Jaeger, and Baker (1995) suggested that the researcher evaluate the quality of the instruments by looking at two statistics. First, the F statistic on the excluded instruments in the first stage should be statistically significant and large. Staiger and Stock (1997) further suggests that a value of less than 5 could signal weak instruments (this is a rule of thumb). Second, the partial r-square (obtained by regressing schooling on the instruments, once the common variables have been partialled out) should be high. Following their suggestion, these two statistics are reported here.

this period appear to have been exogenous to individuals. Although different states might have had different tastes for education, the regressions here include a very large set of controls (cohort dummies, state-of-birth dummies and region-of-birth*cohort interactions are included) which should capture this effect. Also note that the addition of controls (compare column 1 and 2 of Table 3) has little effect on the coefficients of the laws, suggesting that any excluded state-of-birth/cohort level variables are not correlated with the laws. Furthermore, Lleras-Muney (2000) presents evidence consistent with exogenous laws: her results suggest that the laws impacted only the lower end of the distribution of education. She rejects the hypothesis that changes in the laws during this period resulted from (rather than caused) increases in education, using a test inspired by Landes and Solomon (1972).[40]

A final concern is that the laws must affect individual health only through their effect on education. There is no evidence that the laws included any clauses or restrictions that would have affected health independently. For example, there were no lunch programs provided as part of school attendance. Also the states that led in education during this period (the prairie states[41]) were not the same states that led in health (northeastern states).[42] But again, the controls included here are meant to rule out this possibility. Finally, exogeneity tests are performed in the IV estimation (see next section).

Overall the results show that the laws did have an impact on educational achievement, and that their predictive power is large. The important implication is that compulsory education laws can be used as instruments. Therefore I turn now to the question of the effect of education on mortality.

4. Health and Education: Econometric model

4.1. Least Squares Estimation

The econometric model for the relationship between education and health can be written as a linear system of simultaneous equations:

$$H_i = X_{1i}\beta_1 + E_i\pi_1 + \varepsilon_{1i} \tag{4.1}$$

$$E_i = X_{2i}\beta_2 + H_i\pi_2 + \varepsilon_{2i} \tag{4.2}$$

H_i is individual i's health stock, E is his education level, X_1 is a vector of individual characteristics that affect health, such as smoking, and genetic factors. X_2 is a vector of individual characteristics that determine education, such as ability. X_1

[40] The test consists of matching individuals to the laws in place in their state-of-birth when they were 17, 18 and up to 26 years of age, when these laws should no longer have affected them. Lleras-Muney (2000) finds that future laws cannot explain educational attainment, whereas laws at age 14 can.

[41] See Goldin and Katz (1997).

[42] Starr's 1982 book provides anecdotal evidence that the northern states lead in a variety of health aspects. My own data supports this conclusion. For example, the north had the highest number of doctors per capita throughout the period. And the number of doctors per capita in the north did not decline from 1915 to 1939 but did decline in the rest of the country. The north also had the highest declines in infant mortality rates during this period. (Results available upon request.)

and X_2 may contain common factors. This general specification allows for causality to run from education to health and vice-versa.

The purpose of this paper is to determine only whether or not education affects health (i.e. $\pi_1 = 0$?). Therefore I only estimate the health equation (equation 3.1). Although health is unobserved, mortality is observable. Following Grossman's (1972) model of health, death occurs when the stock of health falls bellow a certain threshold. In a less deterministic model, H_i is proportional to the underlying probability (index) of being alive, and death is the observed result. This is the usual limited-dependent-variable set-up.

This mortality equation can be estimated at the individual level using the NHEFS but not the census. If individuals could be followed from the 1960 census to the 1970 census (or from 1970 to 1980), then (based on the previous discussion) the following individual linear probability model could be estimated:

$$D_{t,ics} = b + E_{ics}\pi + (X_{t-1})_{ics}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \varepsilon_{ics} \tag{4.3}$$

where D_{ti} is equal to one if the individual is deceased at time t. E_{ics} is i's education (measured by completed years of education), X_{t-1} are other individual characteristics measured as of (t-1) (including gender), W_{cs} is a set of characteristics of individual i's state-of-birth at age 14, γ_c is a set of cohort dummies, α_s is a set of state-of-birth dummies, b is an intercept, and ε is the error term.

Using the census, individuals cannot be tracked over time, but I can track groups that are constant over time, and calculate their death rates by aggregating the data. I aggregate by gender, cohort, and state-of-birth. This aggregation level uses *all* of the available individual characteristics that are time invariant (except for education), and therefore it maximizes the number of observations in the aggregate data.

The aggregate model is derived from the individual model by averaging over individuals in a given gender/cohort/state-of-birth group as follows:

$$\overline{D}_{tgcs} = b + \overline{E}_{gcs}\pi + \overline{(X_{t-1})}_{gcs}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \overline{\varepsilon}_{gcs} \tag{4.4}$$

where $\overline{D}_{tgcs}$ represents the proportion of individuals that died in a given group or the death rate for that group, and $\overline{X_{t-1}}_{gcs}$ represents the average characteristics of that group at (*t-1*) (for example, the the percentage of people in that group living in urban areas).[43]

Note that I use a linear probability model for the estimation. The existence of negative death rates makes it impossible to use a non-linear model such as a Logit or Normit. However, since the dependent variable (the death rate) is not censored below by 0, the linear probability assumption is less problematic in this case than in general.[44]

In the linear probability model, the error term is heteroskedastic and has the following variance:

$$var(\varepsilon) = \frac{\overline{D}_{tgcs}(1 - \overline{D}_{tgcs})}{n_{gcs}} \tag{4.5}$$

[43] Including a dummy for gender.

[44] Furthermore, in the next section, I will test this assumption by comparing results from the census to those obtained from individual data.

where n_{cse} is the number of individuals in that group and $\overline{D}_{tgcs}$ is the observed probability (death rate) for that group. A standard estimation procedure (the minimum chi-square method[45]) in this case is to run weighted least squares, where the weights are given by $1/\sqrt{var(\varepsilon)}$. Again, due to random sampling and the error it generates, the observed probabilities can be negative, so this estimation is not possible.[46] In order to address the heteroskedasticity problem I estimate the equation by GLS (weighted least squares) using the number of individuals in the group as weights. To correct for further heteroskedasticity, I use White's sandwich estimator.[47]

It is intuitive that GLS estimates at the aggregate level will be biased and inconsistent, since the correlation between the error term and education (at the individual level) will carry over when calculating group means.[48] Now I turn to the Instrumental Variables (IV) estimation which will yield consistent estimates of the causal effect of education on mortality.

4.2. Efficient Wald estimates

One obvious solution to correct for the bias in the GLS coefficient is to use Instrumental Variables (IV). Given that many instruments are available, Two Stage Least Squares (2SLS) would be the preferred estimation method. At the individual level, the 2SLS model is:

$$\begin{aligned} D_{ti} &= b + E_{ics}\pi + (X_{t-1})_{ics}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \varepsilon_i \\ E_{ics} &= b + CL_{cs}\pi + (X_{t-1})_{ics}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \varepsilon_{ics} \end{aligned}$$

where D is equal to one if the individual is deceased at time t. E is i's education (measured by completed years of education), X_{t-1} are other individual characteristics measured as of (t-1) (including gender), W_{cs} is a set of characteristics of individual i's state-of-birth at age 14, γ_c is a set of cohort dummies, α_s is a set of state-of-birth dummies, b is an intercept, and ε is the error term, which is assumed to be normal $N(0,\sigma_1^2 I)$. CL is the set of compulsory education laws that serve as instruments to identify the education equation. This model can be estimated using the individual NHANES data but not with the census.

Since the census data can be used only as grouped data, the Wald estimator is an alternative estimator for the effect of education. Angrist (1991) showed that the Wald estimator for grouped data is efficient and in fact equivalent to 2SLS using individual level data. In the case of many explanatory variables the efficient Wald estimator is found by GLS estimation of the following equation:

$$\overline{D}_{crl} = \overline{E}_{crl}\pi + \gamma_c + \delta_r + \overline{\varepsilon}_{crl} \tag{4.6}$$

[45] See Maddala p. 29, Green p. 895.

[46] An idea is to convert the negative weights into 0. Notice though this procedure will introduce bias in the results: I observe both underestimated and overestimated death rates, but by converting the negative (underestimated) ones into 0, I am "fixing" the problem only for half of the observations. Alternatively, one can run weighted least squares only using the observations for which the observed death rate is positive, but again this will introduce the same kind of bias.

[47] In all the estimations, including the IV estimations, where state-of-birth characteristics are included, the standard errors are also clustered at the state-of-birth and cohort levels.

[48] Proof available upon request.

where $\overline{D}_{\mathsf{crl}}$ is the death rate for individuals born in cohort c in region r under compulsory law l, and $\overline{E}_{\mathsf{csl}}$ is the average education of individuals born in cohort c in region r under compulsory law l. The weights are given by the population in each group. In other words, Wald is estimated by grouping the data by gender/cohort/region-of-birth and compulsory education law. This procedure is equivalent to 2SLS at the individual level, where cohort dummies γ_{c} and region dummies δ_{r} serve as their own instruments (since they are exogenous), and compulsory education laws serve as instruments for education, the endogenous variable. The estimates are referred to as the efficient Wald estimates.

Note that because compulsory education laws are defined at the state-of-birth and cohort level, I cannot control for both state-of-birth and cohort when using this estimator. This is a drawback of the Wald estimator, especially if one thinks that state-of-birth and the laws are correlated. In order to alleviate this problem, I control instead for region-of-birth. But region-of-birth may not be a good proxy for state-of-birth. Furthermore, other individual ($X_{\mathsf{t-1}}$) and state-of-birth characteristics (W_{cs}) cannot be included in this specification.

4.3. Two-Stage Least Squares with Aggregate Data

Alternatively, I can estimate the 2SLS model at the data that has been aggregated at the state-of-birth/cohort and gender level. Estimation at the aggregate level results in less efficient estimates (see Green pp. 433-434) but all the covariates (especially state-of-birth) can be included. Using the aggregate data 2SLS is obtained by estimating the following model:

$$\begin{aligned}
\overline{D}_{\mathsf{gcs}} &= b + \overline{E}_{\mathsf{gcs}}\pi + \overline{X}_{\mathsf{gcs}}\beta + W_{\mathsf{cs}}\delta + \gamma_{\mathsf{c}} + \alpha_{\mathsf{s}} + \overline{\varepsilon}_{\mathsf{gcs}} \\
\overline{E}_{\mathsf{gcs}} &= b_1 + CL_{\mathsf{cs}}\pi + \overline{X}_{\mathsf{gcs}}\beta + W_{\mathsf{cs}}\delta + \gamma_{\mathsf{c}} + \alpha_{\mathsf{s}} + \overline{\varepsilon}_{\mathsf{gcs}}
\end{aligned}$$

where now $\overline{D_{\mathsf{tgcs}}}$ is the proportion of individuals who died in a given gender/cohort and state-of-birth, $\overline{E}_{\mathsf{gcs}}$ is the average education of that group and $\overline{(X_{\mathsf{t-1}})}_{\mathsf{gcs}}$ are other average characteristics. Again, the weights are given by the number of observations in each cell, and the excluded instruments from the mortality equation are the compulsory education dummies, CL_{cs}. The first stage (estimation of $\overline{E}_{\mathsf{gcs}}$) was shown in the previous section.

4.4. Mixed Two-Stage Least Squares Estimation

The census allows me to estimate the first stage using individual data. The intuition behind Mixed-2SLS is that it might be possible to take advantage of this fact and gain efficiency (relative to the previous 2SLS) by estimating the first stage at the individual level (as done in the previous section) and then aggregating the data by gender/cohort/state-of-birth. (See Dhrymes and Lleras-Muney, 2001.)[49]

[49]Dhrymes and Lleras-Muney (2001) compare 2SLS and Mixed 2SLS estimators. The question of which estimator has lower variance turns out to be data dependent, but it is possible for Mixed 2SLS to be more efficient.

Mixed 2SLS is obtained by estimating the following equation through weighted least squares:

$$\overline{D}_{gcs} = b + \overline{\widehat{E}}_{gcs}\pi + \overline{(X)}_{gcs}\beta + W_{cs}\delta + \gamma_c + \alpha_s + \overline{\varepsilon}_{gcs}$$

Again $\overline{D}_{gcs}$ represents the proportion of individuals who died in a given gender/cohort and state-of-birth group, $\overline{X}_{gcs}$ represents the average characteristics of the group, but now I include $\overline{\widehat{E}}_{gcs}$, the average predicted education for that group from the first stage.[50] The weights are given by the number of observations in each cell. The excluded instruments from the mortality equation are the compulsory education dummies. The only difference between standard 2SLS with aggregated data and Mixed-2SLS is in the predicted education term. 2SLS uses predicted average education whereas Mixed-2SLS uses average predicted education.

More formally, define H as the matrix that transforms the data into group means and weights each group mean by the number of individuals in the group, and let $\widehat{X}$ contain all the same variables as in the GLS estimation, but with education replaced by the predicted level of education from the first stage regression ($\widehat{X} = \left[\overline{\widehat{E}} \mid \overline{X_{t-1}} \mid \gamma_c \mid \alpha_s\right]$). Then the estimator β_{Mixed} can be expressed as:

$$\beta_{Mixed} = \left(\widehat{X}'H'H\widehat{X}\right)^{-1}\widehat{X}'H'HD$$

This procedure also results in consistent estimates. As usual the variance-covariance matrix needs to be corrected.

5. Results

5.1. Least Squares Results

Although we have good reason to believe that GLS produces biased estimates, I report them here as the benchmark for comparison with the IV results. Using the census, I estimate the GLS model described above. The results are in the first column of Table 4. The estimated coefficient of the effect of education on the death rate is about -0.012. The coefficient is highly significant and is is robust to the inclusion of more controls.[51] It is a well known fact that there exist persistent differences in mortality rates by gender. I therefore repeat the analysis by gender (Table 5). The effect of education is positive and significant, but this is probably due to the small sample size.

The validity of the aggregation procedure rests on the assumption that the aggregate data can be understood as coming from unobserved individual data. It is important that this intuition be confirmed, so I compare aggregate results from the census with those obtained with the NHEFS individual data. This comparison allows me to check the validity of the linear probability assumption and helps me to interpret the aggregate results.

[50] The expression for the first stage was given in the previous section.
[51] Results available upon request.

Using individual NHEFS data, I estimate a linear probability model and a probit model, where the dependent variable is a dummy indicating whether or not the person died between 1975 and 1985. Then I aggregate the NHEFS data by gender, state-of-birth, and cohort and again estimate the same linear model estimated with the census data. Because of the small number of observations in the NHEFS aggregating by gender, state-of-birth and cohort results in very few observations per cell, so I also reproduce the results only aggregating by state-of-birth and cohort. The results are shown in Table 4.

Comparing the results from LS regressions from the census with results from the NHEFS shows that the census data gives extremely accurate estimates of the effect of education. The census LS estimates are very similar to those obtained using the NHEFS aggregated data, which in turn are similar to those obtained at the individual level, using either LS or probit estimations.

These results suggest that sampling (and the measurement error it generates) does not significantly affect the estimates for education, that there is no aggregation bias and that the linear model is a good approximation of the education-death rate relationship. The comparison is also useful in terms of interpretation: a -0.012 coefficient for education means that increasing the education of a given cell by one year lowers its death rate by 1.2 percentage points. This coefficient also implies that increasing an individual's education by one year will lower his probability of dying between 1960 and 1970 (or between 1970 and 1980) by 1.2 percentage points. This latter interpretation is more intuitive and useful. Note again that the OLS effect is quite large: at the mean, this result implies that a 10 percent increase in education lowers mortality by about 11 percent, therefore an elasticity of about -1.

5.2. IV results

The first column in Table 6 presents the 2SLS results using the NHEFS. This estimation is done at the individual level and using standard 2SLS. The estimate is positive (the effect of education is about -0.02) but not significant: because this sample is small, the first stage estimation[52] is poor. Nonetheless, although the standard errors are high, the estimates from this sample are also larger than the GLS estimates obtained from the same data.

The second column shows the results from the Wald estimation. The Wald estimate of the effect of education is about -0.037 and significant at the 5 percent level. The third column presents the results of 2SLS estimation using aggregated data at the gender/state-of-birth-and cohort level. The effect of education is about -0.045 and significant at the 10 percent level. The Mixed 2SLS results (last column) show that the coefficient on education is approximately -0.059 and is significant at the 5 percent level. All of the previous estimates are significant at the 5% level using a one-tailed test (the null hypothesis is that education is negative) which is perhaps more appropriate in this set-up.[53] Overall the results suggest that increasing education by

[52] Results not shown here but available upon request. In the first stage estimation using the NHEFS, only two of the dummies for compulsory education laws were significant at the 10% level, and the set of dummies was jointly insignificant.

[53] I thank Michael Grossman for this insight.

one additional year lowers the 10-year death rate by at least 3.6 percentage points.[54]

For the last two estimators I perform a test of overidentifying restrictions. The χ^2 statistic for the aggregate 2SLS model is 2.42 and 1.49 for the Mixed 2SLS model. This statistic tests the hypothesis that the model is well specified. It is calculated as the sample size times the R^2 from a regression of the residuals from the second stage on all exogenous variables, including the instruments. In both cases the overidentifying restrictions are not rejected at a 5 percent level (critical value 14.06). This test in conjunction with earlier results from the first stage suggests that compulsory education laws are legitimate instruments.

As a last attempt to address the potential endogeneity of the laws, I repeat the estimations above using a larger set of instruments that include quarter of birth, compulsory attendance and child labor laws, and the interactions of quarter of birth and the laws. Presumably these individual-level instruments will increase the efficiency of the estimates, and they are perhaps less likely to be endogenous.[55] The results (Table 7A) are identical to those presented above.

In table 7B I present the reduced form estimates, i.e. the direct effect of the laws on mortality. The results are consistent with previous estimations: if the effect of childcom on education is about 5 percent, and the effect of education on mortality is about 6 percent, then the direct effect of the laws on mortality should be about 0.3 percent, which is approximately what the reduced form result shows.

In table 7C I present the results excluding ages 40, 50 and 60 since the data showed evidence of age heaping. This is a potential problem if age heaping is correlated with education. The IV results are very similar to the previous results.

The results by gender from the Census are presented in Table 8. The coefficient on education is somewhat smaller for females than for males, confirming the findings in the literature that the effect of education is larger for males. These results are interesting for other reasons. First, they suggest that World Wars I and II did not results in significant selection bias for men. Also, in these estimations the effect of marriage is negative as the literature suggest, whereas the effect is positive in the joint estimations. This is a composition effect: males are both more likely die and to be married.

This section has presented four different estimates of the effect of education on mortality. Three different estimators, using two different data sets and three different levels of aggregation, were used. Although each estimate has weaknesses, all estimates point to the same conclusion: the effect of education is causal and in fact larger than OLS suggests. Given this variety of estimates, this result is very robust.

[54] Estimates by region are comparable in size to those presented here except that they are generally not significant.

[55] Note however that the use of these instruments might be questionable (see papers in footnote 10). Also Lleras-Muney (2001) shows for example that the laws affected whites but not blacks. However quarter of birth does appear to affect blacks' educational attainment. This again raises the issue of whether quarter of birth has an independent effect on education unrelated to compulsory attendance laws.

5.3. Discussion

The results are surprising for two reasons. The first is that the IV estimates are larger than the LS estimates. The second is that the effect of education is quite large. In this section I discuss these two issues.

In all the IV estimations presented here, the effect of education is much larger than the LS estimates suggest. The Mixed 2SLS estimates suggest the effect is as large as -0.058, whereas Wald estimates imply a coefficient of about -0.036. All IV estimates are larger than LS estimates. At first, this could seem to be a surprising result: the a priori expectation was that LS estimates would be too large. However, in the vast literature devoted to the earnings returns to education, researchers have come to similar conclusions: OLS estimates of the effect of education on earnings are too small.[56]

One explanation is that the omitted variable bias is smaller than the bias that results from measurement error in education.[57] The health literature has not been concerned with this potential problem although there is evidence of measurement error in education (Card 1995). If the measurement error is random, then IV estimate will be larger than the OLS estimate. [58]

Another explanation is the choice of instrument. Card (2000) suggests that one possible reason why IV estimates of the return to education are generally larger than OLS is that most instruments are based on policy interventions that affect the education choices of individuals with low levels of education. Under the assumption that different individuals face different returns to education due to unobserved characteristics, IV estimates reflect the marginal rate of return of the group affected by the policies (Imbens and Angrist, 1994; Angrist, Imbens and Rubin, 1996). If individuals choose low levels of education because they face high costs (rather than low returns) then the 2SLS estimates accurately measure those higher returns. In the context of the health returns to education, the results suggest that the individuals affected by the laws also face higher health returns to education than the rest of the population. This would not be surprising if one believes that the health returns to education are larger at lower levels of education.

To find suggestive evidence to support this claim, I re-estimate the model including a quadratic term for education. If the health returns to education are decrease as education increases then this quadratic term should be positive (since the effect of education on mortality is negative). Table 9 shows the results. The quadratic term is postive, suggesting that indeed the health returns to education are larger for lower levels of education. Also note that I cannot reject the hypothesis that education is exogenous using a Hausman test, which also supports the hypothesis that IV is measuring the effect for the bottom half, and that effect is causal.

Larger IV returns can also be explained if there exist health externalities from education. Lleras-Muney (2001) shows that compulsory education and child labor

[56] For a survey of these studies, see Card (1995).

[57] As Card (1995) mentions, the idea that measurement error bias could be just as serious as the omitted variables bias in the returns to education was first noted by Griliches (1977).

[58] Note however that if measurement error is not classical (non random) then IV estimates can also be biased (see Hyslop and Imbens, 2000 and Kane, Rouse and Staiger, 1999)

laws affected those at the lower end of the distribution of education decreasing inequality in education. There is a large literature that indeed suggests that inequality affects health.[59] Again, to find suggestive evidence for this hypothesis, I estimate the OLS model only for those with more than 12 years of schooling,[60] including now the standard deviation of the distribution of education in their state-of-birth and cohort (Table 10). The effect of the standard deviation of education is positive and significant, so that higher inequality results in higher death rates (lower inequality results in lower death rates).

The second issue is that the effect of education is quite large, and it is important to understand why. Education provides individuals with critical thinking skills, which in turn might affect understanding of health risks (Grossman's hypothesis). If this is the case, then the interaction of education with a variety of factors is relevant. For example, although access to information alone cannot explain health differences across education groups (Kenkel, 1991), information available to the more educated will result in greater benefits for them if they can understand it better, or can understand its relevance.[61] Another possibility is that the more educated might be more likely to adopt and implement new medical technologies.[62] Since both the availability of information and the rate of medical innovation dramatically increased in the last century, it is reasonable to think that the more educated were able to capture very high returns during this period. At the same time, this would explain why individuals did not voluntarily acquire education, in spite of the large returns: the increases in medical technology and information were not foreseeable at the time they made their education choices. Other direct mechanisms are documented in the cognitive psychology literature: lack of education is correlated with stress, depression and hostility, all of which have been shown to adversely affect health (Adler et al, 1994).

There are a few other indirect mechanisms through which education might affect health which are also consistent with the results in this paper. One obvious one is that being in the classroom is less of a health risk than working, especially while growing up. Also note that education gives you access to a higher income and different types of jobs, both of which affect health. For example, only high school graduates in the first half of the century had access to white collar jobs, which provided healthier work environments than manufacturing or agriculture.[63] Controlling for income (or occupation) does not change the results in this paper. But, since income is endogenous, it is not possible (given that I have no instruments for income) to distinguish the direct effect of education on health from its indirect effect through income. However, a few pieces of evidence suggest that income alone might not be the sole mechanism.

[59] Deaton and Paxson (1999) review the existing literature in their paper.

[60] I restrict the sample to differentiate the impact of inequality from that of own education. Note that the sample restriction however is not necessarily appropriate.

[61] As a consequence they might seek care earlier, get more medical care, get more preventive care, be more willing to use newly developed medical procedures/medicines, and so on.

[62] This idea was first postulated by Nelson and Phelps (1966). Bartel and Lichtemberg (1987) provide evidence at the plant level that highly educated workers have a comparative advantage with respect to the adjustment to and implementation of new technologies.

[63] Different/better jobs might provide access to health insurance. Note however that health insurance has not been proven to impact health (see introduction).

Grossman (1975) showed that the effects of income on health disappear once a certain level of income has been reached, while the same is not true for education. Standard results suggest that the returns to education are about 10% and that the elasticity of mortality with respect to income is about -0.3.[64] If the sole effect of education is through income, one more year of education should decrease mortality by 0.0033 (for average mortality of 0.11), which is a much smaller effect than what was estimated here.

Finally, the results in this paper do not imply that time preferences do not affect health and education choices nor that there is no reverse causality from health to education. They simply show that there is a causal effect of education on health, and that this effect is not due to time preferences. However, as Becker and Mulligan (1997) argue, education could lower the discount rate, making people more patient.[65] This is yet another indirect mechanism that could explain my results.

6. Conclusion

This paper has shown that there is a large causal effect of education on mortality. In fact, this effect is underestimated by OLS. Instrumental variables estimates show that one more year of education decreases the probability of dying within 10 years by at least 3.6 percentage points. To better understand the impact of education, using the coefficient from the Wald estimation, I calculate how this effect translates into life expectancy gains. I find that in 1960, one more year of education increased life expectancy at age 35 by at least 1.2 years. This is a very large increase.

A few notes of caution on how to interpret these results for public policy purposes are necessary. First, in order to make policy recommendations, we need to know more about the specific mechanisms by which education affects health. This paper analyzes the effects of increasing education from relatively low initial levels. It is unclear what the effects would be at higher initial levels of education. The average education level for white Americans born in 1901 was at most 8.87 years.[66] Today many developing countries, including most Latin American countries,[67] have average levels of education that are similar. This paper implies that more aggressive education policies could dramatically increase adult longevity in such countries. But cost benefit analysis of such policies are extremely complex, since for example we do not know what the cost of increasing education would be, or its effectiveness. Questions such as these are beyond the scope of this paper. But the results presented here suggests that the benefits of education are large enough that we need to consider education policies more seriously as a means to increase health, especially in light of the fact that other factors, such as expenditures on health, have not been proven to be very effective. Finally the results also suggest that the measured effect of technology on

[64] Deaton and Paxson (1999).

[65] This point is cited by Grossman (1999)

[66] This is the average education level of that cohort in 1960. Data for the entire population in 1901 does not exist for the US, but the average was probably much lower.

[67] Average education level of 25 year-olds in many Latin American countries was between 6 and 9: Bolivia, 8; Chile, 8.79; Ecuador, 7.12, Mexico, 6.23; Panama, 8.68; Peru, 7.2; Uruguay, 8.02 and Venezuela, 7.15 (Source: IDB, 1998).

health might actually reflect the effect of increased education rather than the effect of technology.

This evidence that education increases life expectancy implies that the returns to education, measured only in terms of earnings increases, substantially underestimate the true returns to education. In view of the large magnitude of the effect of education on health, it is clear that more attention needs to be devoted to the pathways of influence. Existing models of the relationship between education and health are very imprecise about the mechanisms through which education operates on health. It is crucial that we understand these mechanisms better, so that we can implement effective programs to increase the health of our population.

References

[1] Acemoglu, Daron and Joshua Angrist, "How Large are the Social Returns to Education? Evidence from Compulsory schooling Laws," NBER Working Paper No. W7444, December 1999

[2] Adler, Nancy E. et al, "Socioeconomic Status and Health, the Challenge of the Gradient," American Psychologist, vol 49, No. 1January 1994

[3] Angrist, Joshua D., "Grouped Data Estimation and Testing in Simple Labor Supply Models," Journal of Econometrics, February-March 1991

[4] Angrist, Joshua D. and Alan B. Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," Quarterly Journal of Economics, November 1991

[5] Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," Journal of the American Statistical Association, June 1996

[6] Auster, Richard, Irving Leveson and Deborah Sarachek, "The Production of Health, An Exploratory Study," Journal of Human Resources 4, 1969

[7] Bartel, Ann P. and Frank R. Lichtenberg, "The Comparative Advantage of Educated Workers in Implementing New Technology," Review of Economics and Statistics, February 1987

[8] Becker, Gary S. and Casey B. Mulligan, "The Endogenous Determination of Time Preference," Quarterly Journal of Economics, August 1997

[9] Berger, Mark C. and J. Paul Leigh "Schooling, Self Selection and Health," Journal of Human Resources 24, 1989

[10] Behrman, Jere R., Robin C. Sickles, and Paul Taubman , *Causes, Correlates and Consequences of Death among Older Adults: Some Methodological Approaches and Substantive Analysis*, 1998

[11] Bound, John, David A. Jaeger and Regina Baker, "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variables is Weak," Journal of the American Statistical Association, 90, June 1995

[12] Bound, John, and David A. Jaeger, "On the Validity of Season of Birth as an Instrument in Wage Equations: A Comment on Angrist and Krueger's "Does Compulsory School Attendance Affect Schooling and Earnings"," NBER Working Paper 5835, November 1996

[13] Card, David, "Estimating the Returns to Schooling: Progress on some Persistent Econometric Problems," NBER Working Paper no.W7769, June 2000

[14] Card, David, "Earnings, Schooling, and Ability Revisited," Research in Labor Economics, vol. 14, 1995

[15] Card, David and Alan Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," Journal of Political Economy 100, January 1992

[16] Christenson, Bruce A., and Nan E. Johnson, "Educational Inequality in Adult Mortality: An Assessment with Death Certificate Data from Michigan," Demography 32. May 1995

[17] Currie, Janet and Rosemary Hyson, "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birth weight,"American Economic Review, May 1999

[18] Deaton, Angus and Christina Paxson, "Mortality, Education, Income and Inequality among American Cohorts," NBER Working Paper 7140, May 1999

[19] Dhrymes, Phoebus and Adriana Lleras-Muney (2000), "Estimation of Models with Group Data by Means of '2SLS'," mimeo, Columbia University, 2001

[20] Duleep, H.O., "Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data," Journal of Human Resources 21, 1986

[21] Edwards, Linda N., "An Empirical Analysis of Compulsory Schooling Legislation, 1940-1960," Journal of Law and Economics, April 1978

[22] Eisenberg, M.J. (1988) "Compulsory Attendance legislation in America, 1870 to 1915," Ph.D. Dissertation, University of Pennsylvania

[23] Elo, Irma T. and Samuel H. Preston, "Educational Differentials in Mortality: United States, 1979-85," Social Science and Medicine 42(1), 1996

[24] Ensign, Forest Chester, "Compulsory School Attendance and Child Labor," Iowa City, IA: The Athens press, 1921

[25] Farrell, P. and Victor R. Fuchs, "Schooling and Health: The Cigarette Connection," Journal of Health Economics, 1982

[26] Feldman J. D. Makuc, J. Kleinman and J. Cornoni-Huntley, "National Trends in Educational Differences in Mortality," American Journal of Epidemiology, 1989

[27] Filmer, Deon and Lant Prichett, "Child Mortality and Public Spending on Health: How Much does Money Matter?," World Bank Policy Research Working Papers 1864, December 1997

[28] Fuchs, Victor R., "Time Preference and Health: An Exploratory Study,: in Victor Fuchs, Ed., *Economic Aspects of Health*, Chicago: The University of Chicago Press, 1982

[29] Gilleskie, Donna B. and Amy L Harrison, "The Effect of Endogenous Health Inputs on the Relationship between Health and Education," Economics of Education Review, June 1998

[30] Goldin, Claudia, "How America Graduated From high School: 1910 to 1960," NBER Working Paper No 4762, 1994

[31] Goldin, Claudia and Lawrence Katz, "Why the United States Led on Education: Lessons from Secondary School Expansion, 1910 to 1940," NBER Working Paper 6144, August 1997

[32] Green, William H., Econometric Analysis, 3rd Edition, Prentice Hall, New Jersey 1997

[33] Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems ," Econometrica, January. 1977

[34] Grossman, Michael, "The Human Capital Model of the Demand for Health," NBER Working Paper 7078, April 1999

[35] Grossman, Michael, "The Correlation between Health and Schooling," in *Household Production and Consumption*, Ed N. E. Terleckyj, Studies in Income and Wealth, Vol. 40, Conference on Research in Income and Wealth. New York: Columbia University Press for the National Bureau of Economic Research, 1975

[36] Grossman, Michael, "The Demand for Health: A theoretical and Empirical Investigation," New York: Columbia University (for the NBER), 1972a

[37] Grossman, Michael, "On the Concept of Health Capital the Demand for Health," Journal of Political Economy 80, 1972b

[38] Grossman, Michael and R. Kaestner "Effects of Education on Health," in J.R. Berhman and N. Stacey Eds. *The Social Benefits of education*, University of Michigan Press, Ann Arbor, 1997

[39] Haines, Michael R. and Samuel H. Preston, "The Use of the Census to Estimate Childhood Mortality: Comparisons from the 1900 and 1910 United States Census Public Use Samples," Historical Methods, Vol 30, no.2, Spring 1997

[40] Harmon, Colm, and Ian Walker, "Estimates of the Economic Return to Schooling for the United Kingdom," American Economic Review, Volume 85, Issue 5, December 1995

[41] Hoek, H. W., A.S. Brown and E. Susser, "The Dutch Famine and Schizophrenia Spectrum Disorders," Social Psychiatry and Psychiatric Epidemiology, Volume 33, Issue 8, 1998

[42] Hyslop, Dean R. and Guido W. Imbens, "Bias from Classical and Other Forms of Measurement Error"NBER Technical Working Paper, August 2000

[43] Inter-American Development Bank. *Facing up to Inequality in Latin America: Economic and Social Progress in Latin America. 1998-1999 Report*. Distributed by John Hopkins University Press, Washington D.C., 1999

[44] Imbens, Guido W. and Joshua D. Angrist, "Identification of Local Average Treatment Effects," Econometrica, Volume 62, Issue 2, March 1994

[45] Kane, Thomas J, Cecilia Elena Rouse and Douglas Staiger, "Estiamting Returns to Schooling when Schooling is Misreported," NBER Working Paper 7235, July 1999

[46] Katz, Michael, "A History of Compulsory Education Laws," Phi Delta Kappa Educational Foundation 1976

[47] Kenkel, Donald, "Health Behavior, Health knowledge and Schooling," Journal of Political Economy, Volume 99, Issue 2, April 1991

[48] Kitagawa and Hauser, *Differential Mortality in the United States: a Study in Socioeconomic Epidemiology*. Cambridge, MA: Harvard University Press, 1973.

[49] Landes William and Lewis C. Solomon, "Compulsory Schooling Legislation: An economic Analysis of Law and Social Change in the Nineteenth Century," Journal of Economic History, March 1972

[50] Lang, Kevin and David Kropp, "Human Capital versus Sorting: The Effects of Compulsory Attendance Laws," Quarterly Journal of Economics, August 1986

[51] Leigh, J. Paul and Rachna Dhir, "Schooling and Frailty Among Seniors," Economics of Education Review, Volume 16, No. 1, 1997

[52] Lleras-Muney, Adriana, "Were State Laws on Compulsory Education Effective? An analysis from 1915 to 1939," mimeo, Columbia University, 2001

[53] Maddala, G.S. Limited Dependent and Qualitative Variables in Econometrics, Econometric Society Monographs No. 3, Cambridge University Press, 1997

[54] Margo, Robert A. and T. Aldrich Finegan, "Compulsory Schooling Legislation and School Attendance in Turn-of-The Century America: A 'Natural Experiment'," Economics Letters, October 1996

[55] Meghir, Costas and Marten Palme, "Assessing the Effect of Schooling on earnings Using a Social Experiment," Unpublished Working Paper, University College London, 1999

[56] Menchik, Paul L., "Economic Status as a Determinant of Mortality among Black and White Older Men: Does Poverty Kill?," Population Studies, November 1993

[57] Munasinghe, Lalith and Nachum Sicherman, "Why do Dancers Smoke? Time Preference, Occupational Choice, and Wage Growth," NBER Working Paper 7542, February 2000

[58] Nelson, Richard R. and Edmund S. Phelps, "Investment in Humans, Technological Diffusion, and Economic Growth," American Economic Review, Volume 56, Issue1/2, March 1966

[59] Newhouse, Joseph P., *Free for All? Lessons from the Rand Health Insurance Experiment*. Cambridge: Harvard University press. 1993

[60] Nordhaus, William D., "The Health of Nations: The Contribution of Improved Health to Living Standards," mimeo, Yale University, November 1999

[61] Pappas, Gregory, Susan Queen, Wolber Hadden and Grail Fisher, "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986," The New England Journal of Medicine, 1993,

[62] Perri, Timothy J., "Health Status and Schooling Decisions of Young Men," Economics of Education Review, 1984

[63] Rogers, Richard G., Robert A. Hummer and Charles B. Nam, *Living and Dying in the USA*, Academic Press 2000

[64] Roseboom, T.J. et al, "Coronary heart Disease after Prenatal Exposure to the Dutch Famine, 1944-45," Heart, Volume 84, Issue 6, 2000

[65] Rosenzweig M.R and T.P. Schultz, "Education and Household Production of Child Health," In Proceedings of the American Statistical Association (Social Statistics Section) Washington, DC: American Statistical Association, 1991

[66] Steven Ruggles and Matthew Sobek et. al., *Integrated Public Use Microdata Series: Version 2.0 Minneapolis: Historical Census Projects*, University of Minnesota, 1997

[67] Sander, William, "Schooling and Quitting Smoking," Review of Economics and Statistics, 77, 1995

[68] Schmidt, Stefanie, "School Quality, Compulsory Education Laws, and the Growth of American High School Attendance, 1915-1935," MIT Ph.D. Dissertation 1996

[69] Staiger, Douglas and James H. Stock, "Instrumental Variables Regression with Weak Instruments," Econometrica;65(3), May 1997

[70] Starr, Paul, "The Social Transformation of American Medicine," New York: Basic Books, 1982

[71] Stigler, George, "Employment and Compensation in Education," NBER Occasional Paper. Number 53. 1950

[72] Strauss, John and Duncan Thomas, "Human Resources: Empirical Modelling of Household and Family Decisions," in *Handbook of Development Economics*, Vol. III, Edited by J. Behrman, T.N. Srinivasan. Elsevier Science 1995

[73] Tyack, David *One Best system: A History of American Urban Education.* Harvard University Press. Cambridge, Massachusetts, 1974

[74] Wolfe, Barbara and Jere R. Behrman, "Women's Schooling and Children's Health: Are the Effects Robust with Adult Sibling Control for the Women's Childhood Background?," Journal of Health Economics 6 , 1987

[75] Woltz, Charles, K.,"Compulsory Attendance at School," in Law and Contemporary Problems, School of Law, Duke University Vol. 20 Winter 1955

Appendix: Trends for Compulsory education and Child Labor laws

Compulsory Attendance Laws

Age at which must enter school (enter age)

	States 1915	States 1928	States 1939
6	0	2	2
7	16	28	33
8	25	17	13
9	1	1	0
Total	42	48	48

Child Labor Laws

Minimun age to get work permit (work age)

	States 1915	States 1928	States 1939
12	2		
13	1		
14	38	42	32
15	4	4	4
16	0	2	12
Total	45	48	48

Continuation School Laws

Have Continuation School Laws

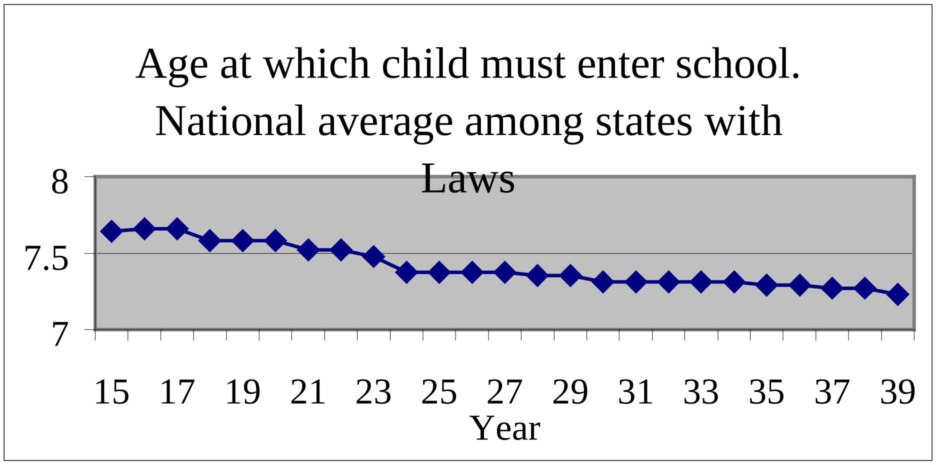
	States 1915	States 1928	States 1939
0	36	20	19
1	12	28	29
Total	48	48	48

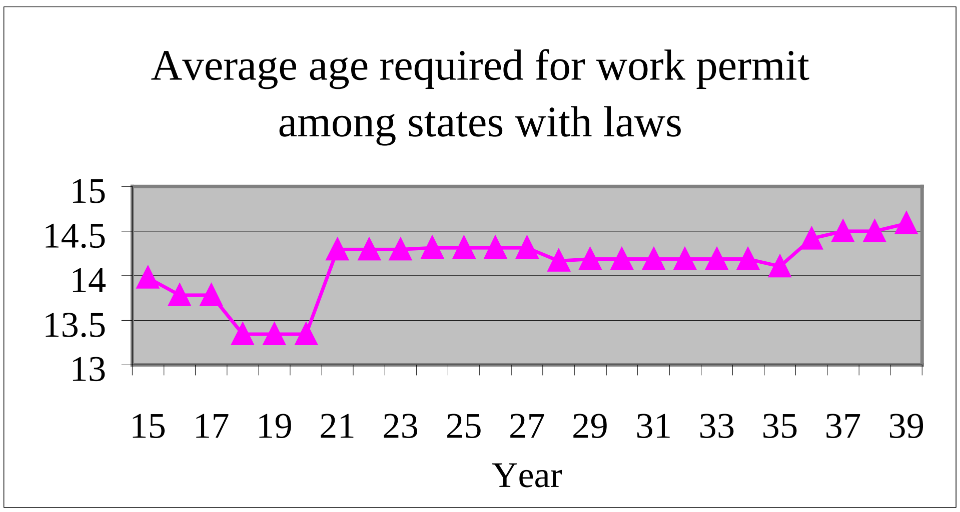
Constructed Variable: Implicit number of years had to attend school

Childcom = work age - enter age

	States 1915	States 1928	States 1939
0	8		
4	1		
5	2	1	
6	21	15	9
7	14	26	23
8	2	5	7
9			8
10		1	1
Total	48	48	48

Time trends for the laws





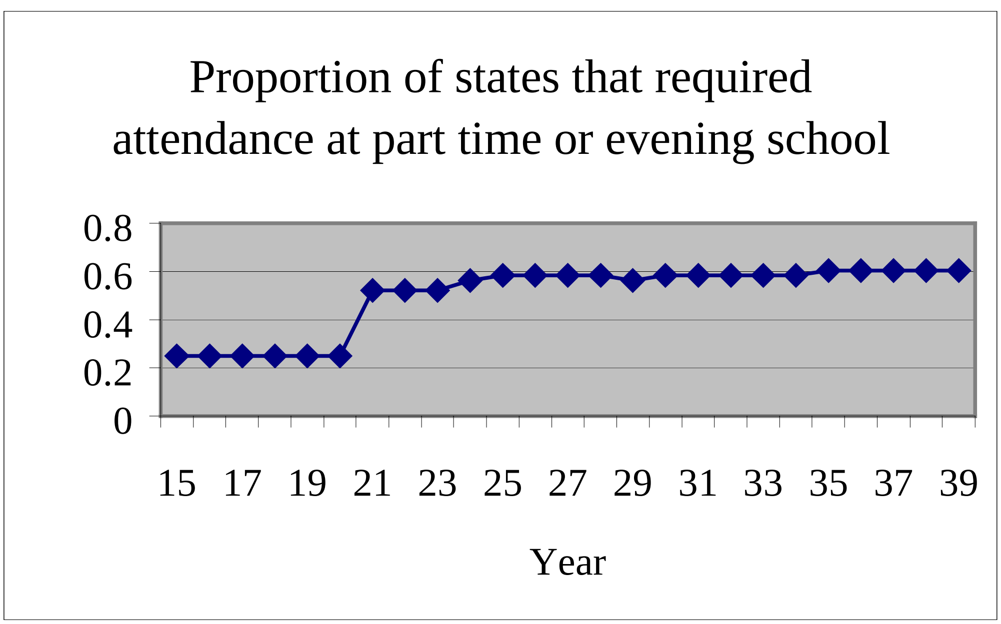


TABLE 1: SUMMARY STATISTICS- AGGREGATED CENSUS DATA

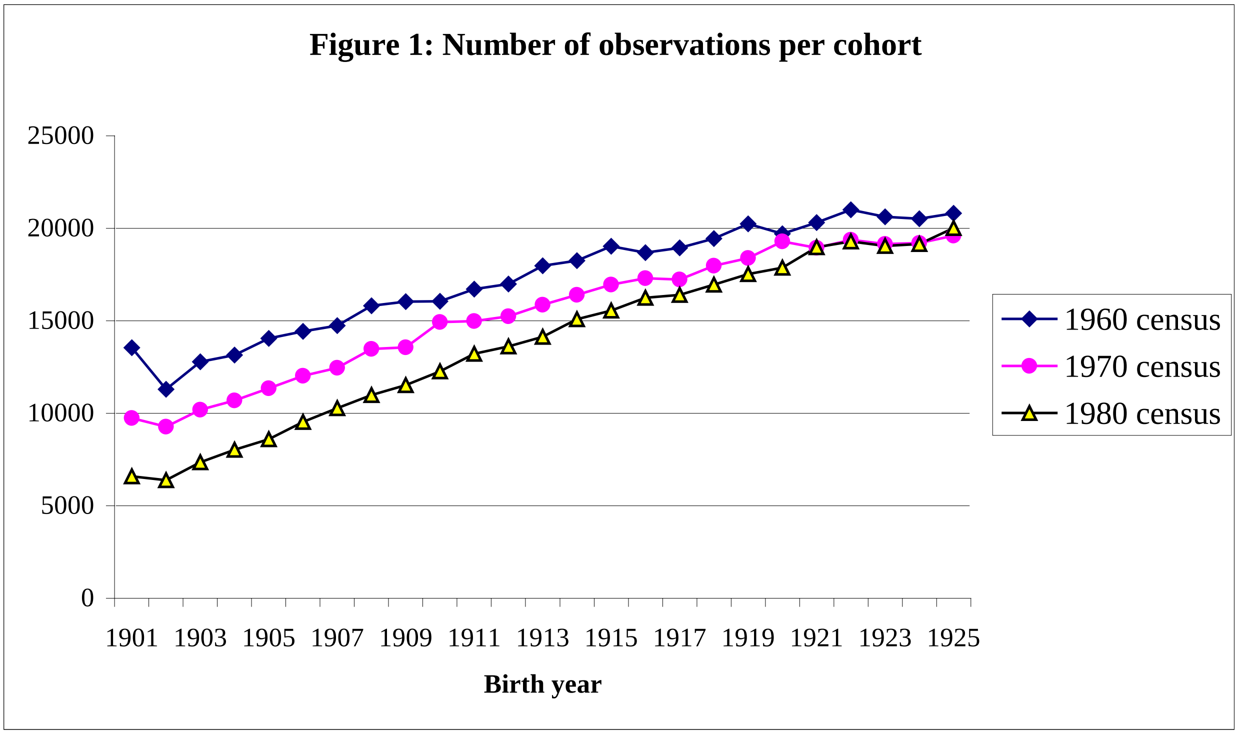
Variables		Mean	Std. Dev.	Min	Max
Individual characteristics	10-year death rate	0.106	0.136	-7	0.875
	Years of completed education	10.697	1.020	4.818	18
	1970 Dummy	0.471	0.499	0	1
	Female	0.517	0.500	0	1
	Married	0.818	0.096	0	1
	Live in North	0.255	0.369	0	1
	Live in West	0.285	0.351	0	1
	Live in South	0.159	0.227	0	1
	Live in an urban area	0.685	0.122	0	1
	Age	50.366	8.482	35	69
	Born in 1901	0.029	0.167	0	1
	Born in 1902	0.025	0.157	0	1
	Born in 1903	0.028	0.166	0	1
	Born in 1904	0.029	0.169	0	1
	Born in 1905	0.031	0.174	0	1
	Born in 1906	0.032	0.177	0	1
	Born in 1907	0.033	0.180	0	1
	Born in 1908	0.036	0.186	0	1
	Born in 1909	0.036	0.187	0	1
	Born in 1910	0.038	0.191	0	1
	Born in 1911	0.039	0.193	0	1
	Born in 1912	0.040	0.195	0	1
	Born in 1913	0.042	0.200	0	1
	Born in 1914	0.043	0.202	0	1
	Born in 1915	0.044	0.205	0	1
	Born in 1916	0.044	0.205	0	1
	Born in 1917	0.044	0.206	0	1
	Born in 1918	0.046	0.209	0	1
	Born in 1919	0.047	0.213	0	1
	Born in 1920	0.048	0.213	0	1
	Born in 1921	0.048	0.214	0	1
	Born in 1922	0.050	0.217	0	1
	Born in 1923	0.049	0.216	0	1
	Born in 1924	0.049	0.215	0	1
	Born in 1925	0.050	0.217	0	1
State-of-Birth Characteristics	% Urban	53.523	21.279	12.300	97.500
	% Foreign	11.737	8.523	0.400	31.300
	% Black	8.983	11.901	0.010	54.200
	% Employed in manufacturing	0.067	0.039	0.003	0.283
	Annual Manufacturing wage	7161.911	1368.253	713.030	12095.160
	Value of farm per acre	540.048	276.353	47.700	1802.575
	Per capita number of doctors	0.001	0.000	0.000	0.003
	Per capita education expenditures	96.474	42.142	5.372	601.391
	Number of school buildings per sq. mile	0.174	0.090	0.002	0.474

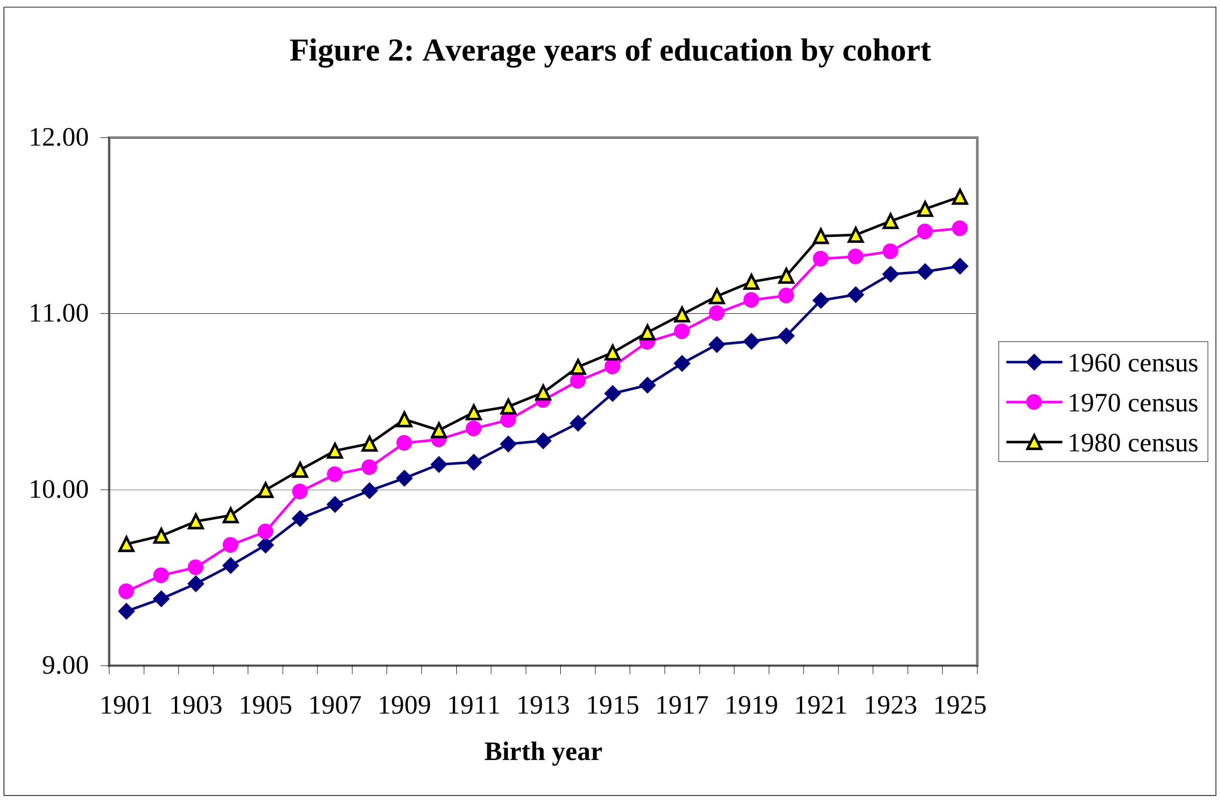
Number of observations: 4795, corresponding to cells defined at the gender, state-of-birth, and cohort. All means calculated using weights, where the weights are given by the number of observations in each cell. Monetary values are in 1982-84 dollars

TABLE 2: SUMMARY STATISTICS- NHEFS

Variables		Mean	Std. Dev.	Min	Max
Individual characteristics	Died between 1975 and 1985	0.254	0.435	0	1
	Years of completed education	10.360	3.326	0	17
	Female	0.540	0.498	0	1
	Married	0.755	0.430	0	1
	Live in North	0.214	0.410	0	1
	Live in West	0.250	0.433	0	1
	Live in South	0.269	0.444	0	1
	Live in an urban area	0.526	0.499	0	1
	Age	62.941	7.561	50	74
	Born in 1901	0.039	0.193	0	1
	Born in 1902	0.054	0.226	0	1
	Born in 1903	0.056	0.230	0	1
	Born in 1904	0.056	0.230	0	1
	Born in 1905	0.061	0.239	0	1
	Born in 1906	0.068	0.251	0	1
	Born in 1907	0.055	0.227	0	1
	Born in 1908	0.042	0.200	0	1
	Born in 1909	0.025	0.155	0	1
	Born in 1910	0.026	0.160	0	1
	Born in 1911	0.027	0.161	0	1
	Born in 1912	0.028	0.165	0	1
	Born in 1913	0.028	0.165	0	1
	Born in 1914	0.031	0.174	0	1
	Born in 1915	0.033	0.178	0	1
	Born in 1916	0.032	0.177	0	1
	Born in 1917	0.034	0.182	0	1
	Born in 1918	0.035	0.184	0	1
	Born in 1919	0.041	0.198	0	1
	Born in 1920	0.037	0.188	0	1
	Born in 1921	0.038	0.192	0	1
	Born in 1922	0.039	0.194	0	1
	Born in 1923	0.034	0.182	0	1
	Born in 1924	0.044	0.206	0	1
	Born in 1925	0.036	0.187	0	1
State-of-Birth Characteristics	% Urban	49.846	20.734	12.3	97.5
	% Foreign	11.489	8.434	0.4	31.3
	% Black	10.108	13.652	0.01	53.8
	% Employed in manufacturing	0.065	0.040	0.003	0.283
	Annual Manufacturing wage	6971.696	1380.099	713.030	11007.230
	Value of farm per acre	549.371	292.371	48.484	1802.575
	Per capita number of doctors	0.0013	0.0003	0.0002	0.0026
	Per capita education expenditures	86.305	44.411	5.372	601.391
	Number of school buildings per sq. mile	0.173	0.092	0.003	0.474

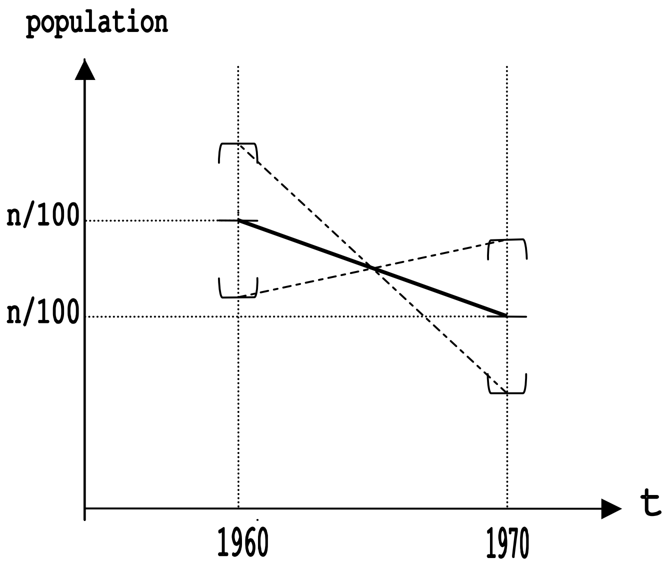
Number of observations: 4554. Monetary values are in 1982-84 dollars





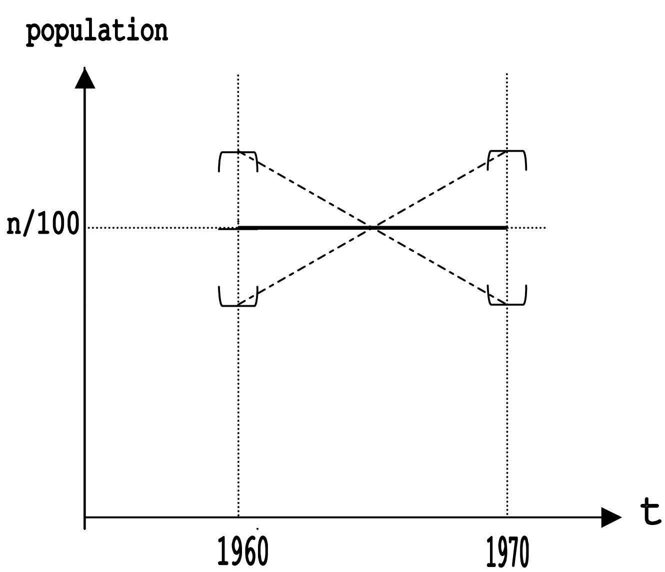
Note: Figures 1 and 2 follow the same cohorts from the 1960 census up to the 1980 census.
In Figure 1 we can observe that that 10-year mortality increases with age: for older cohorts the number of individuals observed in 1980 is much smaller than in 1960 or 1970.
In figure 2 we can see that the average level of education is higher in 1980 than in 1960 for all cohorts, suggesting that those who died in each cohort had bellow average levels of education.

Figure 3: Calculating Death Rates with the Census



The 1960 and 1970 census are 1/100 random samples of the population, therefore the number of individuals in any given group is always observed with error. Because of this sampling error the death rates for any given group are over-estimated 50% of the time and underestimated 50% of the time. However, since the sampling is truly random, the observed death rates are consistent estimates of the true death rates.

Figure 3B: An example for a young cohort: 0 death rate



If the true death rate is 0 then I observe 50% negative death rates. As cohorts age, the death rate increases (see example above) the number of negative death rates falls.

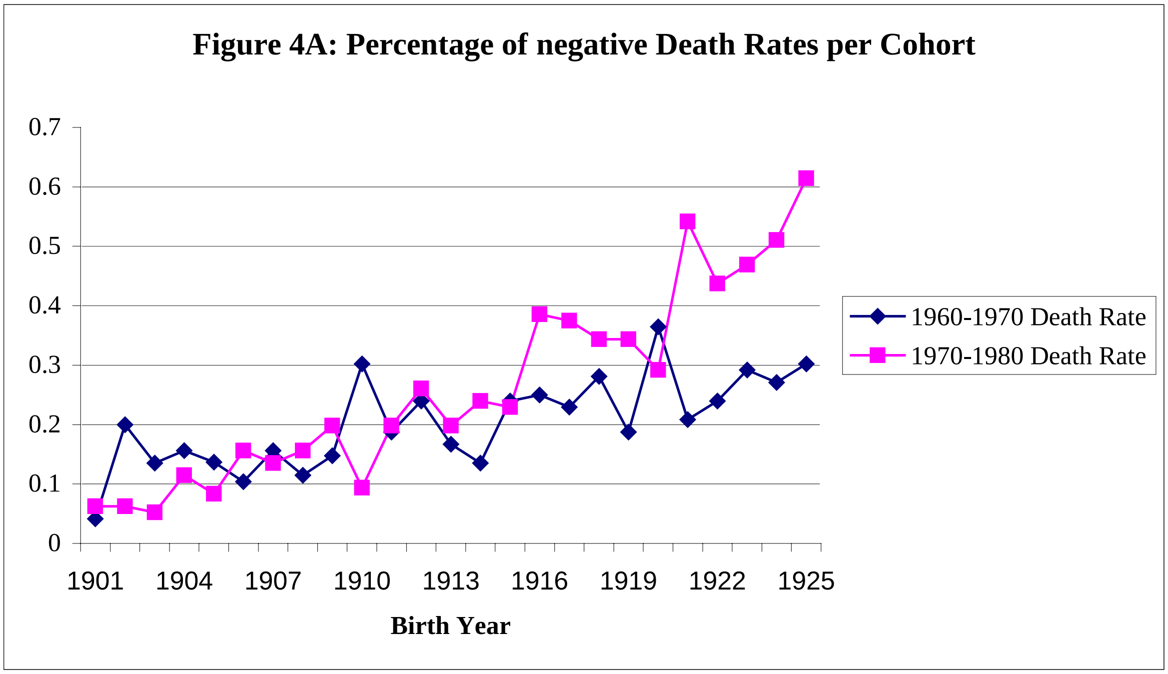


Figure 4A: Percentage of negative Death Rates per Cohort
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1901
1904
1907
1910
1913
1916
1919
1922
1925
Birth Year
1960-1970 Death Rate
1970-1980 Death Rate

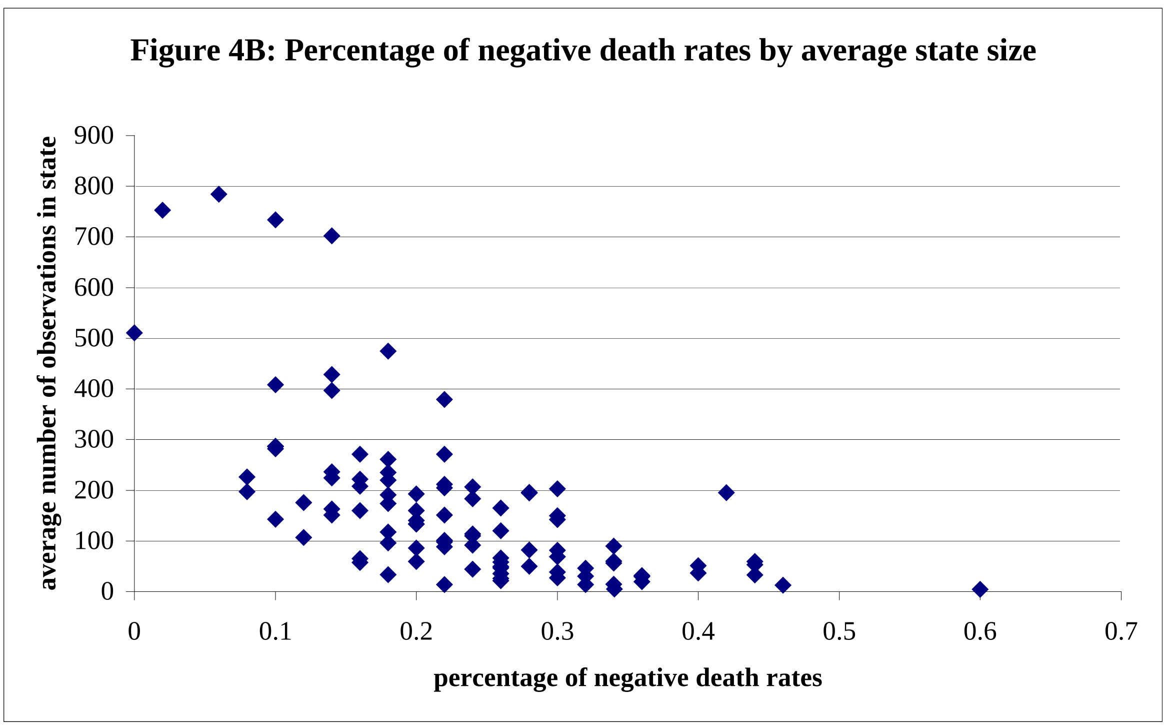


Figure 4B: Percentage of negative death rates by average state size
average number of observations in state
900
800
700
600
500
400
300
200
100
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
percentage of negative death rates

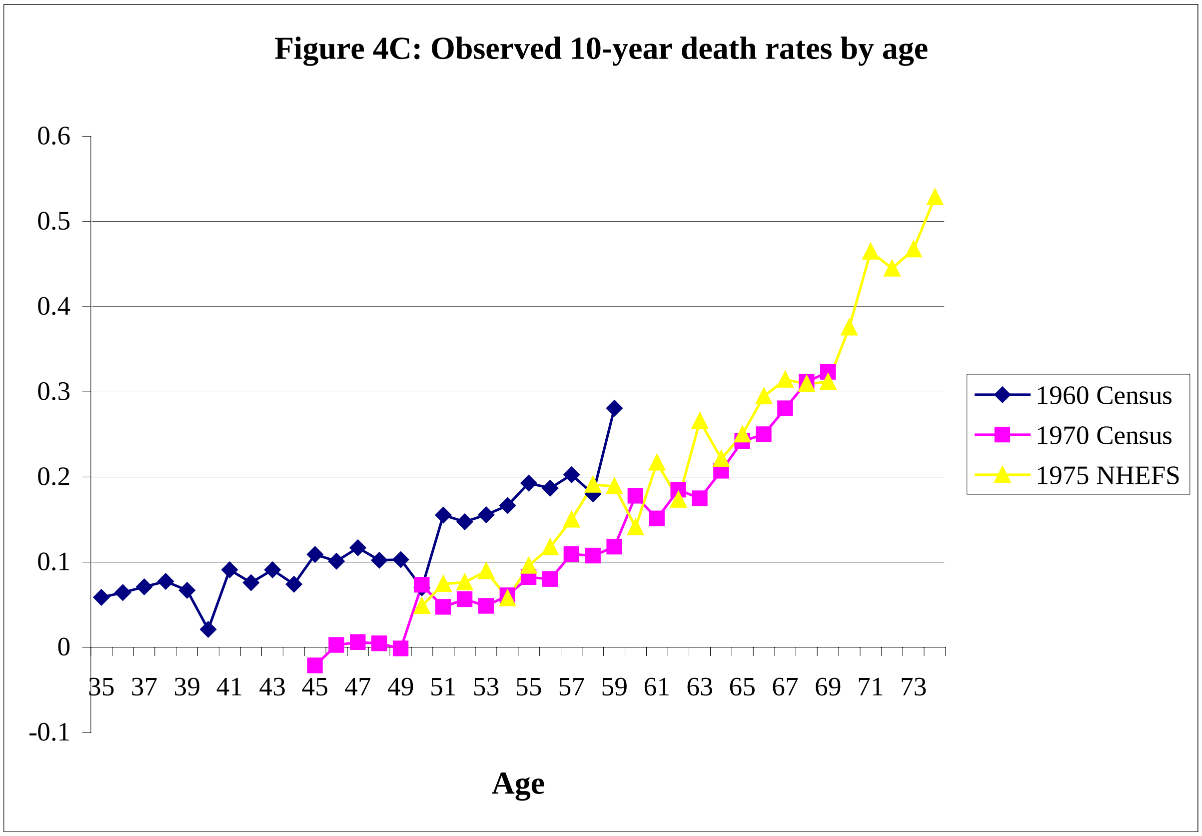


Figure 4C: Observed 10-year death rates by age
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73
Age
1960 Census
1970 Census
1975 NHEFS

Figure 5: Average education level by years of compulsory education

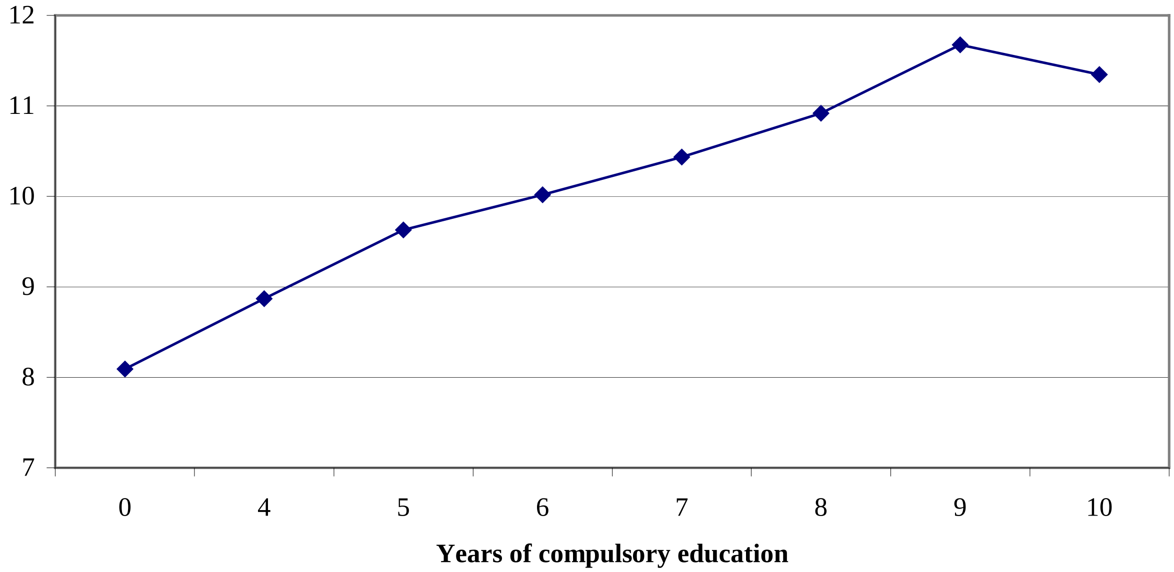


Figure 6: Average education level by compulsory education for selected cohorts

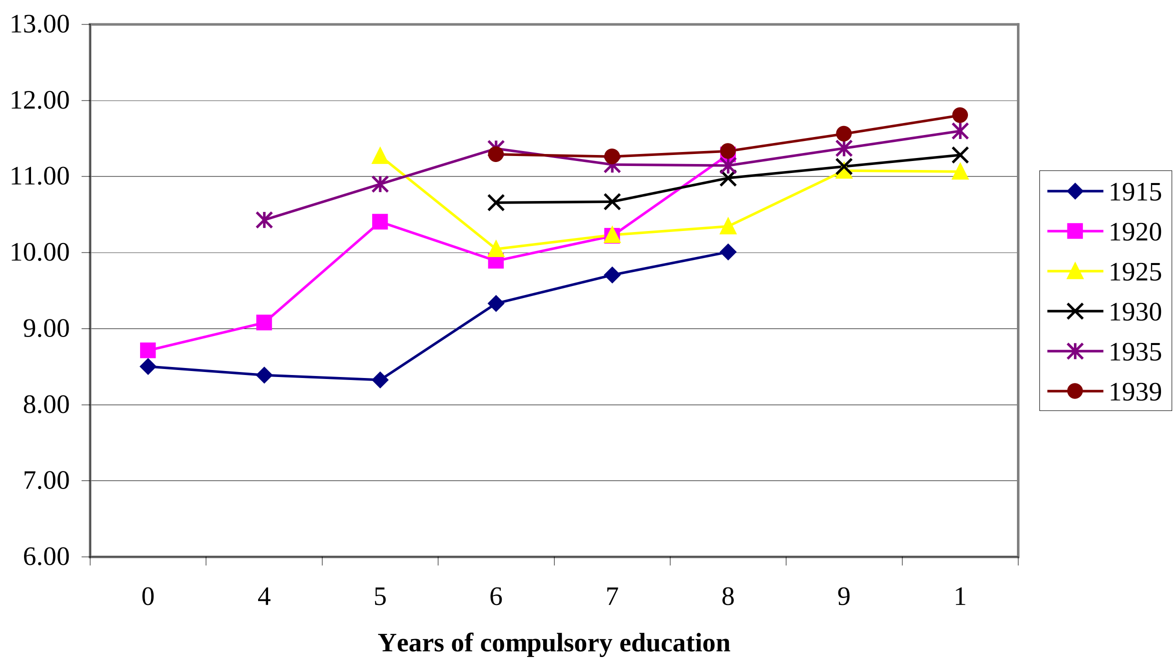


TABLE 3: EFFECT OF COMPULSORY EDUCATION LAWS ON EDUCATION

Variables		Individual data	Individual data	Aggregate data
Dependent Variable	Education			
Education Laws	Childcom Category 4[(a)]	0.347**	0.355**	0.323**
		(0.075)	(0.088)	(0.077)
	Childcom Category 5	0.323**	0.260**	0.321**
		(0.090)	(0.099)	(0.098)
	Childcom Category 6	0.274**	0.302**	0.266**
		(0.069)	(0.087)	(.074)
	Childcom Category 7	0.385**	0.408**	0.369**
		(0.070)	(0.087)	(0.074)
	Childcom Category 8	0.416**	0.398**	0.315**
		(0.075)	(0.089)	(0.078)
	Childcom Category 9	0.580**	0.512**	0.470**
		(0.076)	(0.092)	(0.084)
	Childcom Category 10	0.325**	0.328**	0.318**
		(0.080)	(0.095)	(0.095)
	Continuation School Required (=1)	0.027	0.017	0.027
		(0.026)	(0.028)	(0.031)
Individual characteristics	Female	0.116**	0.140**	-0.012
		(0.014)	(0.014)	(0.014)
	Married		0.433**	-1.111**
			(0.011)	(0.097)
	Live in an urban area		0.968**	0.278**
			(0.019)	(0.138)
State-of-Birth Characteristics	% Urban		0.021**	0.028**
			(0.004)	(0.005)
	% Foreign		-0.002	0.004
			(0.008)	(0.010)
	% Black		0.024**	0.020**
			(0.009)	(0.010)
	% Employed in manufacturing		-0.350	-1.220**
			(0.513)	(0.621)
	Annual Manufacturing wage		0.000	0.000
			(0.000)	(0.000)
	Value of farm per acre		0.000	0.000
			(0.000)	(0.000)
	Per capita number of doctors		150.018**	188.615**
			(71.840)	(66.2)
	Per capita education expenditures		0.001**	0.000
			(0.000)	(0.000)
	Number of school buildings per sq. mile		-0.359	-0.166
			(0.289)	(0.373)
	3 region of residence dummies	No	Yes	Yes
	Region of birth*cohort dummies	No	Yes	Yes
	R-Squared	0.0811	0.1052	0.888
	F-statistic on instruments	14.93**	8.37**	4.49**
	Partial R-squared	0.0003	0.0001	0.0108

* significant at 10% ** significant at 5%. All regressions include a dummy for the 1970 census, state-of-birth dummies, cohort dummies and an intercept. For the individual-level regressions (1 and 2) N=814805 and the standard errors (in parenthesis) are clustered at the state-of-birth and cohort level.

(a) Childcom=work permit age - entry age. Childcom=0 is the excluded category.

(b) Data aggregated by gender/cohort/state-of-birth. Robust standard errors. N=4792

TABLE 4: EFFECT OF EDUCATION ON MORTALITY-LEAST SQUARE RESULTS

Variables						
Data		Census	NHEFS	NHEFS	NHEFS	NHEFS
Method		WLS	OLS	Probit[a]	WLS	W LS
Level[d]		Aggregate[b]	Individual	Individual	Aggregate[b]	Aggregate[c]
Dependent Variable		10-year death rate	died 75-85	died 75-85	death rate 75-85	death rate 75-85
Individual characteristics	Education	-0.012**	-0.012**	-0.011**	-0.017**	-0.013**
		(0.004)	(0.002)	(0.002)	(0.004)	(0.005)
	Female	-0.048**	-0.147**	-0.151**	-0.137**	-0.139**
		(0.004)	(0.013)	(0.013)	(0.015)	(0.030)
	Married	0.227**	-0.044**	-0.053	-0.005	-0.015
		(0.030)	(0.016)	(0.015)	(0.030)	(0.037)
	Live in an urban area	-0.136**	0.037**	0.039**	0.056**	0.080**
		(0.044)	(0.015)	(0.015)	(0.024)	(0.030)
State-of-Birth Characteristics	% Urban	0.000	-0.003	-0.003	-0.002	-0.002
		(0.001)	(0.005)	(0.005)	(0.005)	(0.005)
	% Foreign	0.000	0.005	0.012	0.005	0.005
		(0.002)	(0.007)	(0.008)	(0.007)	(0.007)
	% Black	0.000	-0.014*	-0.012	-0.014	-0.014
		(0.002)	(0.008)	(0.008)	(0.009)	(0.009)
	% Employed in manufacturing	-0.075	-0.085	-0.060	-0.091	-0.100
		(0.105)	(0.590)	(0.563)	(0.621)	(0.640)
	Annual Manufacturing wage	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Value of farm per acre	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Per capita number of doctors	-2.043	1.058	17.451	-0.762	-2.139
		(14.384)	(48.228)	(39.746)	(49.833)	(52.857)
	Per capita education expenditures	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	# of school buildings per sq. mile	0.045	0.712**	0.744**	0.725**	0.758**
		(0.064)	(0.345)	(0.334)	(0.360)	(0.380)
	N	4792	4554	4554	1557	942
	R-Squared	0.3685	0.1736		0.3952	0.5219

All regressions include 24 cohort dummies, 47 state of birth dummies, region-of-birth * cohort, region of residence dummies and an intercept. Standard errors (in parenthesis) are clustered at the state-of-birth and cohort level. The census regressions also include a dummy for the 1970 census.

(a) The reported coefficients are the mean marginal effects. The standard errors are calculated using the Delta Method.
(b) Data are aggregated at the cohort/gender and state-of-birth level.
(c) Data aggregated at the cohort and state-of-birth level only.
(d) All regressions at the aggregate level are weighted by the number of observations in the original cell

* significant at 10% ** significant at 5%.

TABLE 5: EFFECT OF EDUCATION ON MORTALITY-LEAST SQUARES RESULTS BY GENDER

Variables		Males	Females
Dependent Variable	10-year death rate		
Individual characteristics	Education	0.008	0.011
		(0.007)	(0.008)
	Married	-0.293**	-0.324**
		(0.081)	(0.057)
	Dummy for 1970	0.017**	-0.063**
		(0.005)	(0.008)
	Live in an urban area	-0.080	-0.086
		(0.058)	(0.065)
State-of-Birth Characteristics	% Urban	-0.001	0.000
		(0.002)	(0.002)
	% Foreign	-0.002	0.001
		(0.004)	(0.004)
	% Black	-0.001	-0.002
		(0.004)	(0.004)
	% Employed in manufacturing	-0.023	-0.072
		(0.313)	(0.261)
	Annual Manufacturing wage	0.000	0.000
		(0.000)	(0.000)
	Value of farm per acre	0.000	0.000
		(0.000)	(0.000)
	Per capita number of doctors	11.124	-18.978
		(24.294)	(28.178)
	Per capita education expenditures	0.000	0.000
		(0.000)	(0.000)
	Number of school buildings per sq. mile	-0.063	0.096
		(0.144)	(0.151)
	N	2397	2395
	R-Squared	0.4668	0.2297

All regressions include 24 cohort dummies, 47 state-of-birth dummies, region-of-birth * cohort interactions, region-of-residence dummies and an intercept. All regressions are weighted by the number of observations in the original cell. Standard errors (in parenthesis) are robust.

* significant at 10% ** significant at 5%.

TABLE 6: EFFECT OF EDUCATION ON MORTALITY-IV RESULTS

Variables					
Data		NHEFS[b]	Census[a][c]	Census[a][b][c]	Census[a][b][c]
Method		2SLS	Wald	2SLS	Mixed 2SLS
Level		Individual	Aggregate	Aggregate	Aggregate
Dependent Variable		Died 1975-1985	10-year death rate	10-year death rate	10-year death rate
Individual characteristics	Education	-0.020	-0.037**	-0.045*	-0.059**
		(0.054)	(0.006)	(0.026)	(0.027)
	1970 Dummy		0.003	0.012**	0.021**
			(0.004)	(0.005)	(0.007)
	Female	-0.142**	-0.071**	-0.048**	-0.040**
		(0.030)	(0.004)	(0.004)	(0.005)
	Married	-0.040		0.190**	0.266**
		(0.027)		(0.041)	(0.031)
	Live in an urban area	0.046		-0.126**	-0.080
		(0.055)		(0.045)	(0.054)
State-of-Birth Characteristics	% Urban	-0.002		0.001	0.001
		(0.005)		(0.001)	(0.001)
	% Foreign	0.005		0.001	0.001
		(0.007)		(0.002)	(0.002)
	% Black	-0.014		0.001	0.001
		(0.008)		(0.002)	(0.002)
	% Employed in manufacturing	-0.089		-0.118	-0.080
		(0.605)		(0.113)	(0.137)
	Annual Manufacturing wage	0.000		0.000	0.000
		(0.000)		(0.000)	(0.000)
	Value of farm per acre	0.000		0.000	0.000
		(0.000)		(0.000)	(0.000)
	Per capita number of doctors	7.298		6.078	6.675
		(62.347)		(15.337)	(17.31)
	Per capita education expenditures	0.000		0.000	0.000
		(0.000)		(0.000)	(0.000)
	Number of school buildings per sq. mile	0.698**		0.051	0.044
		(0.350)		(0.066)	(0.075)
	State-of-birth Dummies	Yes	No	Yes	Yes
	Region of Birth Dummies	No	Yes	No	No
	Cohort Dummies	Yes	Yes	Yes	Yes
	Region-of-birth*cohort	Yes	No	Yes	Yes
	Region of residence dummies	Yes	No	Yes	Yes
	N	4554	1396	4792	4792

All regressions include an intercept.

(a) Regressions are weighted by the number of observations in the original cell.

(b) Standard errors (in parenthesis) are clustered at the state-of-birth and cohort level and have been corrected in the second stage.

(c) Note: P2SLS and aggregate 2SLS use data aggregated at the gender/cohort/state-of-birth. Wald uses data aggregated at the gender/cohort/region-of-birth/compulsory education laws level.

* significant at 10% ** significant at 5%.

TABLE 7: ADDITIONAL ESTIMATIONS

A: Quarter of birth, laws and interactions used as instruments

		2SLS	Mixed 2SLS
Variables			
Dependent Variable	10-year death rate		
Individual	Education	-.067**	-.062**
		(.0260)	(.024)

B: Reduced form Results (OLS)

Variables		
Dependent Variable	10-year death rate	
	Childcom	-0.0027**
		(0.0013)
	Continuation school	-0.0032
		(0.005)

C: Age Heaping: Exclude ages 40, 50, and 60

		2SLS	Mixed 2SLS
Variables			
Dependent Variable	10-year death rate		
Individual	Education	-.040	-.052
		(.026)	(.026)

All regressions include the same controls as in Table 6.

* significant at 10% ** significant at 5%.

TABLE 8: EFFECT OF EDUCATION ON MORTALITY-IV RESULTS BY GENDER

		2SLS	2SLS	Mixed 2SLS	Mixed 2SLS
Variables		Males	Females	Males	Females
Dependent Variable	10-year death rate				
Individual characteristics	Education	-0.047	-0.044	-0.077	-0.054
		(0.051)	(0.063)	(0.040)	(0.039)
	Married	-0.292**	-0.311**	-0.263**	-0.297**
		(0.082)	(0.059)	(0.087)	(0.065)
	Dummy for 1970	0.025	-0.057**	0.032**	-0.052**
		(0.009)	(0.011)	(0.009)	(0.011)
	Live in an urban area	-0.096	-0.102	-0.002	-0.036
		(0.059)	(0.067)	(0.079)	(0.078)
State-of-Birth Characteristics	% Urban	0.002	0.001	0.001	0.001
		(0.003)	(0.002)	(0.002)	(0.002)
	% Foreign	-0.001	0.001	-0.001	0.001
		(0.004)	(0.004)	(0.003)	(0.003)
	% Black	0.000	-0.001	0.000	-0.001
		(0.004)	(0.004)	(0.003)	(0.003)
	% Employed in manufacturing	-0.095	-0.128	-0.056	-0.103
		(0.328)	(0.264)	(0.182)	(0.194)
	Annual Manufacturing wage	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
	Value of farm per acre	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
	Per capita number of doctors	22.270	-6.219	27.940	-5.860
		(24.002)	(32.441)	(22.782)	(26.29)
	Per capita education expenditures	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
	Number of school buildings per sq. mile	-0.018	0.078	-0.057	0.094
		(0.148)	(0.155)	(0.103)	(0.106)
	N	2397	2395	2397	2395

All regressions include 24 cohort dummies, 47 state-of-birth dummies, region-of-birth * cohort interactions, region-of-residence dummies and an intercept. All regressions are weighted by the number of observations in the original cell. Standard errors (in parenthesis) are robust.

* significant at 10% ** significant at 5%.

TABLE 9: TESTING FOR THE FUNCTIONAL FORM OF THE EFFECT OF EDUCATION ON MORTALITY

Variables		
Method		OLS
Dependent Variable	10-year death rate	
Individual characteristics	Education	-0.064**
		(0.029)
	Education Squared	0.003*
		(0.001)
	1970 Dummy	0.010
		(0.005)
	Female	-0.047
		(0.004)
	Married	0.224
		(0.030)
	Live in an urban area	-0.138
		(0.044)
State-of-Birth Characteristics	% Urban	0.000
		(0.001)
	% Foreign	0.000
		(0.002)
	% Black	0.000
		(0.002)
	% Employed in manufacturing	-0.058
		(0.107)
	Annual Manufacturing wage	0.000
		(0.000)
	Value of farm per acre	0.000
		(0.000)
	Per capita number of doctors	-4.526
		(14.546)
	Per capita education expenditures	0.000
		(0.000)
	Number of school buildings per sq. mile	0.041
		(0.064)

All regressions include 24 cohort dummies, 47 state-of-birth dummies, region-of-birth * cohort interactions, region-of-residence dummies and an intercept. All regressions are weighted by the number of observations in the original cell. Standard errors (in parenthesis) are clustered at the state-of-birth and cohort level. Estimated using census data aggregated at the gender/state-of-birth/cohort level. N=4792

TABLE 10: EFFECT OF THE DISTRIBUTION OF EDUCATION ON MORTALITY: ARE THERE EXTERNALITIES?

Variables		
Method		OLS
Dependent Variable	10-year death rate of those with more than 12 years of schooling	
	Average Education (12 years of schooling or more)	.010
		(0.011)
	Standard deviation of education of the entire distribution of education (education from 0 to 17)	-.092**
		(.025)

All regressions include the covariates in previous table plus 24 cohort dummies, 47 state-of-birth dummies, region-of-birth * cohort interactions, region-of-residence dummies and an intercept. All regressions are weighted by the number of observations in the original cell. Standard errors (in parenthesis) are clustered at the state-of-birth and cohort level. Data aggregated at the gender/state-of-birth/cohort level. N=4792