NBER WORKING PAPER SERIES

EDUCATION AND HEALTH OVER THE LIFE CYCLE

Robert Kaestner
Cuiping Schiman
Jason M. Ward

Working Paper 26836
http://www.nber.org/papers/w26836

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2020

Cuiping Schiman was supported during this project by a NHLBI grant (R01-60036640) awarded to Dr. Norrina Allen. Sponsors did not participate in the design or conduct of the study, or in the collection, analyses, and interpretation of the data, or in the preparation and review of the manuscript. The views expressed herein are those of the authors and do not necessarily reflect the views of the sponsors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26836.ack

Education and Health Over the Life Cycle
Robert Kaestner, Cuiping Schiman, and Jason M. Ward
NBER Working Paper No. 26836
March 2020
JEL No. I12,I14,I26

ABSTRACT

There is little theoretical and empirical research on the effects of education on health over the life cycle. In this article, we extend the Grossman (1972) model of the demand for health and use the extended model to analyze the effect of education on health at different ages. The main conclusion from our model is that it is unlikely that the relationship between education and health will be constant over the life cycle and that education is likely to have little effect on health at younger ages when there is little depreciation of the health stock. We also present an extensive empirical analysis documenting the association between education and health over the life cycle. Results of our analysis suggest that in terms of mortality, education has little effect until age 60, but then lowers the hazard rate of death. For measures of morbidity, education has an effect at most ages between 45 to 60, but after age 60 has apparently little effect most likely due to selective mortality. In addition, most of the apparent beneficial effect of education stems from obtaining a high school degree or more. It is the health and mortality of lowest education group—those with less than a high school degree—that diverges from the health and mortality of other education groups. Finally, we find that the educational differences in health have become larger for more recent birth cohorts.

Robert Kaestner
Harris School of Public Policy
University of Chicago
1307 East 60th Street (Room 3057)
Chicago, IL 60637
and NBER
kaestner@uchicago.edu

Cuiping Schiman
Georgia Southern University
11935 Abercorn Street
Savannah, GA 31419
cschiman@georgiasouthern.edu

Jason M. Ward
RAND Corporation
1776 Main Street
PO Box 2138
Santa Monica, CA 90407-2138
jward@rand.org

1. Introduction

There is an extensive literature on the effects of education on health.[1] Although it is generally accepted that education is a cause of health—more education leads to better health—this view is not universally accepted. Some scholars have delineated mechanisms linking more education to better health, for example, greater cognition to seek out, understand and implement health improving treatments (investments), but others have noted that there are common causes of both education and health that provide an alternative explanation for the positive association between education and health. The empirical evidence supporting a causal interpretation of this well-documented relationship is also not uniform. For example, several recent papers that used plausibly valid research designs exploiting educational reforms that increased educational attainment reported mixed results. Albouy and Lequien (2009), Clark and Royer (2013), Lager and Torsander (2012) and Meghir et al. (2018) concluded that educational reforms that increased education levels in France, United Kingdom (UK), and Sweden, respectively, had no significant effect on health. In contrast, Lleras-Muney (2005), Kemptner et al. (2011), and Davies et al. (2018) found that similar, and sometimes the same (UK), educational reform had a significant effect on health. Similarly, inconsistent findings were reported in twin studies, which also have some plausibility in terms of a research design to obtain a causal estimate (see Grossman 2015).

The debate over the causal link between education and health has many dimensions. For example, is education a cause of health because it provides specific knowledge of health and its determinants, or because it provides general knowledge and greater cognitive skills that are used to improve health? Another example is whether the association between years of schooling and health is approximately linear or, alternatively, are there educational thresholds at which the association is particularly pronounced? The complete list of questions and studies related to education and health is too long to review here and has been reviewed elsewhere (e.g.,Grossman 2006).

Despite the large literature on the relationship between education and health, surprisingly little of it is concerned with the life cycle nature of the relationship. This is an important gap in the literature because there are theoretical reasons to expect the relationship between education and health to change over the life cycle. In addition, ignoring the possibility that the relationship between education and health is likely to change over the life cycle may explain why several empirical studies conclude that education has no effect on health. In this article, we address these issues. First, we extend the canonical model of Grossman (1972) to show theoretically that the relationship between education and health is likely to change over the life cycle. Specifically, we allow the productivity of investments in health to differ by age, which is itself a novel extension that has implications for the life cycle pattern of health and health investments, and allow the effect of education to differ by the type of investment. Only one other study that we are aware of has provided a theoretical analysis of the relationship between education and health over the life cycle, but does so in a fundamentally different way than we do (Galama and van Kipperslius 2015).

Second, there is relatively little empirical research documenting the age profile of the relationship between education and health. We analyze multiple data sets using an approach to measure the effect of education on health that, to our knowledge, has not been previously used despite it being a relatively simple derivation obtained from the health production function of the human capital model of the demand for health (Grossman 1972). We estimate the effect of education at various ages over the life cycle of several birth cohorts. For each birth cohort or, in some analyses, each person, we measure the change in health between two ages, and then compare this change in health between these ages for those with more or less education. While seemingly trivial, this measure of the effect of education on health differs fundamentally from most estimates in the literature and it identifies the association between education and health at each age. The importance of our empirical contribution is illustrated by contrasting it to the large, descriptive literature analyzing education-related differences in

[1] For reviews, see: Grossman 2000; Grossman 2006; Cutler and Lleras-Muney 2008; Grossman 2015, and Galama et al. 2018.

earnings by age and period. Education-related differences in health are arguably as, if not more, important than education-related differences in earnings. It is also worth noting that, like our analysis, virtually all of the studies examining education-related differences in earnings over time and birth cohort are mainly descriptive (e.g., Katz and Murphy 1992; Card and Lemieux 2001; Autor et al. 2008).

Our empirical analysis also makes a few other contributions. First, restricting the sample to people in the same birth cohort, as we do, eliminates compositional changes within educational groups that have occurred over time (across cohorts) and that may have been significant. For example, it is possible that the underlying ability distribution of persons with a Bachelor's degree has changed over time as a greater proportion of persons have obtained this level of education. This and similar changes in composition of educational categories by birth cohort may confound estimates of the effect of education on health that combine birth cohorts. Second, focusing on specific birth cohorts allows us to examine whether the effect of education has changed over time, which may be expected given the changing medical technology available to different birth cohorts and plausible causal links between education and health.[2] Third, our approach obtains estimates of the effect of education at each age, which is consistent with our theoretical model and the changing scope for education to influence health that occurs over the life cycle, as the disease burden grows with age and the nature of investments in health changes. Identifying the effects of education on health by age can also clarify the role of selective mortality on the association between education and morbidity.

Results of the study suggest that education has a positive effect on health and mortality, but that this effect manifests at older ages. Early in adulthood, say before age 45, education seems to have little effect on health and mortality. This is an important finding because studies that combine samples of adults across a wide age range and that restrict the effect of education to be constant across ages will produce estimates of the effect of education on health and mortality that are attenuated. The second finding of our study is that most of the effect of education on health and mortality stems from the worse health and higher mortality of those with less than a high school degree. Among those with a high school degree or more, more education tends not to have substantial effects on health and mortality, although there are some, small beneficial effects of increased education beyond high school. Third, there is some evidence that the effect of education on health has become larger for more recent birth cohorts, although this was not found for mortality. Fourth, we find different patterns of the effect of education on physical and mental health, at least between ages 40 and 50, which are the only ages for which we used measures of both physical and mental health. Specifically, we find that education has little effect on mental health at those ages, which is consistent with the predictions of our theoretical model and the lower rate of depreciation of mental health and a smaller potential role for education to influence mental health vis-à-vis physical health at those ages. Finally, we found evidence of a convergence of morbidity between more or less educated persons as they age. As a greater proportion of low-educated persons die as a cohort ages, the remaining members of the low-educated group are likely positively selected. This selective mortality effect makes it appear as if education has no effect on measures of morbidity later in life, which is a possibility that many prior studies have not considered.

2. Education and Health over the Lifecycle: Previous Literature

There have been few theoretical analyses of the life cycle relationship between education and health. A prominent hypothesis in sociology is that socioeconomic status, and in particular education, has a cumulative effect on health as people age (e.g., House et al. 1994; Ross and Wu 1996). The logic underlying this "cumulative advantage" hypothesis is that education leads to greater control of resources that influence health and these resources cumulate with age. Therefore, the difference in health between those with more and less education will

[2] Whether the relationship between education and health (mortality) has changed over time (by birth cohort) has been a prominent research question, although few studies integrate this question with analyses that allow education to have different effects of the life cycle (i.e. age). Some recent studies in this area include Preston and Elo (1995), Manton et al. (1997), Lauderdale 2001, Meara et al. (2008), Cutler et al. (2011), and Montez et al. (2011).

grow with age. An alternative hypothesis in the sociological literature is the "age-as-leveler" hypothesis that posits that, as the burden of illness grows with age, education will have less effect, and that selective mortality will result in convergence of health at older ages between education groups (Dupre 2007).

In economics, there is almost no theoretical research on the nature of the relationship between education and health over the life cycle. To the best of our knowledge, Galama and Van Kippersluis (2015) is the only study to address the question. Galama and Van Kippersluis (2015) extend Grossman's (1972) model to derive hypotheses about the life cycle trajectory of health by socioeconomic status (not education per se). While predictions of the model are ambiguous, Galama and van Kipperslius show that under one set of assumptions, there is a predicted divergence of health (stock of health capital) between education groups in middle age and convergence at older ages. An important distinction between Galama and Van Kipperslius (2015) and our model is that we allow the productivity of health investments to differ by age and the effect of education on that productivity to differ by age. In contrast, predictions in Galama and Van Kipperslius (2015) stem from the earnings increase associated with education and not because of education's effect on the productivity of investments.[3]

A similar sparseness characterizes the empirical literature, as there have been relatively few studies, and none in the economics literature, that document the life cycle profile of the association between education and health. Here, we describe several of the most cited studies. House et al. (1994) and Ross and Wu (1996) were two of the first studies to examine the age-profile of the education-health gradient, but these studies used cross-sectional data and ignored cohort differences.

Mirowski and Ross (2008), Herd (2006) and Wilson et al. (2007) used longitudinal samples that followed individuals for six to 17 years to examine the age-profile of the education and health relationship. These studies focused on the effect of education on self-rated poor health or functional status. These studies generally reported evidence supportive of the "cumulative advantage" hypothesis—diverging age-profiles of health by education. Lynch (2003) examined the association between education and self-rated health by age and birth cohort using synthetic, birth cohorts derived from repeated cross-sections of the National Health Interview Survey (NHIS). He concluded that the results of the analysis are most consistent with the "cumulative advantage" hypothesis. Masters et al. (2012) used the linked NHIS-mortality files to examine the effect of education on mortality. Their results indicated that education is associated with a lower rate of death and that this effect is relatively constant over the life cycle—neither convergence nor divergence. More recently, Case and Deaton (2017) examined educational differences in death rates by age and birth cohort. Case and Deaton (2017) focused on deaths due to drugs, alcohol and suicide (i.e., "deaths of despair") and found a marked increase in in deaths from these causes over time (by birth cohort) among lower educated (< BA) persons in middle age. For low-educated whites, the increase in deaths of despair contributed to an increase in all-cause death rates for cohorts born between 1940 and 1970.

Three studies have used European cohorts to study the relationship between education and health over the life cycle. Bijwaard et al. (2015) used a Dutch cohort born between 1937 and 1941 to examine the effect of education on mortality between ages 55 and 75. They found that education exceeding the primary level decreased mortality, but mainly after age 65. Martikainen et al. (2007) used Finnish cohorts aged 30-59 between 1971 and 2000 to study the effect of education on 5-year mortality rates at ages 30-39, 40-49 and 50-59. Results of this study indicated that education had a beneficial effect on mortality at each age, but that the effect was largest at ages 30-39. There was inconsistent evidence as to whether the educational differences changed by birth cohort. Finally, Leopold and Leopold (2018) used German longitudinal data and found that education was associated with

[3] The Galama and Van Kipperslius (2015) prediction of an income-related divergence of health by education as people age is difficult to square with the evidence on earnings differentials by education and age. Education related earnings differentials grow early in the life cycle and become constant at around age 40 (Card and Lemieux 2001). As we show below, education-related differences in health are constant until around age 50 and then diverge.

better self-rated health at all ages between 30 and 80 and that the education effect grew with age, evidence supportive of the "cumulative advantage" hypothesis.

This brief review leads to three points. First, there has been little theoretical analysis of the relationship between education and health over the life cycle. Here, we provide a novel analysis of this relationship using a modified version of the Grossman (1972) model. We link this theoretical model directly to a flexible empirical regression model. Second, there is relatively little empirical research that examines the age profile of the relationship between education and health. Third, studies that focus on obtaining an estimate of the effect of education on health largely ignore the possibility that this relationship may change over the life cycle and by cohort with Case and Deaton (2017) being a notable exception with respect to cohort. We address these last two points and add to the existing empirical literature. We use several data sets to estimate the age-varying relationship between education and a broad set of health and mortality outcomes for different birth cohorts.

3. Theory

Grossman (1972) is the canonical theory of the demand for health and is the model used almost exclusively to study the effect of education on health. In the Grossman (1972) model, the key relationship is the first order condition for investment in health. A person chooses to invest in health to alter the stock of health, which provides utility because health itself is valued by the person, and because health influences income. Optimal investment occurs where the marginal benefit of investment equals the marginal cost of investment. A slightly revised version of the Grossman (1972) first order condition for investment in health at age t is:[4]

$$(1)\quad \frac{U_H}{\lambda} + Y_H = MC_{I(t-1)}(\delta_t + r) - \Delta MC_I$$

In equation (1), the following definitions are used:

$$(2)\quad U_H = \frac{\partial U}{\partial H}(1+r)^t,\ MC_{It} = \frac{p_t}{f_{It}},\ \Delta MC_I = MC_{I(t-1)} - MC_{It}$$

In equation (1), the optimal investment in health (I) at age t is chosen to equate the marginal benefit of investment to marginal cost. The marginal benefit includes the additional monetary value of utility (U_H / λ) and the greater income (Y_H) from a higher stock of health (H) that is produced by additional investment *(I)*. The marginal cost is the price of health investment (p) divided by the marginal product of investment in producing health (f_I) multiplied by the rate of interest (r) plus the depreciation rate on stock of health (δ_t) minus the change in MC_I (ΔMC_I) between ages t-1 and t.

It is important to recognize that the only exogenous factors that vary with age (t) in equation (1) are the depreciation rate and MC_I. Depreciation of health generally grows with age, although this may not be the case with mental health (distinct from cognition). To maintain the equilibrium in (1) with rising depreciation, the level of health stock must decline with age so that $\frac{\partial U}{\partial H}$ and $\frac{\partial Y}{\partial H}$ increase. Eventually, the health stock declines to a level

[4] Similar to Grossman (1972), the lifetime utility function depends on health (H) and consumption (Z): $U = U(H_0,...,H_t,Z_0,...,Z_t)$. The production function of health depends on investment (I) and depreciation (δ): $H(t+1) = f(I_t) - \delta(t)H(t)$, $f_{It} = \frac{\partial f}{\partial I_t}$. The budget constraint is: $\sum_t \frac{p_t I_t + Z_t}{(1+r)^t} = \sum_t \frac{Y_t(H_t)}{(1+r)^t}$. The price of consumption is 1 and income is exogenous and depends on health ($Y_H = \frac{\partial Y}{\partial H}$).

where death occurs (Grossman 1972). Investment in health is chosen to accomplish this change in the health stock. As Grossman (1972) shows, investment, for example, medical care, may increase or decrease with age (i.e., with increasing depreciation) depending on model parameters, although empirical evidence suggests that it increases.

3.a. An Extension—Age-varying Productivity of Investment

Notably, most studies assume that the marginal cost of investment (MC_I) is constant over time (with age), for example, because prices are assumed constant. However, this is unlikely to be the case and not because prices are changing, but because the productivity of investment (f_I) is likely to change. Up to now, we have used the generic term investment to refer to activities and goods that augment health, but investments in health take many forms, such as nutrition, exercise and medical care. It is also likely that the nature of investment changes with age as the nature of the depreciation of health (illness) changes. For example, exercise may be more important at younger ages than older ages because the depreciation of health at younger ages may be due to moderate, slowly evolving effects of aging more than abrupt changes in health due to the onset of illness. Conversely, medical care may be more important at older ages than younger ages because the depreciation of health that occurs at older ages may be due to serious illness and be more amenable to medical care than exercise. Accordingly, the productivity of investments is likely to change with age simply because the types of investments change, but also because of other reasons such as the fact that older persons are more experienced investors (made prior investments of same type).

If the productivity of investments changes with age, then the life cycle pattern of investment and health may differ significantly from that predicted based on a model with only an increasing rate of depreciation with age. For example, assume that the productivity of investment grows with age. If so, then this increase in productivity will offset some, perhaps all, of the higher cost of investment associated with a rising depreciation rate with age and investment will be higher, and the health stock will decline more slowly with age, than in the model with just an increasing rate of depreciation with age. Alternatively, if the productivity of investment decreases with age then the health stock will decrease more rapidly with age than predicted by a model with only depreciation rising with age. Of course, how the productivity of investment changes with age is not really known and it may change non-monotonically with age. The extended model also has implications for the life cycle patterns of, and correlation between, health care spending and health. Again, consider the case in which the productivity of investment increases with age, and is therefore negatively correlated with the depreciation rate. In this case, health care spending is more likely to increase, perhaps sharply, with age even as health deteriorates (i.e., when depreciation rate is high). In fact, empirically, there is a strong negative correlation between health and health care spending that changes with age.

3.b. The Effect of Education

One way to incorporate education into the model is to make the marginal product of investment a function of education, as in Grossman (1972). A common, and we think reasonable, assumption is that more education raises the productivity of investment. If so, and assuming other things the same (e.g., p, δ_t , r), then the marginal cost of investment is lower, investment is higher and the health stock is larger for those with more education.[5] This result will hold at all ages, and suggests that more educated people will be healthier at all ages and live longer than less educated people. However, whether the age-profiles of health for persons with more or less education diverge or converge over the life cycle has largely been ignored in the literature with the exception of Galama and van Kippersluis (2015).[6]

[5] There will also be an income effect associated with the lower marginal cost (shadow price). This may lead to more investment that would increase health. Education may also raise wages/income and cause an increase in investment.
[6] Galama (2017) and Galama and van Kippersluis (2018) extend the Grossman (1972) to address previous criticisms of the model. These extensions yield a richer set of predictions about the age profiles of investment and health than in Grossman

One reason the difference in health between those with more and less education may change over the life cycle is that education may have different effects on the productivity of investment by age. As noted, the nature of investment likely changes with age, and education may have different effects on the productivity of different types of investments in health, for example, raising the productivity of investments in medical care more than investments in exercise. If so, then the effect of education on the productivity of investment will change with age as the types of investment in health change with age. As a result, investment in health will differ by age and education, and the age-profile of health will differ by education.

The difference in the age-profile of health (and investment) by education will depend on the age-profile of the productivity of investment and how that productivity is affected by education. To understand how these factors will change over the life cycle would require information about the causes of the depreciation of health at various ages, the productivity of different types of investments associated with the types of health depreciation with age, and how education affects the productivity of different types of investments in health. Such information is not readily available. Regardless, the relationship between education and health is unlikely to be constant over the life cycle. Empirical analyses of the effect of education on health should take this into account and allow the effect of education on health to differ by age.

3.c. Interaction Between Depreciation and the Effect of Education

Depreciation of health is a fundamental biological fact and central to the Grossman (1972) model and in particular to the age-profile of investment and health. The extended model that allows for changing productivity of investments with age alters the predictions of that model, and it could even result in an increasing rate of investment *and* health over certain age ranges. It may make sense for a person to invest more at older ages and for their health to increase with age if the productivity of investment increases with age.

The effect of education on the productivity of health investment will interact with the changing productivity of investment and depreciation. Consider a case where the depreciation rate is constant (e.g., zero) over an age range. Under this assumption, and all else equal, health would remain unchanged for both low- and high-educated persons in the Grossman (1972) model. In the extended model health would remain unchanged for both types of persons unless the productivity of investment changed between these ages. But there is little reason to expect the productivity of investment to change because the depreciation rate is the same at both ages and the types of health shocks and investments to offset those shocks are the same. Thus, over this age range, education will have the same effect on investment and health and the difference in health between low and high educated persons will remain constant.

The upshot of this discussion is that the effect of education on health is likely to depend on the rate of depreciation because it is the change in depreciation that causes changes in investment and in the productivity of investments that provide scope for education to have an effect. This result has significant implications for empirical analyses of the relationship between education and health. At young ages when there is little to no depreciation, we may expect education to have little to no effect on health. Another example pertains to mental health, which does not change monotonically with age. Depreciation of mental health (i.e., mental illness) occurs at all ages and is often more pronounced at younger ages than older ages. In this case, the difference in mental health between low- and high-educated persons may be larger at younger than older ages.

3.d. Summary

(1972). However, like here and in the Grossman (1972) model, beyond the prediction that more educated people will be healthier and live longer than less educated people, the shape of the age-profile of health and the difference in the age-profile of health by education is ambiguous. Galama and van Kippersluis (2018) also analyze the effects of education on health only through the income effect of education.

To summarize, while economic theory has provided a clear hypothesis about the relationship between education and health, predicting that more educated persons will be healthier and live longer than less educated persons, a hypothesis about how the relationship between education and health changes over the life cycle has not been well developed. We extended the Grossman (1972) model to suggest that the age-profile of the education and health relationship is unlikely to be constant. Predictions about the life cycle profile of the relationship between education and health are ambiguous and underscore the need for empirical analysis to document a set of facts. We argued that at young ages when there is little depreciation of health, education is unlikely to have much of an effect on health and the health of low- and high-educated persons should change in a largely parallel fashion. At older ages, when depreciation is high and investments are high, education is likely to have a more pronounced effect and the health of more- and less-educated persons is likely to diverge.

4. Measuring the Effect of Education on Health over the Lifecycle

The theory discussed above provides a rationale for why the relationship between education and health may change over the life cycle. Here, we outline how to measure that effect empirically. We focus on the health production function that relates investment in health to the stock of health. The effect of education on health is through investment and therefore reflected in the production function. Any difference in health between more and less educated persons has to be because of differences in the productivity and quantity of investment. The health production function measures these differences.

One description of the health production function is as follows:

$$\begin{aligned}(3)\quad H_t &= H(1-\delta_0)...(1-\delta_{t-1}) + \alpha I_0(1-\delta_1)...(1-\delta_{t-1}) + \alpha I_1(1-\delta_2)...(1-\delta_{t-1}) + \\ &\quad ... + \alpha I_{t-2}(1-\delta_{t-1}) + \alpha I_{t-1}\end{aligned}$$

In equation (3), health at age t is a function of initial health (H), the quantity of all past investments in health (I_0, …, I_{t-1}), the effects (productivity) of those investments (α) and the depreciation rate in each period (δ_0, …, δ_{t-1}). Health is determined by a cumulative process in which investments in health augment the stock of health and depreciation erodes the stock.

Note that the increment to health from an investment, for example, medical care at a specific age, will differ for those at different ages because of depreciation.[7] The effect of an investment, α, may also differ by age because the productivity of an investment—the same investment and the same amount of investment—may be age-specific for biological reasons. For example, calorie and nutrient consumption at earlier ages may add more to health than at older ages because of differences in biological development at these ages, or for clinical reasons, for example, if effects of pharmacotherapy differ by age (Goldman and Lakdawalla 2005). This possibility is reflected in the following:

[7] To illustrate this, consider two people, one who is age 20 and one who is age 30. The health production function for these two individuals are given by:

$$\begin{aligned}H_{20} &= H(1-\delta_0)...(1-\delta_{19}) + \alpha I_0(1-\delta_1)...(1-\delta_{19}) + \alpha I_1(1-\delta_2)...(1-\delta_{19}) + \\ &\quad ... + \alpha I_{18}(1-\delta_{19}) + \alpha I_{19} \\ H_{30} &= H(1-\delta_0)...(1-\delta_{29}) + \alpha I_0(1-\delta_1)...(1-\delta_{29}) + \alpha I_1(1-\delta_2)...(1-\delta_{29}) + \\ &\quad ... + \alpha I_{19}(1-\delta_{20})...(1-\delta_{29}) + ... + \alpha I_{28}(1-\delta_{29}) + \alpha I_{29}\end{aligned}$$

As shown, the increase in health caused by an investment at any age (e.g., 19) will differ for those age 20 and age 30 because depreciation of health capital between ages 20 to 30 has eroded the value of previous investments (e.g., age 19) for those age 30 more than for those age 20.

$$H_t = H(1-\delta_0)...(1-\delta_{t-1}) + \alpha_0 I_0(1-\delta_1)...(1-\delta_{t-1}) + ... + \alpha_{(t-1)} I_{t-1}$$

$$(4) \qquad = H \prod_{j=k}^{t-1}(1-\delta_j) + \alpha_0 I_0 \prod_{j=k+1}^{t-1}(1-\delta_j) + ... + \alpha_{(t-1)} I_{t-1}$$

$$\prod_{j=k+1}^{t-1}(1-\delta_j) = (1-\delta_{k+1})...(1-\delta_{t-1})$$

The point to note about equation (4) is that the health production function is age-specific. Also, the effect of an investment, α, differs by age so the same investment made at, for example, ages 14 and 24, will have a different effect on health at each age. One empirical implication of the age-specificity of the health production function is that it may be inappropriate in empirical analyses to use samples of people of different ages without allowing the effects of (the same) investments to differ by age (for both reasons noted earlier).

The same approach applies to health at age t+1:

$$(5) \quad H_{t+1} = H \prod_{j=k}^{t}(1-\delta_j) + \alpha_0 I_0 \prod_{j=k+1}^{t}(1-\delta_j) + ... + \alpha_t I_t$$

And the difference in health between ages t+1 and t is:

$$(6) \quad \begin{aligned} & H_{t+1} - H_t = \alpha_t I_t - \delta_t D = \alpha_t I_t - \delta_t H_t \\ & D = H_t = \{H[\prod_{j=k}^{t-1}(1-\delta_j)] + \alpha_0 I_0[\prod_{j=k+1}^{t-1}(1-\delta_j)] + ... + \alpha_{t-2} I_{t-2}(1-\delta_{t-1}) + \alpha_{(t-1)} I_{t-1}\} \end{aligned}$$

The difference in health between two ages depends on the last period's investment ($\alpha_t I_t$) minus the extra depreciation (δ_t) that has occurred on past investments as a result of being one year older. Note that if depreciation is zero then the difference in health between ages is just the addition to health from last period investment.[8] It is also evident from equation (6) that, in the absence of investment, health declines with age. Additional investment slows any biological decline.

To integrate education into the production function, we assume, as we did in the theoretical model, that education influences health by increasing the productivity of investment. For example, those who are more educated may process information from health providers more efficiently and have better decision-making skills to implement treatment (Cutler and Lleras-Muney 2008). In addition, because the productivity of investment is higher, then the amount of investment will be higher for more- versus less-educated persons, as shown by the first-order condition of equation (1). To reflect this influence of education on investments in health, we can rewrite the health production function assuming, for simplicity, that the effect of education is zero up to age 18 and then constant after that age:

$$(7) \quad H_t = H \prod_{j=k}^{t-1}(1-\delta_j) + \alpha_0 I_0 \prod_{j=k+1}^{t-1}(1-\delta_j) + ... + \alpha_{19t}(E) I_{19}(E) \prod_{j=k+1}^{t-1}(1-\delta_j) + ... + \alpha_{(t-1)}(E) I_{t-1}(E)$$

Given the specification in equation (7), which makes the effect and quantity of investment depend on education, which is consistent with theory above, the effect of education on health at age t is:

---

[8] We note, however, that there is likely to be little to no investment when depreciation is zero (see the earlier theoretical discussion).

$$(8)\quad \frac{\partial H_t}{\partial E} = \sum_{k=19}^{t-2} [(\frac{\partial \alpha_k}{\partial E} I_k + \alpha_k \frac{\partial I_k}{\partial E}) \prod_{j=k+1}^{t-1} (1-\delta_j)\ ] + \frac{\partial \alpha_{t-1}}{\partial E} I_{t-1} + \alpha_{t-1} \frac{\partial I_{t-1}}{\partial E}$$

This equation shows that the effect of education on health is the sum of the effects of education on productivity and quantity of investment. It is the cumulative effect of education on health at that age. Estimates of equation (8) can be obtained from a regression model such as the following estimated on a sample of people age t:

$$(9)\quad H_{it} = \pi_0 + \pi_t EDUCATION_i + e_{it}$$

Note that there is a subscript on the coefficient of education indicating that the estimate is age specific. The coefficient on education in equation (9) is an estimate of the cumulative effect of education on health (i.e., equation 8) at a certain age. In many applications, an estimate of the effect of education is obtained from a regression model such as equation (11) using a sample of people with varied ages. In such analyses, the effect of education on health is restricted to be the same at every age. This restrictive specification yields an estimate of what may be thought of as an average of the cumulative effect of education on health at each age in the sample. If, at many ages in the sample, the cumulative effect of education on health is zero, then estimates from such models (a weighted average of the effect of health across many ages) will be close to zero and difficult to detect empirically.

The effect of education on health at age t+1, is:

$$(10)\qquad \frac{\partial H_{t+1}}{\partial E} = \sum_{k=19}^{t-1} \{[\frac{\partial \alpha_k}{\partial E} I_k + \frac{\partial I_k}{\partial E} \alpha_k] \prod_{j=k+1}^{t} (1-\delta_j\} + \frac{\partial \alpha_t}{\partial E} I_t + \frac{\partial I_t}{\partial E} \alpha_t$$

And using equations (8) and (10), the difference in the effect of education on health between age t+1 and age t is:

$$(11)\qquad \begin{aligned} \frac{\partial H_{t+1}}{\partial E} - \frac{\partial H_t}{\partial E} &= \frac{\partial \alpha_t}{\partial E} I_t + \frac{\partial I_t}{\partial E} \alpha_t - \delta_t \{ \sum_{k=19}^{t-2} [(\frac{\partial \alpha_k}{\partial E} I_k + \frac{\partial I_k}{\partial E} \alpha_k) \prod_{j=k+1}^{t-1} (1-\delta_j)] + \frac{\partial \alpha_{t-1}}{\partial E} I_{t-1} + \frac{\partial I_{t-1}}{\partial E} \alpha_{t-1} \} \\ &= \frac{\partial \alpha_t}{\partial E} I_t + \frac{\partial I_t}{\partial E} \alpha_t - \delta_t \frac{\partial H_t}{\partial E} \end{aligned}$$

As described by equation (11), the difference in the effect of education on health between two ages is the effect of education on health at age t (i.e., in the last period). It consists of the effect of education on the productivity of health investment, on the amount of health investment at age t, and on the initial (i.e., period t) stock of health at age t.[9]

An estimate of equation (11) can be obtained from the regression:

$$(12)\qquad H_{i(t+1)} - H_{it} = \tilde{\pi}_0 + \tilde{\pi}_{t+1} EDUCATION_i + (e_{i(t+1)} - e_{it})$$

Note that equation (12) is a first difference specification between two ages, but it still includes education, which is time invariant. This is because education modifies the effect and quantity of investment in every period (education has a time varying effect). This time-invariant variable does not drop out of the first-difference specification. This

[9] Note that, in the absence of differential investment in health by education at age t (last period), the difference in health by education will become smaller with age. Persons that are more educated will have a higher health stock due to past investments and depreciation will be greater for them causing a convergence of the age profiles of health between low- and high-educated persons. This result is due to the multiplicative form of depreciation and may not hold empirically.

is a feature of all human capital production functions, but is often ignored in empirical applications, particularly in health but also in other analyses (e.g., Dee 2007; French et al. 2010; Ozkan and Henderson 2011; Meraya et al. 2018; Austin et al. 2019). The parameter estimate from the first-difference model measures the effect of education on health *at* age t+1.[10]

One issue related to interpreting the coefficients on education in regression analyses such as those represented by equations (9) or (12) is that education may have an indirect effect on health because it raises wages (income), which would likely lead to greater investment and better health. However, these "indirect" effects are still caused by education and are conceptually part of the total effect of education on health.[11] In fact, this pathway is similar to the effect of education on the quantity of investment ($\frac{\partial I_t}{\partial E}$) shown in equation (11).

5. Empirical Approach
5.a. Research Design

We obtain estimates of the effect of education on health that are consistent with the conceptual model and health production function just described. To do so, we use a sample of persons all born in the same year, or a group of years (a birth cohort), that are followed over time as they age. To facilitate description of the empirical analysis, we assume that there is a sample of people followed for 40 years between ages 30 and 70 and that there are only two levels of education. With this sample, we estimate regression models, such as the following:

$$\text{(13)}\quad \begin{aligned} H_{ikt} &= \alpha_1 FEM_i + \alpha_2 BASE\_H_i + \sum_{k=30}^{70} \delta_k AGE_{ikt} + \sum_{k=30}^{70} \beta_k (AGE_{ikt} * EDUCATION_i) \\ &+ \sum_{k=30}^{70} \rho_k (AGE_{ikt} * BASE\_H_i) + \sum_{k=30}^{70} \sigma_k (AGE_{ikt} * FEM_i) + e_{ikt} \end{aligned}$$

In equation (13), health of person (cohort) *i* in year *t* who is age *k* depends on a dummy variable indicating person is a female, a variable measuring baseline (e.g., age 30) health, age dummy variables, interactions between age and education, interactions between female and age, and interactions between baseline health and age. Gender, baseline health and education do not vary with age. Basically, equation (13) estimates the mean of each outcome by age and education adjusted for different age profiles of health by gender and baseline health. In practice, we pool samples of persons within relatively narrow bands of birth year and control for common year effects in equation (13).

From equation (13), the effect of education on health at a specific age ($\tilde{\pi}_{k+1}$) is:

$$\text{(14)}\qquad \tilde{\pi}_{k+1} = \beta_{k+1} - \beta_k$$

[10] Another empirical issue is that many studies, and even those that allow the effect of education on health to differ by age, often impose a specific functional form, for example, linear or quadratic, on the modifying effect of age on the association between education and health (Lynch 2003; Masters et al. 2012). However, there is no reason to believe that age will have a moderating influence that is linear or quadratic.

[11] Altering education is not the only way to alter the quantity of investment and, thus, health. Investments in health could be changed directly, or through other means such as income transfers or information interventions. This point is relevant to the policy implications of the association between education and health.

Equation (14) measures the difference in the change in health between ages k+1 and k for those with more education versus those with less education. It is a difference-in-differences estimate and can be obtained from estimates of equation (13).[12]

Estimates of the effect of education on health from a regression model such as equation (13) may be biased if there are omitted variables correlated with education that also augment, in an age-specific way, the productivity or quantity of investments in health. Note that several confounding influences commonly noted in the literature on education and health, such as ability, time preference and early childhood influences (e.g., genetic, in utero environment) are often viewed as time-invariant. If the effects of these variables do not change with age, then these variables would not result in bias of estimates form equation (14), which is a difference-in-differences estimator that accounts for time-invariant effects. However, if these variables, like education, affect health in an age-specific way then their omission will bias estimates.

In this article, we address this issue only partly. When examining mortality (hazard rate) of a specific birth cohort, and as indicated in equation (13), we allow the effect of baseline health to differ by age and for the effect of gender to differ by age. We note here, that including these variables has little effect on estimates, which in the case of the interactions between baseline health and age is an important piece of evidence that our estimates may not be too confounded by age-varying factors. In addition, in an analysis of NLSY data, we allow several variables, such as mother's education, armed forces qualification test (AFQT), self-esteem and locus of control to have age-varying effects and results remain significant and consistent with other estimates. Nevertheless, we do not want to assert that we have obtained causal estimates. Our main contribution is to highlight the theoretical prediction that the relationship between education and health is unlikely to be constant over the life cycle, to describe the empirical implications of that prediction, and to estimate the age-profiles of health by education. In this sense, our analysis parallels the large, descriptive literature documenting education-related earnings differential by age and over time (see references earlier). The sparseness of research on educated-related health differences that considers time-varying effects of education over the life cycle underscores why our descriptive study is important.

However, it is worth commenting on the recent literature that attempts to estimate the causal effects of education on health. Most of these studies have exploited educational reforms that changed the compulsory school leaving age and an instrumental variables research design. However, these studies have generally estimated the cumulative effect of education on health using a regression model such as equation (9). These estimates may be misleading because education may have significant effects on health only at certain ages, and therefore, the cumulative effect of education may be small. This issue may be particularly important if samples include a large fraction of young persons because of the biological nature of the burden of disease that appears to be nonlinear in age—health spending and the incidence of disease grow exponentially with age (Solé-Auró et al. 2015).[13]

Consider the results in Clark and Royer (2013) and Davies et al. (2018). Both studies examined the effect of a 1972 educational reform in the UK on mortality. Clark and Royer (2013) reported estimates of the effect of education for persons between the ages of 20 and 44 and found no statistically significant effects of the reform on mortality, although estimates for those ages 40 to 44 were negative, marginally significant and indicated that reform was associated with a 5.3% reduction in death rates. Davies et al. (2018) used a similar statistical approach, but for an older sample with an average age of 53 and maximum age of 62. Davies et al. (2018) reported that the educational reform was associated with a significant, 42% decrease in death rates. The differences in estimates between Clark and Royer (2013) and Davies et al. (2018) may be due to the rapidly rising death rates after age 50 and the potentially greater role for education to affect health after that age (see Figure 1).

[12] Note that the standard errors of difference-in-differences estimates account for the covariance of estimates across ages.
[13] Also see: https://seer.cancer.gov/statfacts/html/all.html

Clark and Royer (2013) also reported estimates of the effect of a 1947 UK reform on mortality for several age groups. These estimates were not statistically different from zero except for a *positive* 3% estimate for persons ages 45 to 49.[14] However, estimates did differ by age group suggesting the possibility that education had different effects on the log odds of death at different ages. We can use equation (13) and estimates of the cumulative effect of education at different ages reported in Clark and Royer (2013) to obtain estimates of the effect of the 1947 reform on the log odds of death at ages 50-54, 55-59, 60-64 and 65-69. Respectively, these constructed estimates are: -0.013, -0.025, 0.005, and 0.020. These are estimates of the effect of educational reform at those ages (not cumulative estimates) and these are not particularly small estimates, for example, a 2.5% (0.025) decrease in mortality at ages 55-59. If, for argument's sake, we assume that these estimates are true, then it highlights how the interpretation and implications of analyses that use our approach can differ markedly from most previous studies.

5.b. Data

To estimate the model given by equation (13), we use three data sets. The first is the Chicago Heart Association Detection Project in Industry (CHA). It is a novel, longitudinal data set with linkages of baseline survey data to Vital Statistics Death file (2012) and the National Death Index (2004 to 2008). The CHA was a milestone epidemiological study that assessed cardiovascular health risk factors (e.g., blood pressure and serum cholesterol) and collected demographic information for approximately 40,000 employees in Chicago, IL between 1967 and 1973 (Stamler et al. 1993 and Stamler et al. 1999). These data provide us with a unique opportunity to investigate the education-health gradient on mortality because individuals were followed for up to 40 years after the baseline survey.

We focused on two birth cohorts from the CHA sample: those born between 1927 and 1934 and those born between 1935 and 1942. We limited the sample to those aged 30 to 40 at the time of the baseline survey and who lived until at least age 40.[15] The 1927-34 cohort sample size was 4097 and the 1935-42 cohort sample size was 4814. In the 1927-34 cohort, 16.6 percent had less than high school degree, 36.9 percent had a high school degree or equivalent, 16.5 percent had some college, and 30.04 percent had at least a college degree. Analogous figures for education for the 1935-42 cohort are 13.8 percent, 33.1 percent, 19.9 percent and 33.1 percent, respectively. Other survey information collected at baseline that was used in the analysis includes gender, race, height and cardiovascular health. Cardiovascular health is a composite measure of health status and incorporated health risk factors including BMI, blood pressure, whether an individual has diabetes, serum cholesterol, and smoking status. It is classified into favorable, elevated risk, high risk and very high risk (see Allen et al. 2017).

The information on mortality for all CHA participants was updated until 2012. We examined the education-mortality gradient from age 40 to 75 for the 1927-1934 cohort and 40 to 70 for the 1935-1942 cohort. To do this, we created a panel data set that started from age 40 up until age 75 (age 70 for the 1935-1942 cohort) or the age of death, if he/she died before 75 (70). Mortality is measured by a dichotomous variable that equals one in the year/age a person died, zero otherwise. Descriptive statistics for the CHA sample are provided in Table 1.

The second source of information is the National Health Interview Survey (NHIS). We used the NHIS surveys linked to the National Death Index to study the effect of education on mortality (Masters et al. 2012). In this case, we focused on respondents aged 40 in the 1986 to 1990 NHIS surveys. The first NHIS survey to be linked to the mortality data was 1985. The data allow us to follow two cohorts of persons born between 1945 and

[14] It is not implausible that education would increase mortality. For example, if depreciation of health is zero and there is no investment in health there would still be an effect of education that works through income. If education raises income and consumption, then the risk of death and actual death may rise with the greater economic activity (e.g., driving). See Snyder and Evans (2006) for evidence consistent with this hypothesis.

[15] Approximately 0.17% of the 1927-34 cohort died before age 40 and approximately 0.62% of the 1935-42 cohort died before age 40.

1949 and 1950 to 1954 until 2011 to identify age of death.[16] We observe this cohort between the ages of 40 (35) and 64 (59).

For the analysis of the effect of education on mortality using both the CHA and the NHIS, we model the (discrete time) hazard rate of death and obtained differences in the hazard rate of death between ages and the difference-in-differences in the hazard rate of death between ages by education. In the main set of models, we controlled for gender, baseline health (composite measure in CHA and self-reported health in NHIS) and interactions between these variables and age. The unit of analysis in these samples is the person-year (age). Each person contributes an observation for each year they are alive.

We also used NHIS surveys from 1976 to 2016.[17] The NHIS is a cross sectional survey and we use successive surveys to form synthetic, birth cohorts (Lynch 2003). We focus on several cohorts: 1930-34, 1935-39, 1940-44, 1945-49, 1950-54, 1955-59 and 1960-64. In each year of the survey we assign birth cohort based on year of birth. In 1976, the oldest birth cohort is approximately 46 years old and the youngest cohort is approximately 17 years old. In 2016, the oldest cohort is around 82 years old and the youngest cohort is around 50 years old. The analysis is limited to the period of life between ages 30 and 70, although not all birth cohorts are observed for the entire 40-year period. For example, someone born in 1930 is observed between ages 46 (in 1976) and 70 (2000). A person born in 1950 is observed between ages 30 (in 1980) and 66 (2016). In order to reduce the likelihood of confounding factors within-cohort such as immigration and differences in the quality and composition of education, we restrict the sample to non-Hispanic whites. Note that the synthetic cohorts are affected by selective mortality as those who die before a given year (age) are not observed in the data. We return to this point in the discussion of results.

The NHIS surveys provide information on educational attainment and health. All health-related measures are self-reported. We constructed four educational categories: less than high school, high school, some college and Bachelor's degree or more. For health, we used the following measures: a dichotomous indicator of poor health (self-rated health equal to 4 or 5); a dichotomous indicator of the presence of any limitation; dichotomous indicators for whether a person reported having hypertension (diagnosis) or diabetes (diagnosis); and, finally, a dichotomous indicator for whether a person was widowed. The last outcome merits some comment. It is not a measure of a person's own health, but there are two reasons for it to be sensitive to a person's education. First, there is some degree of positive assertive mating, so a person's education reflects the spouse's education. Second, a person's education may have a direct effect on spousal health for the same reasons that it may affect their own health. We note here that results using this variable are highly consistent with analyses of mortality using the CHA and NHIS cohorts. The use of several NHIS measures is motivated by the fact that the causes and treatments of diseases such as diabetes, hypertension and activity limitations differ. If education had very different effects on these measures of morbidity it may point to the mechanisms through which education works.

Using these NHIS, synthetic cohorts, we obtained estimates of the effect of education on the difference in health between contiguous ages and the difference-in-differences between education groups, as described above. In this analysis, however, we only include interactions between gender and age because it is a synthetic cohort and we do not know baseline characteristics. We do this separately for each cohort. Descriptive statistics for this sample are presented in Table 2.

The third data set we use is the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). This is a longitudinal data set that has followed approximately 12,000 persons between the ages of 14 and 21 in

[16] Restricted access NHIS data with linked National Death Index data are available from 1985-2015 but our data come from IPUMS (Blewett et al. 2016), which have mortality linkage from 1986-2009.

[17] We use data from 1982 on for measures of self-reported health due to a change in the scaling that was not amenable to harmonization.

1979 to the present. Attrition is remarkably low for such a long follow-up period. [18] The data provide extensive information about demographic and socioeconomic characteristics, family background, educational attainment, cognitive (e.g., ASVAB test scores) and non-cognitive (e.g., self-esteem) attributes, work, marital history, fertility and many other aspects of a person's life. It is a well-documented and widely used data set.[19] We make use several of these characteristics. Specifically, we construct measures of race/ethnicity, family background, self-esteem, locus of control (Rotter scale), and AFQT.

For our purposes, the NLSY79 is useful because, for each respondent, it collected information on educational attainment and health at ages 40 and 50, as measured by the following: self-reported health, depression (CESD instrument), and physical and mental health (SF-12 instrument sub-scales). Therefore, we can use these data to analyze the effect of education on health and the change in health between ages 40 and 50. The information on health was collected at the time of the first interview after reaching age 40 and age 50. So, there is some small variation in age at the time health was measured. We adjust for these age differences in all models. Descriptive statistics for these data are presented in Table 3.

6. Results

6.a. Chicago Heart Association Cohort

We begin the discussion of results with those from analyses that use the Chicago Heart Association (CHA) cohorts: 1927-1934 and 1935-1942. The basic approach is to estimate equation (13), which we repeat here but now showing that we include year dummy variables because we combine several single year birth cohorts into one larger cohort:

$$\begin{aligned} H_{ikt} &= \alpha_1 FEM_i + \alpha_2 BASE\_H_i + \sum_{k=30}^{70} \delta_k AGE_{ikt} + \sum_{k=30}^{70} \beta_k (AGE_{ikt} * EDUCATION_i) \\ &+ \sum_{k=30}^{70} \rho_k (AGE_{ikt} * BASE\_H_i) + \sum_{k=30}^{70} \sigma_k (AGE_{ikt} * FEM_i) + e_{ikt} \end{aligned} \tag{13}$$

Estimates of equation (13), which are obtained using Ordinary Least Squares (OLS) methods, are shown in Figure 2.[20] The dependent variable is the hazard rate of death for each cohort and the figure is divided into panels based on birth cohort and whether one-year or five-year age categories are used. The graphs using the five-year age categories are smoother and reveal more clearly the age-profile of the hazard rate of death.

For the 1927-1934 cohort, the hazard rate of death rises gradually from ages 40 to ages 55-59 except for persons with a Bachelor's degree or more—their hazard rate remains relatively constant. There is a slight widening of the gap between the hazard rates of death by education. At ages 55-59, the hazard rate of death starts to increase for each education group and at ages 60-64, there is a noticeable increase in the hazard rate of death for the lowest education group. For other education groups, the hazard continues a gradual rise. For the 1935-1942 cohort, the hazard rate of death remains relatively flat (constant) from ages 40-44 to ages 55-59 for all education groups. At ages 55-59, the hazard rate of death among the lowest education group diverges from the hazard rates of death for the other education groups, although the divergence in the rate of change of the hazard is observed only until ages 60-64.

[18] In round 20 of the survey (26 years on), overall attrition was 22.5% (NLS Handbook 2005, Chapter 3).
[19] https://www.nlsinfo.org/content/cohorts/nlsy79/using-and-understanding-the-data/nlsy79-documentation
[20] In the Appendix (Appendix Figures 1 and 2), we present estimates of equation (13) that exclude interactions between baseline health and age and gender and age. Results are similar to those presented in the text.

Difference-in-difference estimates are shown in Figure 3, which has six panels showing results for several comparisons for each birth cohort: BA or more versus some college; BA or more versus high school; BA or more versus less than high school; some college versus high school; some college versus less than high school; and high school versus less than high school. The difference-in-differences (DiD) estimates are estimates of equation (14): the difference in the effect of education on health between two ages.[21]

To reduce year-to-year noise in the results, we present DiD estimates of the five-year change in health. The main result revealed by Figure 3 is that there are no significant differences in the change in the hazard of death by age across educational groups until age 70. At that age (70 to 75), those with less than a high school degree have a higher hazard rate of death than those in other education groups, although estimates are only statistically significant for the differences between those with a BA and those with a high school degree or less. While those with more education have a lower hazard rate of death than those with less education at all ages, the effect of education is virtually zero over the life cycle, at least until age 70. This result suggests that education is not modifying significantly the productivity of health investments (or increasing health investments) related to mortality between ages 40 to 70. This result can also be illustrated by constructing survival rates. In Table 4 (also see Appendix Table 1), we provide estimates of the difference in survival (product of one minus previous hazard rates) until various ages by education. For each cohort, education has little effect on survival to age 50. However, education does begin to have an effect on survival to age 60, although it remains relatively small. The change in survival rate between ages 50 and 60 and the difference by education is analogous to the DiD estimates in Figure 3. For example, in the 1927-34 cohort, those with a BA are 6.7 percentage points (8%) more likely to live until age 60 than those with less than a high school degree. By age 70, the survival advantage of those with a BA or more versus those with less than a high school degree for this cohort is 13.4 percentage points (19.1%). For the 1935-42 cohort, the analogous figures are 4.5 percentage points (5.2%) and 12.2 percentage points (16.3%). The widening of the difference in survival rates by education between ages 60 and 70 is reflecting the fact that education is having a larger effect at those ages than at previous ages.

Overall, the results from the analysis of the effect of education on mortality (the hazard rate of death) show that education has distinctive age-specific effects. Education starts to have a particularly important effect after age 60. The main education-related difference is between those with less than a high school degree and higher education levels. Finally, there was no noticeable difference in the effect of education on mortality across the 1927-34 and 1935-42 cohorts.

6.b. NHIS 1945-49 Cohort

As noted, we linked the NHIS data to the National Death Index and, using these data, we followed two cohorts of people born between 1945 and 1949 and 1950 to 1954 until 2009 or the age they died. Education was measured in the baseline surveys between 1986 and 1989. Estimates of equation (13) were obtained using the hazard rate as a dependent variable. Figure 4 presents estimates for this cohort.[22]

As shown in Figure 4, the hazard rates of death of the different education groups are relatively constant up to ages 50-54 (note the different age categories on X axis for the two cohorts). After that age, the hazard rate of death starts to grow markedly faster for those with less than a high school education. Figure 5 presents difference-in-differences estimates for the NHIS cohorts. There are six panels showing the six comparisons that are possible between the four education groups. We again note that these are estimates of the effect of education on the hazard rate at each age, or between ages when we use more than a one-year difference. To reduce variation related to sample size, we present estimates of the five-year change in health. There is one notable finding revealed in

[21] The standard errors of the DiD estimates account for covariances between estimates and are obtained by imposing linear restrictions on estimates from equation (13). These are identical to equation (14).
[22] In the Appendix (Appendix Figures 3 and 4), we present estimates that exclude interactions between baseline health and age and gender and age. Results are similar to those presented in the text.

Figure 5 and it reflects the divergence of the hazard rates between those with less than a high school degree and those in other education categories in Figure 4. Education appears to have a particularly important effect on the hazard rate of death from the mid-50s until age 64, but only in comparisons between those having less than a high school degree and other categories. There seems to be a threshold effect of education, as differences in estimates related to comparisons among those with a high school degree or more are small and not statistically significant.

In Table 5 (and Appendix Table 2), we provide estimates of the difference in predicted survival until various ages by education. As was the case for the CHA cohorts, education has very little effect on survival to age 50. However, education does have an effect on survival to age 60 (64), and the change in survival grows with education, but most of the effect is between those with a less than a high school degree and the other education categories. For example, those who have Bachelor's degree or more have a probability of surviving to age 60 that is 3.4 percentage points (4%) and 8 percentage points (9.5%) greater than those with less than a high school degree. There is a clear divergence in the hazard rate of death between those with less than a high school degree and those with greater educational attainment, but among those with a high school degree or more, education does not have much effect.

Combined, estimates of the effect of education on health from the CHA and NHIS cohorts reveal a consistent pattern. Those with less than a high school degree have a higher rate of death that manifests when people are approximately 60 and that remains through ages 70 to 75. While there are some differences between other education groups with respect to mortality (see Tables 4 and 5), the differences among groups with more than a high school degree are relatively smaller (see Hayward et al. 2015 for similar finding). This suggests that education has little effect on mortality until age 60 and particularly large effects after that age. Finally, we do not see much difference in the effect of education across the four birth cohorts, which differs from some previous studies (Preston and Elo 1995, Meara et al. 2008; Montez et al. 2011).

6.c. NHIS Synthetic Cohort

The next set of analyses make use of the NHIS synthetic cohorts and self-reported measures of health (morbidity). The analyses and presentation of results are similar to those previously presented. The first outcome we discuss is self-rated poor health. Figure 6 shows the means of this variable by age and education for several cohorts spanning 1930-34 and 1960-64. We omit the means for those with some college education to aid the visual presentation.

While a bit noisy, Figure 6 shows that there were significant differences in self-rated poor health at age 30 by education (a level difference that is eliminated in difference-in-differences estimate), particularly between those with less than a high school degree compared to other groups. This initial difference in health was not as pronounced in the case of mortality. Also evident in Figure 6 is a growing divergence in self-rated poor health between education groups as people age. Most striking is the divergence in self-rated poor health between those with less than a high school degree and those with a Bachelor's degree or more. Another point to note about Figure 6 is that there are some substantial cohort differences. At any age, the distribution of estimates by education category reflects the differences in poor health at that age by birth cohort. For example, at age 55-59 among those with less than a high school degree, the mean of poor health ranges from approximately 32% to 40% for cohorts born between 1930-34 and 1950-54. And consistent with the findings of Case and Deaton (2017), the health of more recent cohorts of low-educated persons is worse than earlier cohorts

In Figure 7, we show similar results, but for birth cohorts collapsed into two groups 15-year groups (1930-44, 1945-59), which makes the figures less noisy. While there is some variation across birth cohorts in the age-education profiles of self-rated poor health, the general pattern is the same across cohorts. Figure 7, however, reveals clearly the divergence in self-rated poor health by education at around age 40 and that this divergence became more pronounced in recent birth cohorts.

Figure 8 presents difference-in-differences estimates for birth cohorts collapsed into two groups. We include results from the less detailed cohort groupings in the text to highlight the basic patterns we found. Estimates for the more detailed cohort groupings in the Appendix. We note in the text if there are any differences that arise using the more detailed cohort groupings. Estimates in Figure 8 show that at each age between ages 35-39 and 60-64, education has a positive effect on health with the largest effects for those with a Bachelor's degree or more versus those with less than a high school degree. Between these ages, having a Bachelor's degree appears to lower the probability of poor health by almost 10 percentage points per five-year period. This is a large relative effect. In fact, in all comparisons between those with a less than high school degree and other educational categories, estimates indicate a positive effect of education on self-rated health at the specified age and almost all estimates are statistically significant.[23] These findings suggest that education is have a moderating effect on health at each age and causing age profiles of health to diverge. This finding contrasts with the for mortality that showed effects of education only at much older ages. In general, differences in self-reported health between contiguous educational categories are smaller and usually not significant, although the difference between those with a Bachelor's degree or more and other education groups are significant.

Finally, there is evidence that selective mortality may be playing a dominant role at older ages. The evidence is found in two ways. First, among the oldest cohorts who are not observed in the data until ages 40 to 45, educational differences in self-reported poor health are smaller and not as statistically significant. One explanation for this is that selective mortality has left a cohort of low-educated persons that is unusually healthy. The second piece of evidence is the convergence of the age-health profiles by education at older ages within cohorts. We can make a back-of-the-envelope calculation using the predicted survival rates in Tables 4 and 5 to gauge the potential importance of selective mortality. Estimates in Tables 4 suggest that by age 70, 13% more of the low-educated (LTHS) cohort has died compared to the high-educated (BA) cohort. If that "missing" share of low-educated people had a rate of poor health of 0.5, which is plausible given the mean of those living and the fact that this group died, then adding them back into the average would raise the average of the complete cohort by 0.065 percentage points. This is a large adjustment and suggests that selective mortality can have significant effects on estimates of associations between education and morbidity. Education may have large effects that are obscured by not accounting for selective mortality. Few studies looking at the effect of education on health that use measures of morbidity consider this possibility.

Results for self-reported activity limitations are presented in Figures 9 and 10 (Appendix Figures 6 and 7). The pattern of estimates for activity limitations are very similar to those for poor health. At each age between ages 45-49 and 60-64, education has an effect and the effects tend to be larger the larger the educational difference. Again, effect sizes are relatively large, for example, the Bachelor's degree versus a less than high school degree comparison suggests approximately a five-percentage-point reduction in the probability of reporting any limitations per five-year period change for those with a Bachelor's degree. We also observe the same evidence related to selective mortality, and some evidence that the effects of education have become larger among more recent birth cohorts.

Results for hypertension are shown in Figures 11 and 12 (Appendix Figures 8 and 9). There is a substantial age-gradient in diagnosed hypertension, as the prevalence of hypertension rises noticeably with age. It is not clear in Figure 11 whether there are significant differences in the age-profile of hypertension by education. Figure 12 (Appendix Figure 8) confirms that education is not having an effect at these ages. Just to be clear, what this implies is that education is not modifying the productivity or quantity of investment in health related to hypertension during these ages.

[23] We examine five outcomes for each age and cohort (i.e., five potentially non-independent estimates). Therefore, multiple testing bias is not likely to be a major problem. In addition, in many cases where estimates are statistically significant, the p-value is 0.01 or less, which would remain statistically significant after applying a Holm-Bonferroni correction.

In Figures 13 and 14 (Appendix Figures 10 and 11), we present estimates for diabetes (diagnosed). In this case, there is some divergence in the prevalence of diabetes by age by education. At age 45-49, the prevalence of diabetes grows significantly faster with age for those with less than a high school degree than those in other educational categories. For this outcome, the birth cohort differences are more pronounced than in other outcomes. The primary difference between birth cohorts is the worsening of health among recent cohorts of persons with less than a high school degree, which echoes the findings of case and Deaton (2017). Difference-in-differences estimates are shown in Figure 14. Here, we see more evidence of a threshold effect. Having a less than high school degree is associated with a greater prevalence of diabetes between ages 45-49 and ages 55-59 with the effect being larger for larger differences in education. Effect sizes are also relatively large, for example, at ages 50-54, having a Bachelor's degree or more is associated with an approximately four percentage point lower probability of diabetes. For other educational differences, there is little evidence that education has a positive effect on diabetes prevalence. Finally, we again observe evidence consistent with an important selective mortality effect.

The last outcome for the NHIS synthetic cohorts we discuss is the likelihood of being widowed. Figure 15 (Appendix Figure 12) shows the age profile by education groups of this variable. At around age 50, the likelihood of being widowed increases with age noticeably faster for those with a high school degree or less relative to those with a Bachelor's degree or more. The figure is surprisingly similar to the earlier figures on the hazard rates of death by age and education for the CHA and NHIS cohorts. It is also notable that this effect is present despite the fact that marriage rates are lower among the less educated relative to those with greater education attainment (results not shown). Figure 16 (Appendix figure 13) shows difference-in-differences estimates for this outcome. Estimates indicate significant differences between those with a Bachelor's degree or more versus other education categories. Starting in the mid-50s, the probability of being widowed is significantly higher (approximately two percentage points) for lower educated groups than for the group with a Bachelor's degree or more. The "mortality" advantage among the more educated groups remains through age 70.

6.d. NLSY79 Cohort

Estimates of the effect of education on health using the NLSY79 cohort are presented in Tables 6 to 9.[24] We use four measures of health that are available in NLSY79 data and each table presents results from one of the four outcomes. The format of each table is the same. In each table, we present estimates of the effect of education on health at age 40, at age 50 and on the difference in health between ages 40 and 50. For each age, estimates from four regression model specifications are shown. The first specification includes controls for age (as noted above, there is little variation), gender and race/ethnicity. Estimates from this specification are closest to those for the CHA and NHIS analyses. The second specification adds controls for family background (mother's education, whether household had magazines, newspapers or library card while the respondent was growing up, and family structure). The third specification adds locus of control scale and self-esteem scale measured in 1980 and 1981 when the sample was quite young (i.e., mostly before completing education). And the final specification adds the AFQT scores. The purpose of estimating these different models is to assess the amount of confounding of the estimate of the effect of education on health from these covariates. Theoretically, all these baseline factors may have an independent effect on the productivity of investments in health that are similar to that of education. Importantly, they are also allowed to have an age-specific effect like education.

In Table 6, the dependent variable is the physical health score from the SF-12 instrument. The mean of this variable at age 40 is approximately 50 and the mean at age 50 is approximately 45, which is consistent with declining health with age (higher SF-12 score indicates better health). As indicated by estimates in the first four columns, at age 40, education is significantly associated with better physical health. For example, having a

[24] For this analysis, we did not estimate equation (13) because we observed people at only two ages. The estimates in Tables 6-8 are equivalent to previously presented difference-in-differences estimates, but are from a model of first differences of age 50 health minus age 40 health (i.e., equation 14).

Bachelor's degree versus not finishing high school (the omitted group) is associated with a 3 to 4 point increase in the SF-12 physical health score, which is an increase of approximately a third of a standard deviation. Having a high school degree instead of not finishing high school is associated with a significant 1.5 to 2 point increase in the SF-12 physical health score. While the addition of covariates reduces magnitudes of the estimates by approximately 25% to 50%, all estimates remain statistically significant. Estimates of the effect of education on physical health at age 50 are similar, but substantially larger (e.g., 50%) than the corresponding estimates at age 40. Estimates of the effect of education on health between ages 40 and 50, which are shown in the last four columns of Table 6, are also significant, but smaller than estimates at either age 40 or 50.

The smaller estimates of the effect of education on physical health between ages 40 and 50 are consistent with the conceptual model. These estimates measure the effect of education during a 10-year period, whereas estimates of the effect of education on health at age 40, or age 50, measure the cumulative effect. Consider the estimate of the effect of education on physical health between ages 40 and 50 associated with a Bachelor's degree or more of 2.22 points (last column). If we assume that this 10-year effect is constant between ages 20 and 50, then we would expect the cumulative effect at age 40 to be approximately 4.44 points assuming that education was completed around age 20. Compared to the actual estimate of 3.18 points, the 4.44 point estimate is just outside its confidence interval. A similar extrapolation to simulate the age 50 cumulative effect would yield an estimate of 6.66 points, which is also just outside the confidence interval of the corresponding actual estimate of 5.45 points. While we do not want to place too much weight on this extrapolation, it does suggest that the effect of education on physical health is likely not constant across ages. At both age 40 and age 50, the simulated estimate was above the actual estimate suggesting that the effect of education on health is larger at older ages than younger ages. This observation also holds true for other education categories. These results, albeit based on back-of-the-envelope estimates, are consistent with the depreciation of health between ages 40 and 50 and the conceptual model that has education modifying the effect of investments in health that are made to offset depreciation. Finally, note that if we had combined the age 40 and age 50 samples, estimates of the effect of education on health would have been a weighted average of the effects at each age.

Estimates of the effect of education on mental health at ages 40 and 50 are shown in Table 7. Although it is common to assume that health generally declines with age, the SF-12 measure of mental health in our data remained similar between ages 40 and 50. Given no depreciation of mental health with age, theory predicts that there may be little change in investment in health during these ages and little scope for education to affect mental health. Our estimates are largely consistent with this hypothesis. At age 40, education is significantly associated with better mental health, but effect sizes are relatively small. For example, a person with a Bachelor's degree or more has a SF-12 mental health score that is 1.2 percentage points (approximately 12% of a standard deviation) higher than a person with a less than high school degree. At age 50, we see similar estimates of the association between education and mental health, although somewhat larger than at age 40. Estimates of the effect of education on the change in mental health between ages 40 and 50 are small and most are not statistically significant. There is also not much of a gradient in the relationship between education and the change in mental health between ages 40 and 50, although having a Bachelor's degree or more does appear to be associated with better mental health during this 10-year period.

Again, it is worth noting that the relatively small effects of education on mental health between ages 40 and 50 are consistent with the conceptual model and the lack of depreciation (investment) in mental health between ages 40 and 50. Additionally, a simulation exercise similar to the one previously described suggests that for mental health, the effect of education is relatively constant across ages. This too is expected if depreciation in mental health and, thus, investment in mental health, is not increasing with age.

Table 8 presents estimates of the effect of education on the CESD scale of depression. Higher CESD scores indicate higher risks for clinical depression and in our sample the mean score increases slightly from age 40 to age 50. Estimates of the effect of education on the CESD score follow a very similar pattern to those for the SF-12 mental health score. At ages 40 and 50, education is associated with better mental health. For example, those with a Bachelor's degree or more have a CESD score that is 1.25 (25% of a standard deviation) and 1.68 (33% of a standard deviation) lower than a person with less than a high school degree at ages 40 and 50, respectively. However, the effect of education on depression between ages 40 and 50 is small and not statistically significant, and there is little evidence of a gradient in the association between education and health. For this outcome, the exercise that simulates estimates of the cumulative effect of education on health at age 40 and age 50 suggests that education may have larger effects on depression at younger ages.

The final measure of health we used in the NLSY79 analysis is an indicator of self-reported poor health (fair or poor). Estimates of the effect of education on this outcome are shown in Table 9. Education has similar effects on self-reported poor health at ages 40 and 50, although in relative terms estimates at age 50 are smaller than those at age 40. Despite of the depreciation in this measure of health with age (which would imply investment based on the conceptual model), education is not statistically significantly associated with changes in self-reported poor health from age 40 to age 50. These findings differ from those for the SF-12 physical health measure reported earlier and suggest that education has a limited role in moderating investments between these ages. While the exact cause of discrepancy is unclear to us, we note that self-reported poor health is arguably the least well validated health measure used in the NLSY79 analysis.

7. Conclusion

In this article, we extended the Grossman (1972) theoretical model to allow the productivity of investments in health to differ by age and the effect of education on that productivity to differ by age. We used the model to analyze the effect of education on health over the life cycle. The main conclusion from this model is that it is unlikely that the relationship between education and health will be constant over the life cycle. We argued that the effect of education is likely to grow with age as depreciation of health increases, the quantity of investments in health grow and the types of investment in health change allowing more scope for education to affect health. This prediction has important implications for empirical analyses of the effect of education on health. It suggests that such analyses should be age specific and that the effect of education on health should be allowed to change with age. This possibility has largely been overlooked in many previous analyses of the effect of education on health. As a result of this oversight, previous studies have produced estimates that are best viewed as an average of the cumulative effect of education on health over the range of ages in the sample. However, at many ages, the effect of education on health is likely zero, or close to zero, particularly at younger ages because depreciation of health is zero or close to zero. Therefore, previous studies are likely to conclude that education has little effect on health, particularly if the sample consists of younger persons when, in fact, education may have substantial effects, but only at some ages.

We also presented an extensive empirical analysis that, while largely descriptive, documents, arguably more thoroughly than any previous study, the association between education and health over the life cycle. Here, we focused on the human capital production function of health to highlight its usefulness in terms of empirically measuring the effect of education on health over the life cycle. Our formulation of the problem provides a way to estimate the effect of health at every age over the lifecycle. To our knowledge, we are the first to propose and conduct such an analysis. In addition, we highlight the possible confounding of the effect of education on health due to using samples that include a wide range of birth cohorts. The significant increase in educational attainment over time directly implies that the composition of groups defined by fixed educational categories has shifted over time.

Results of our analysis suggest that education does have a beneficial effect on health and mortality, but that those benefits do not manifest until older ages. For mortality, divergence in the age-profile of health occurs at around age 60. The divergence occurs at earlier ages, for example age 45, for morbidity. These findings are consistent with the conjecture that education has an impact on health that works through investment and that education has a more pronounced effect on the productivity of investments at older ages when there is more investment. Results from the NLSY analysis that show that education has a different association with measures of physical and mental health at ages 40 to 50 are also consistent with the conjecture that the effects of education are likely to be found when depreciation is large and when investment in health are being made. Between ages 40 and 50, mental health in the NLSY cohort was largely unchanged (i.e., little depreciation) and education was not significantly related to mental health.

We find that most of the beneficial effect of education stems from obtaining a high school degree or more. It is the health and mortality of lowest education group—those with less than a high school degree—that diverges most from the health and mortality of other education groups. Among those with a high school degree or more, differences in education have modest effects on health and mortality. We also found suggestive evidence that selective mortality may cause educational differences in morbidity to narrow at older ages. Not accounting for this potential effect, as most analyses do, can lead to incorrect conclusions related to whether education affects morbidity (e.g., hospitalization) even when using samples of older people. Finally, we find that the effects of education on health have tended to become larger for more recent birth cohorts, although this was not a uniform finding, for example, it was not the case for mortality.

References

Albouy, Valerie and Lequien, Laurent, 2009. Does compulsory education lower mortality?, *Journal of Health Economics*, **28**, issue 1, p. 155-168.

Allen NB, Zhao L, Liu L, Daviglus M, Liu K, Fries J, Shih T, Garside DB,Vu TH, Stamler J, Lloyd-Jones DM. 2017. Favorable Cardiovascular Health, Compression of Morbidity and Healthcare Costs: 40-Year Follow-up of the Chicago Heart Association Detection Project in Industry. Circulation, May 2017.

Austin, Wes, Garth Heutel, Daniel Kreisman, School bus emissions, student health and academic performance, Economics of Education Review, Volume 70, 2019, Pages 109-126

Autor, D.H., Katz, L.F., Kearney, M.S.: Trends in U.S. wage inequality: revising the revisionists. Rev. Econ. Stat. 90(2), 300–323 (2008)

Card, David, and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men?" Quarterly Journal of Economics 116 (May 2001), 705–746

Case, Anne and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." Brookings Papers on Economic Activity. Spring397-476.

Clark, D. O. & G. L. Maddox. 1992. "Racial and Social Correlates of Age-Related Changes in Functioning." The Journals of Gerontology: Social Sciences. 47: S222–S232

Clark, Damon, and Heather Royer. 2013. "The Effect of Education on Adult Mortality and Health: Evidence from Britain." American Economic Review 103 (6): 2087–2120.

Cutler, David M., and Adriana Lleras-Muney. 2008. "Education and Health: Evaluating Theories and Evidence." In Making Americans Healthier: Social and Economic Policy as Health Policy, edited by Robert F. Schoeni, James S. House, George A. Kaplan, and Harold Pollack, 29–60. New York: Russell Sage Foundation.

Cutler, David M., Fabian Lange, Ellen Meara, Seth Richards-Shubik, and Christopher J. Ruhm. 2011. "Rising Educational Gradients in Mortality: The Role of Behavioral Risk Factors." Journal of Health Economics 30(6):1174–87.

Dee, Thomas S. "Teachers and the Gender Gaps in Student Achievement." Journal of Human Resources 42, 3 (Summer 2007): 528-554

Davies, N, Dickson, MR, Smith, GD, van den Berg, GJ & Windmeijer, F, 2018, 'The causal effects of education on health outcomes in the UK Biobank'. *Nature Human Behaviour*, vol 2., pp. 117-125

Dupre, M.E., 2007. Educational differences in age-related patterns of disease: reconsidering the cumulative disadvantage and age-as-leveler hypotheses. J Health Soc Behav. 48(1):1-15.

French, M. T., Norton, E. C., Fang, H. and Maclean, J. C. (2010), Alcohol consumption and body weight. Health Econ., 19: 814-832. doi:10.1002/hec.1521

Galama, Titus, lleras-Muney, Adriana, and Hans van Kipperslius. 2018. "The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence", Oxford Research Encyclopedia of Economics and Finance, doi:10.1093/acrefore/9780190625979.013.7

Goldman, D. & Lakdawalla, D. (2005). A Theory of Health Disparities and Medical Technology. Contributions in Economic Analysis & Policy, 4(1) doi:10.2202/1538-0645.1395

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

Grossman, Michel. 2006. "Education and Nonmarket Outcomes." In Handbook of the Economics of Education, Vol. 1, edited by Eric Hanushek and Finis Welch, 577–633. Amsterdam: North-Holland.

Grossman, Michael. 2015. "The Relationship between Health and Schooling: What's New?" Nordic Journal of Health Economics 3, 1:7-17

Hayward, Mark D., Robert A. Hummer and Isaac Sasson. 2015. "Trends and Group Differences in the Association between Educational Attainment and U.S. Adult Mortality: Implications for Understanding Education's Causal Influence." Social Science & Medicine 127: 8-18. PMCID: PMC4324094

Herd, P. 2006. "Do Functional Health Inequalities Decrease in Old Age? Educational Status and Functional Decline Among the 1931-1941 Birth Cohort." Research on Aging. 28: 375–392

House, James S., James M. Lepkowski, Ann M. Kinney, Richard P. Mero, Ronald C. Kessler, and A. Regula Herzog. 1994. "The Social Stratification of Aging and Health." Journal of Health and Social Behavior 35:213-34

Katz, Lawrence F., and Kevin M. Murphy, "Changes in Relative Wages, 1963–87: Supply and Demand Factors," Quarterly Journal of Economics 107 (February 1992), 35–78

Kemptner, Daniel & Jürges, Hendrik & Reinhold, Steffen. (2011). Changes in Compulsory Schooling and the Causal Effect of Education on Health. Journal of health economics. 30. 340-54

Lauderdale, Diane S. 2001. "Education and Survival: Birth Cohort, Period, and Age Effects." Demography 38(4):551–61.

Leopold, Liliya and Leopold, Thomas. 2018. Education and Health across Lives and Cohorts: A Study of Cumulative (Dis)advantage and Its Rising Importance in Germany. iJournal of Health and Social Behavior. 59(1): 94-112

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." Review of Economic Studies 72 (1): 189–221.

Lynch, Scott M. Cohort and Life-Course Patterns in the Relationship between Education and Health: A Hierarchical Approach. Demography. 2003

Martikainen, Pekka, Jenni Blomgren, and Tapani Valkonen. 2007. "Change in the Total and Independent Effects of Education and Occupational Social Class on Mortality: Analyses of All Finnish Men and Women in the Period 1971–2000." Journal of Epidemiology and Community Health 61(6):499–505.

Masters, Ryan K., Robert A. Hummer, and Daniel A. Powers. 2012. "Educational Differences in U.S. Adult Mortality: A Cohort Perspective." American Sociological Review 77(4):548–72.

Meara, Ellen R., Seth Richards, and David M. Cutler. 2008. “The Gap Gets Bigger: Changes in Mortality and Life Expectancy, by Education, 1981–2000.” Health Affairs 27(2):350–60

Meghir, Costas, Mårten Palme, and Emilia Simeonova. 2018. "Education and Mortality: Evidence from a Social Experiment." *American Economic Journal: Applied Economics*, 10 (2): 234-56.

Meraya AM, Dwibedi N, Tan X, Innes K, Mitra S, Sambamoorthi U. The dynamic relationships between economic status and health measures among working-age adults in the United States. *Health Econ*. 2018;27(8):1160–1174. doi:10.1002/hec.3664

Mirowsky, J. & C. E. Ross. 2008. “Education and Self-Rated Health: Cumulative Advantage and Its Rising Importance.” Research on Aging. 30: 93–122.

Montez, Jennifer Karas, Robert A. Hummer, Mark D. Hayward, Hyeyoung Woo, and Richard G. Rogers. 2011. “Trends in the Educational Gradient of U.S. Adult Mortality from 1986 through 2006 by Race, Gender, and Age Group.” Research on Aging 33(2):145–71.

Montez, Jennifer Karas, Robert A. Hummer, and Mark D. Hayward. 2012. “Educational Attainment and Adult Mortality in the United States: A Systematic Assessment of Functional Form.” Demography 49(1):315–36.

Ozkan Eren, Daniel J. Henderson, Are we wasting our children's time by giving them more homework?, Economics of Education Review, Volume 30, Issue 5, 2011, Pages 950-961.

Preston, Samuel H. and Irma T. Elo. 1995. “Are Educational Differentials in Adult Mortality Increasing in the United States?” Journal of Aging and Health 7(4):476–96

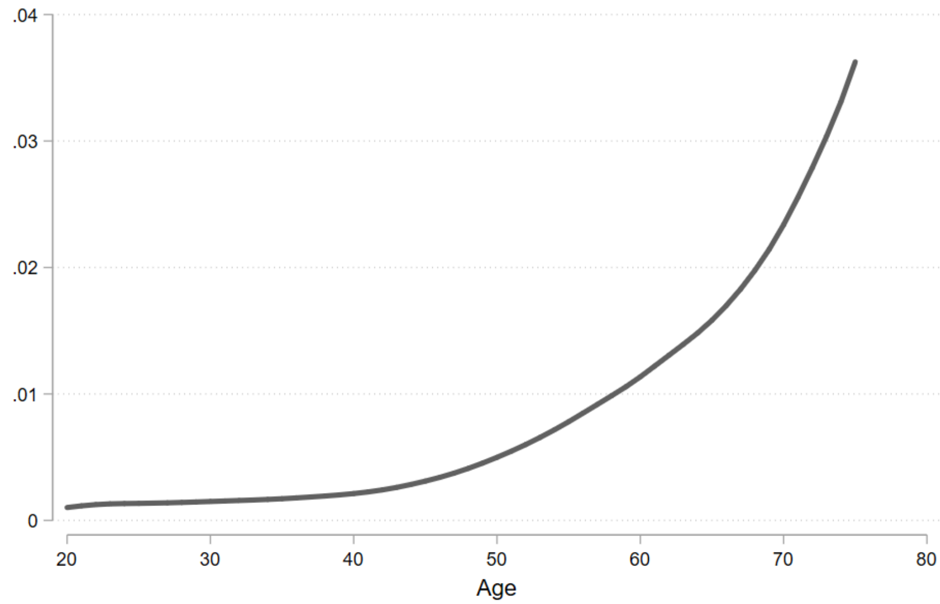
Ross, C. E. & C.-L. Wu. 1996. “Education, Age, and the Cumulative Advantage in Health.” Journal of Health & Social Behavior. 37: 104–120.

Snyder, Stephen E. and William N. Evans. "The Effect Of Income On Mortality: Evidence From The Social Security Notch," Review of Economics and Statistics, 2006, v88(3,Aug), 482-495

A Solé-Auró, PC Michaud, M Hurd, E Crimmins 2015. “Disease incidence and mortality among older Americans and Europeans.” Demography 52:593-611.

Willson, Andrea E.; Shuey, Kim M.; & Elder, Glen H., Jr. (2007). “Cumulative Advantage Processes as Mechanisms of Inequality in Life Course Health.” *American Journal of Sociology, 112(6),* 1886-1924.

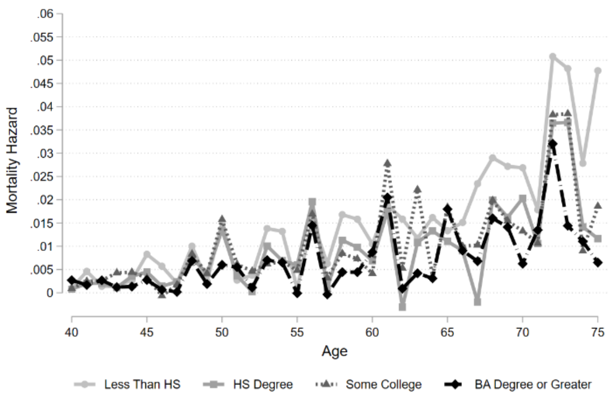
Figure 1. Male Death Rate



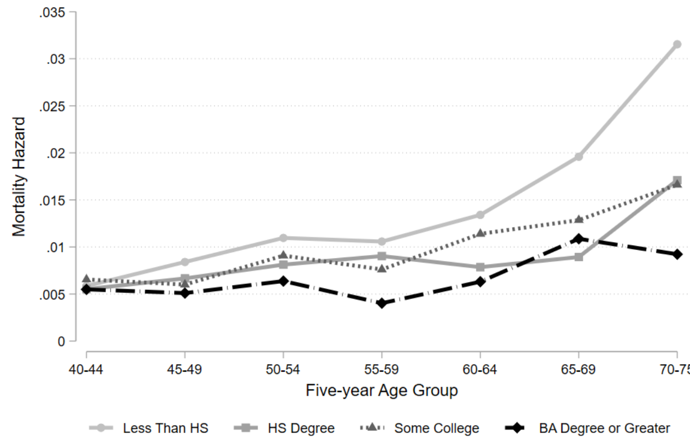
Notes: Data are from the 2014 period life table at the Social Security Administration (https://www.ssa.gov/oact/STATS/table4c6.html).

Figure 2. Predicted Hazard Rate of Death by Education and Age – CHA Cohorts

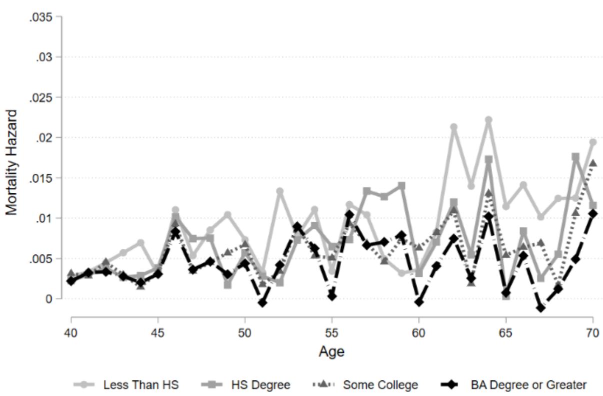
a. 1927-1934 Birth Cohort



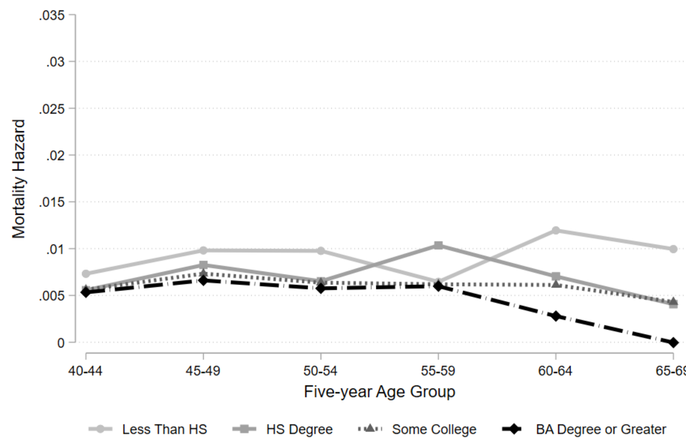
b. 1927-1934 Birth Cohort



c. 1935-1942 Birth Cohort



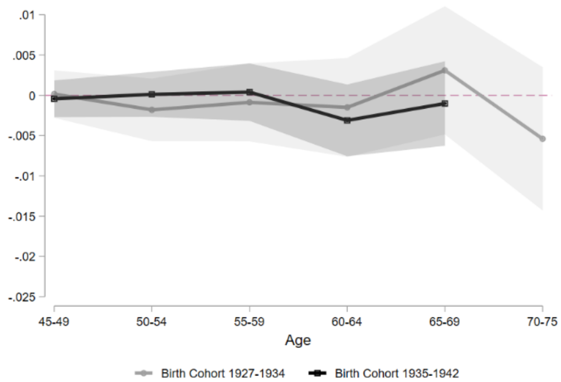
d. 1935-1942 Birth Cohort



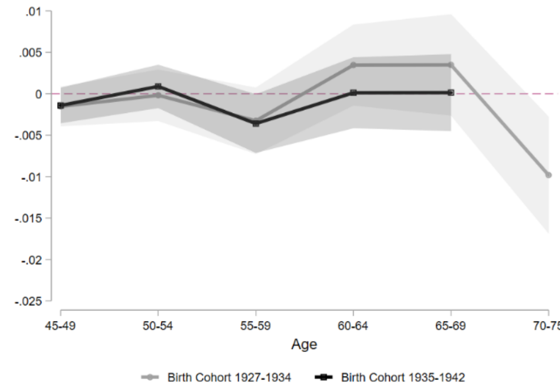
Notes: Figures present single-year and 5-year hazard rates of death by age and education for indicated birth cohorts. Predicted hazard rates of death are obtained from an OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, race dummy variables, height, baseline health, the interaction between female and age dummy variables, and the interaction between baseline health (composite measure) and age dummy variables. Predicted values are calculated at the mean of covariates (female, year, race, height, baseline health, interaction terms between female and age, and interaction terms between baseline health and age).

Figure 3. Difference-in-Differences in the Hazard Rate of Death by Education and Age – CHA Cohorts

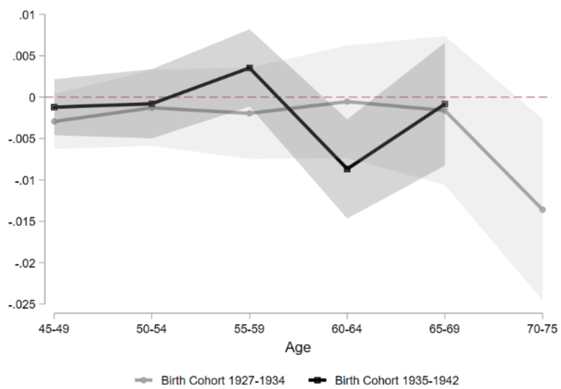
a. BA Degree or Greater – Some College



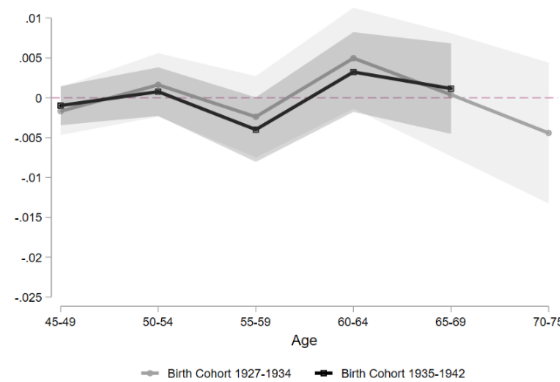
b. BA Degree or Greater – HS Degree



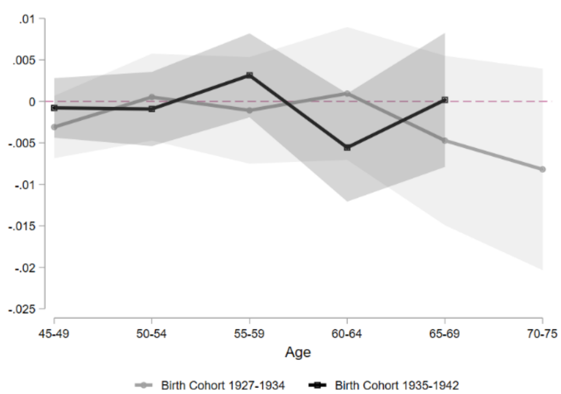
c. BA Degree or Greater – Less Than HS Degree



d. Some College – HS Degree



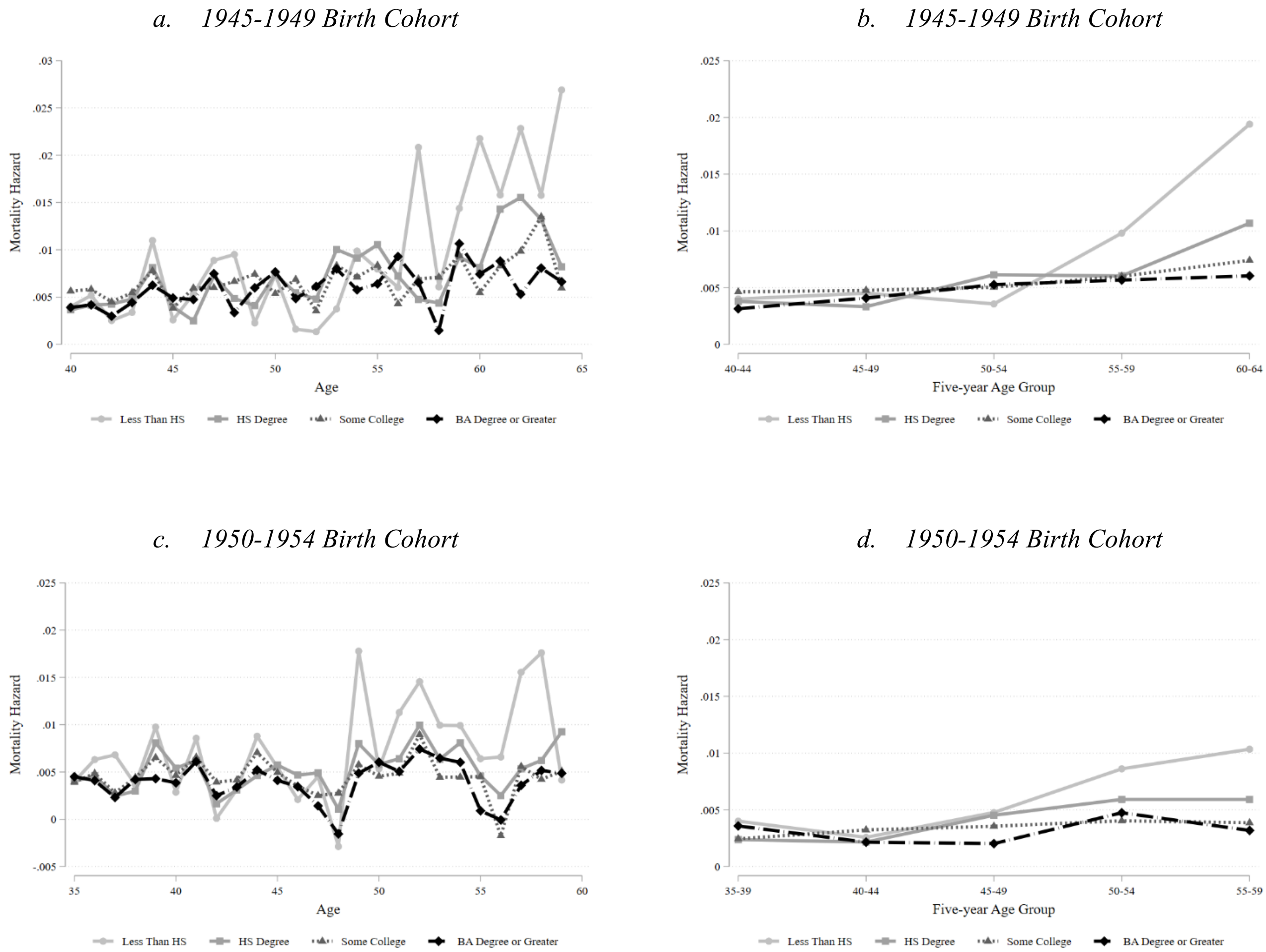
e. Some College – Less Than HS Degree



f. HS Degree – Less Than HS Degree

Notes: Figures present differences in the changes in hazard rate of death by age between education groups for CHA 1927-1934 birth cohort and 1935-1942 birth cohort. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 2.

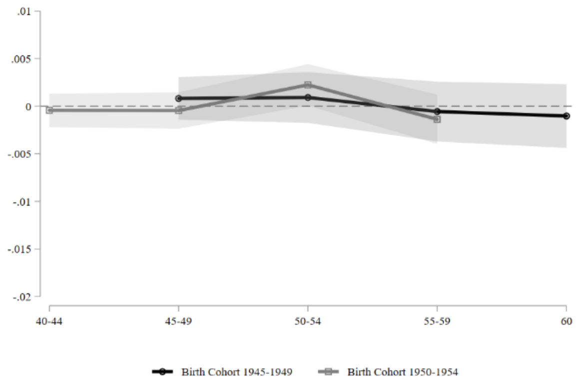
Figure 4. Predicted Hazard Rate of Death by Education and Age – NHIS Cohorts



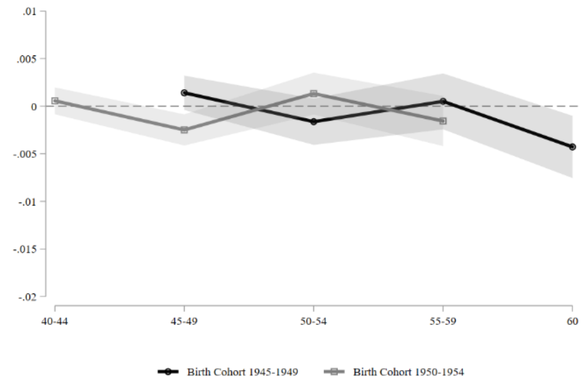
Notes: Figures present single-year and 5-year hazard rates of death by age and education for non-Hispanic white NHIS respondents in indicated birth cohorts who were initially surveyed between 1986 and 1990. Predicted hazard rates of death are obtained from an OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, baseline self-reported health, the interaction between female and age dummy variables, and the interaction between baseline self-reported health and age dummy variables. Predicted values are calculated at the mean of covariates (female, survey year, baseline self-reported health, interaction terms between female and age, and interaction terms between baseline self-reported health and age).

Figure 5. Difference-in-Differences in the Hazard Rate of Death by Education and Age - NHIS Cohorts

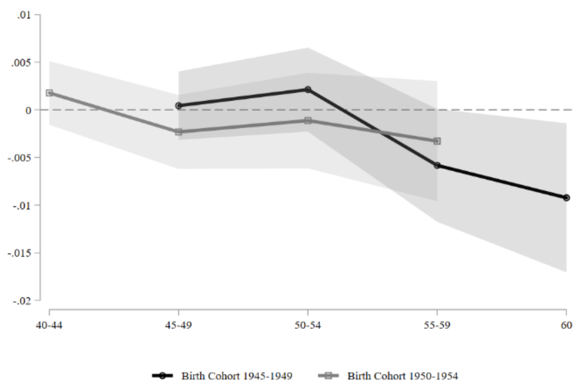
a. BA Degree or Greater – Some College



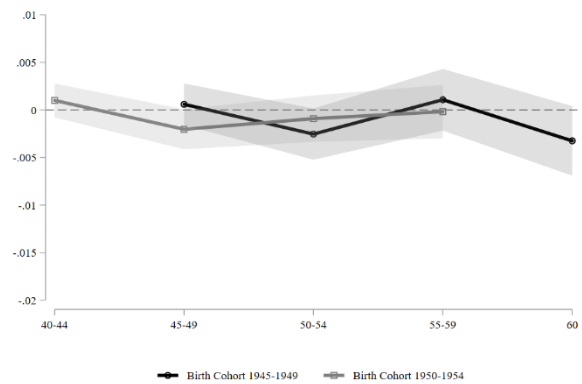
b. BA Degree or Greater – HS Degree



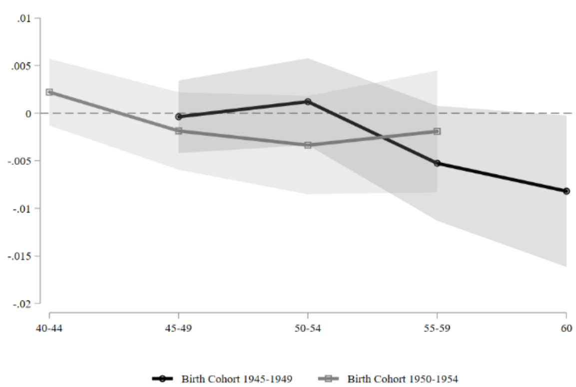
c. BA Degree or Greater – Less than HS Degree



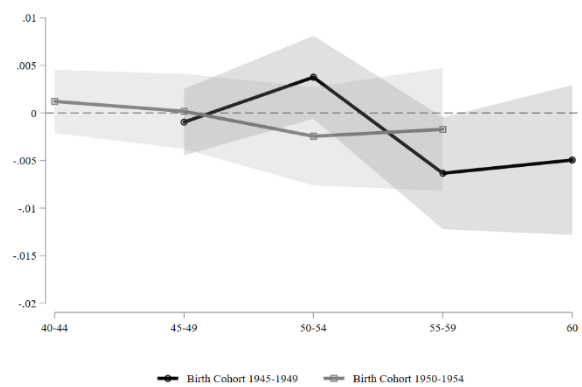
d. Some College – HS Degree



e. Some College – Less Than HS Degree

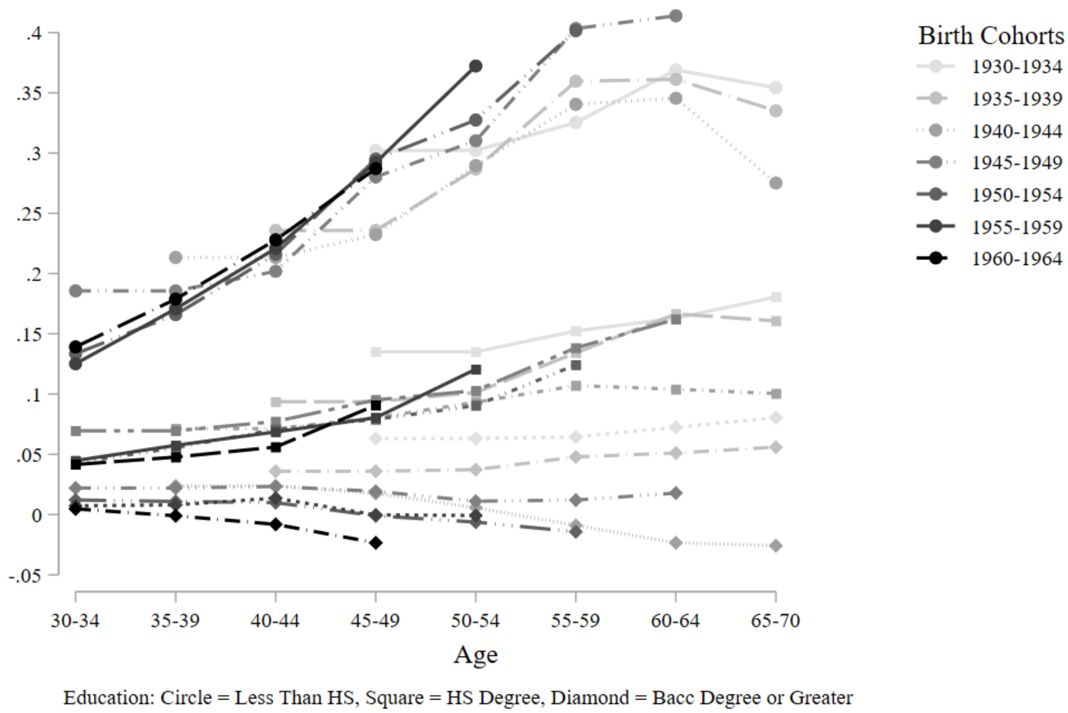


f. HS Degree – Less than HS Degree



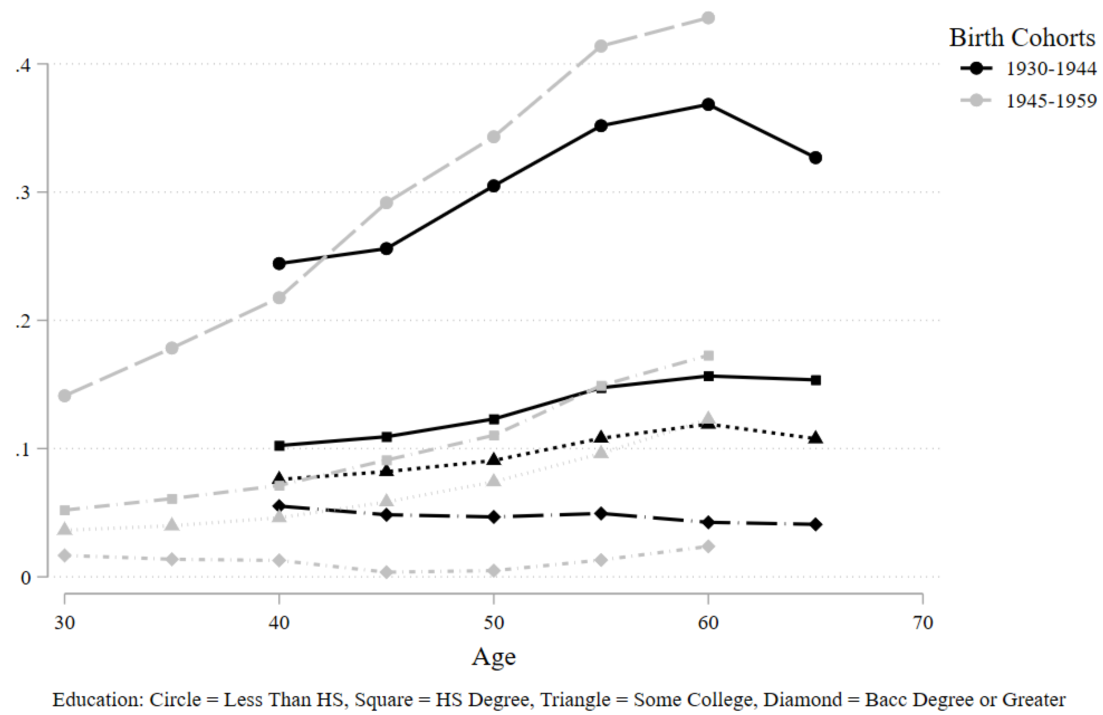
Notes: Figures present differences in the changes in hazard rate of death by age between education groups for NHIS 1945-49 birth cohort and 1950-54 birth cohort. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 4.

Figure 6. Proportion of Persons Reporting Poor Health by Five-Year Age Group and Education Level Among NHIS Respondents Born Between 1930 and 1964 (Five-Year Birth Cohorts)



Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1964 and surveyed between 1982 and 2016. To aid in interpretation, "Some College" cohorts are excluded from figure. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

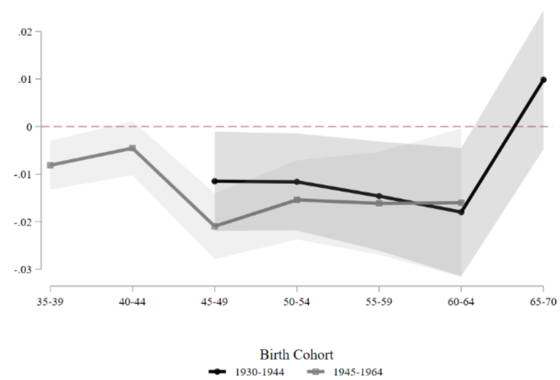
Figure 7. Proportion of Persons Reporting Poor Health by Five-Year Age Group and Education Level Among NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)



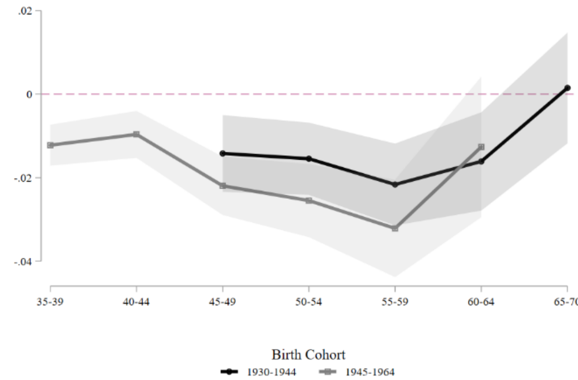
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1959 and surveyed between 1982 and 2016. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Figure 8. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Poor Health by Age for NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)

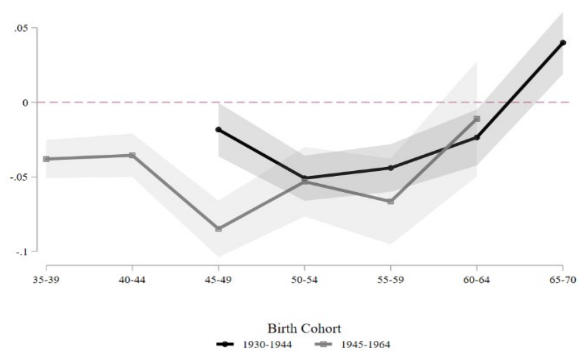
a: BA Degree or Greater – Some College



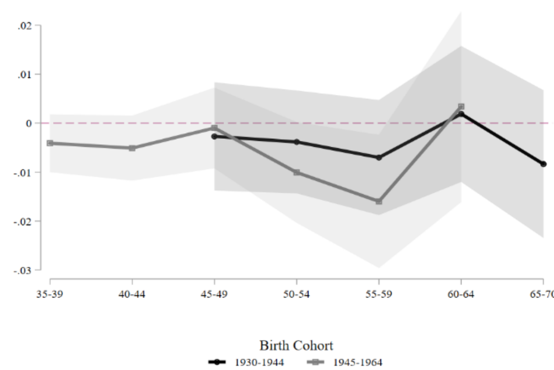
b: BA Degree or Greater – HS Degree



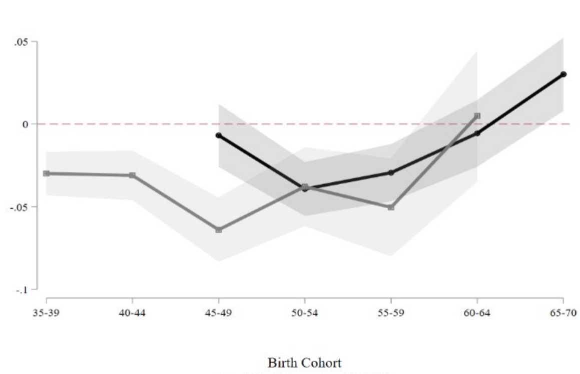
c: BA Degree or Greater – Less Than HS Degree



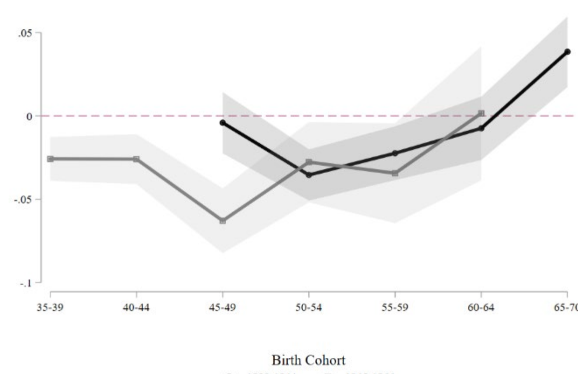
d: Some College – HS Degree



e: Some College – Less Than HS Degree

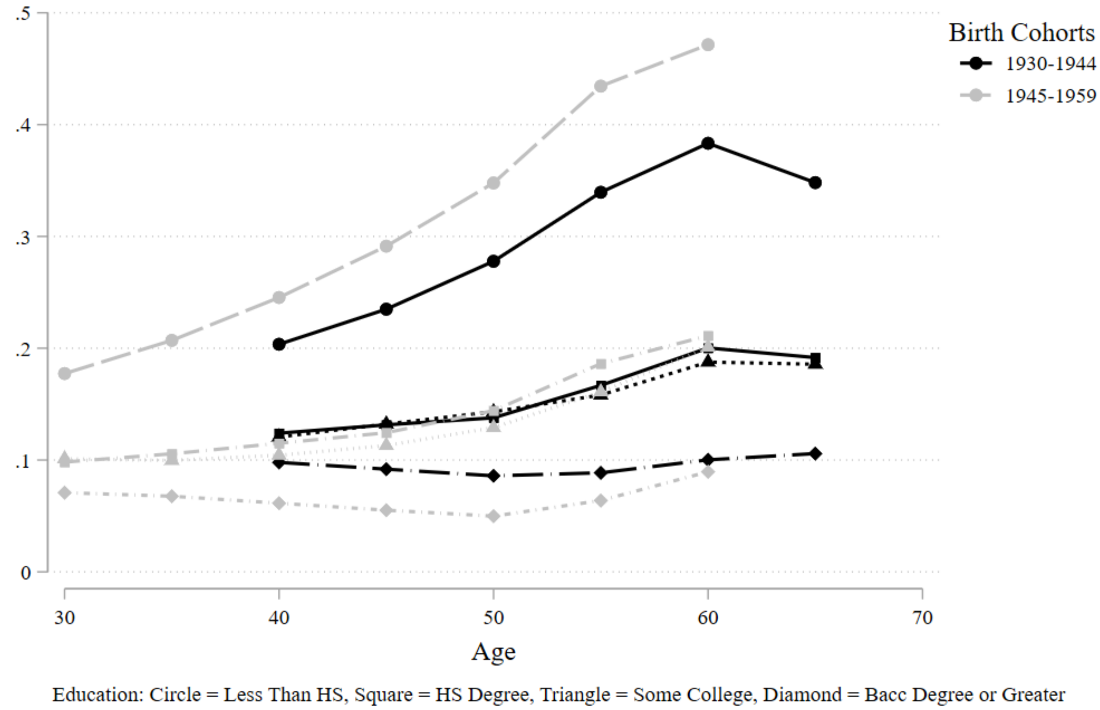


f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for or non-Hispanic, white NHIS respondents born between 1930 and 1959 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 7.

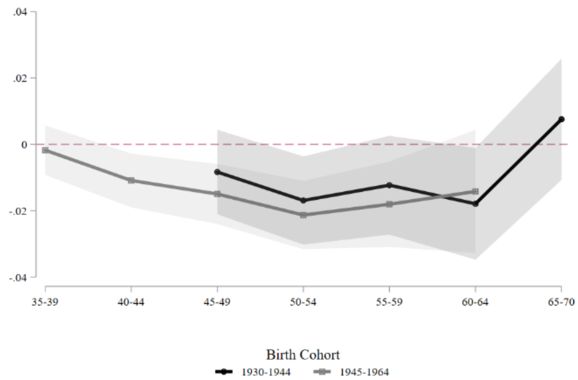
Figure 9. Proportion of Persons Reporting Any Limitation by Five-Year Age Group and Education Level Among NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)



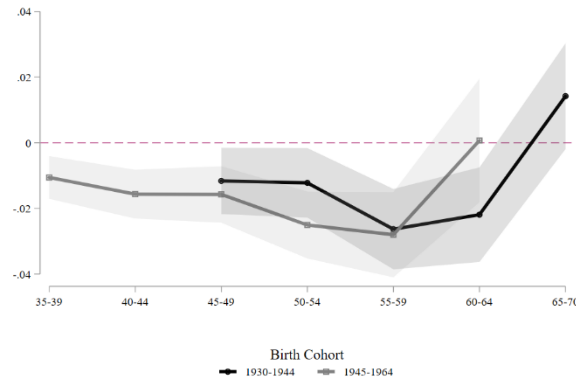
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1959 and surveyed between 1982 and 2016. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Figure 10. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Any Limitation by Age for NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)

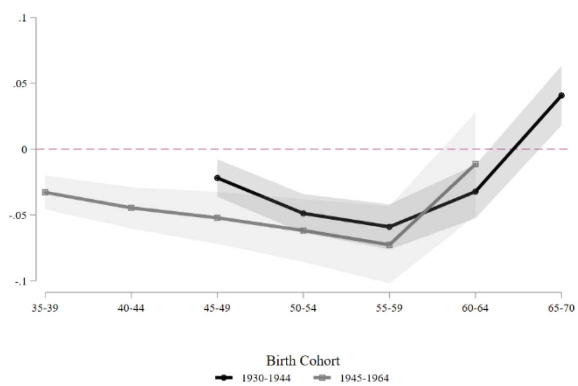
a: BA Degree or Greater – Some College



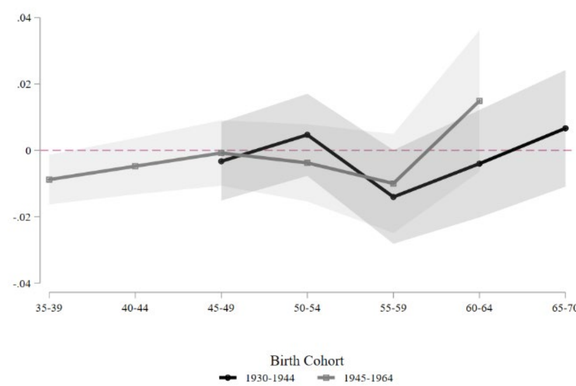
b: BA Degree or Greater – HS Degree



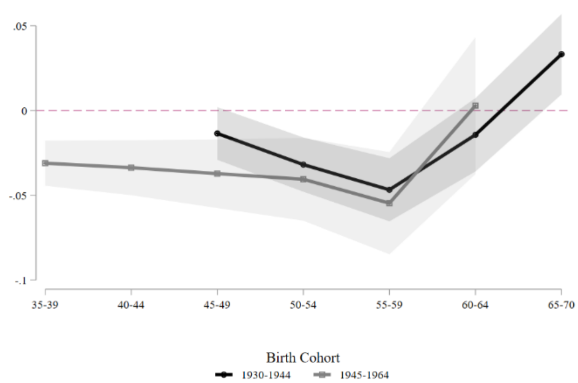
c: BA Degree or Greater – Less Than HS Degree



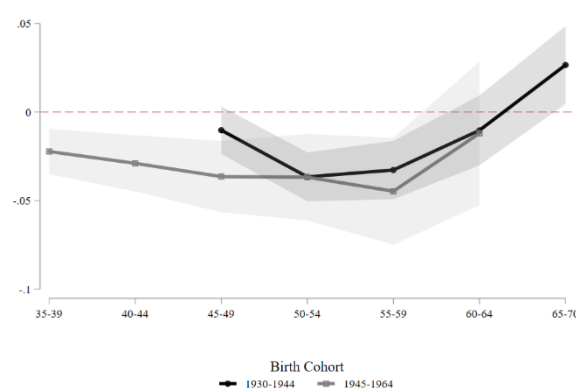
d: Some College – HS Degree



e: Some College – Less Than HS Degree

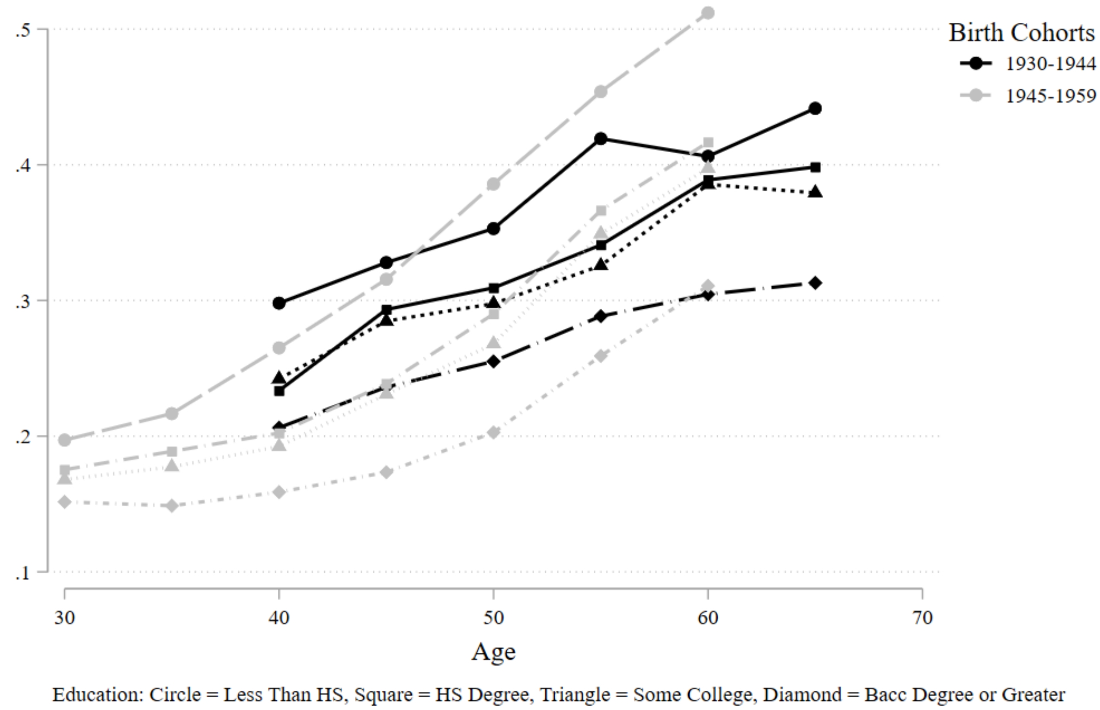


f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for or non-Hispanic, white NHIS respondents born between 1930 and 1959 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 9.

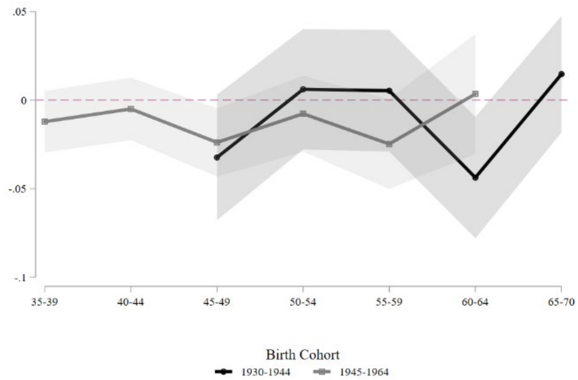
Figure 11. Proportion of Persons Reporting Hypertension by Five-Year Age Group and Education Level Among NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)



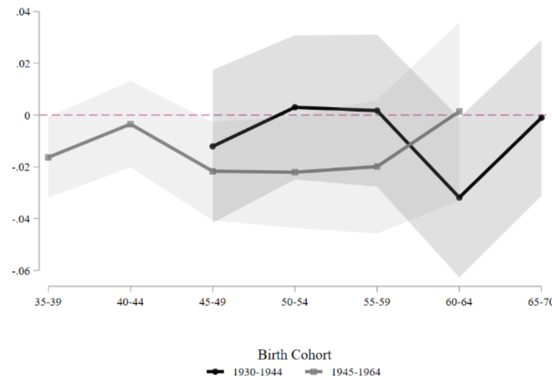
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1959 and surveyed between 1982 and 2016. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Figure 12. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Hypertension by Age for NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)

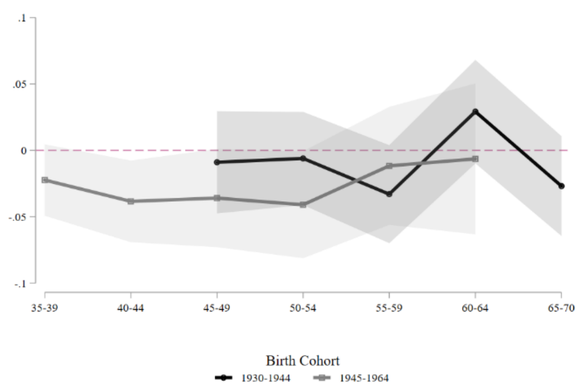
a: BA Degree or Greater – Some College



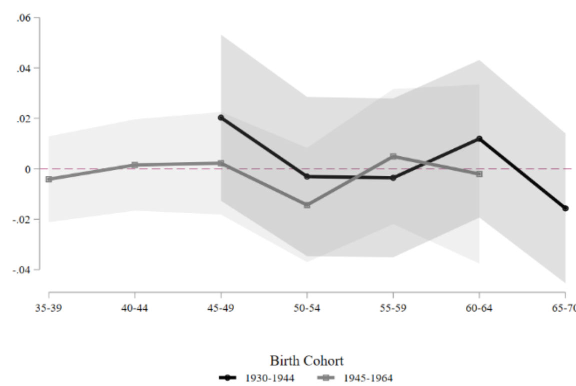
b: BA Degree or Greater – HS Degree



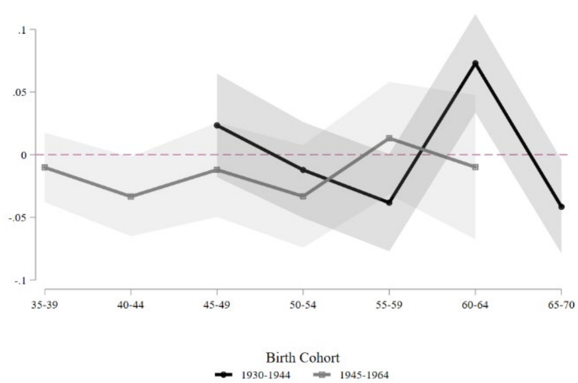
c: BA Degree or Greater – Less Than HS Degree



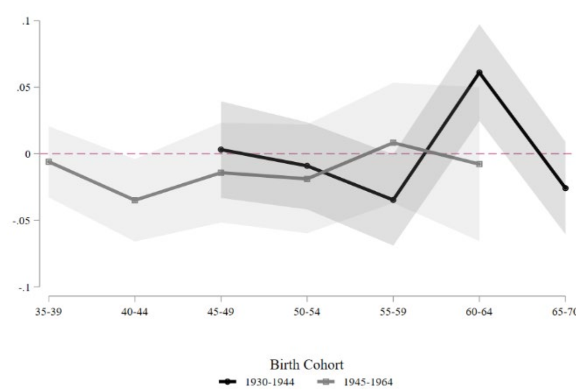
d: Some College – HS Degree



e: Some College – Less Than HS Degree

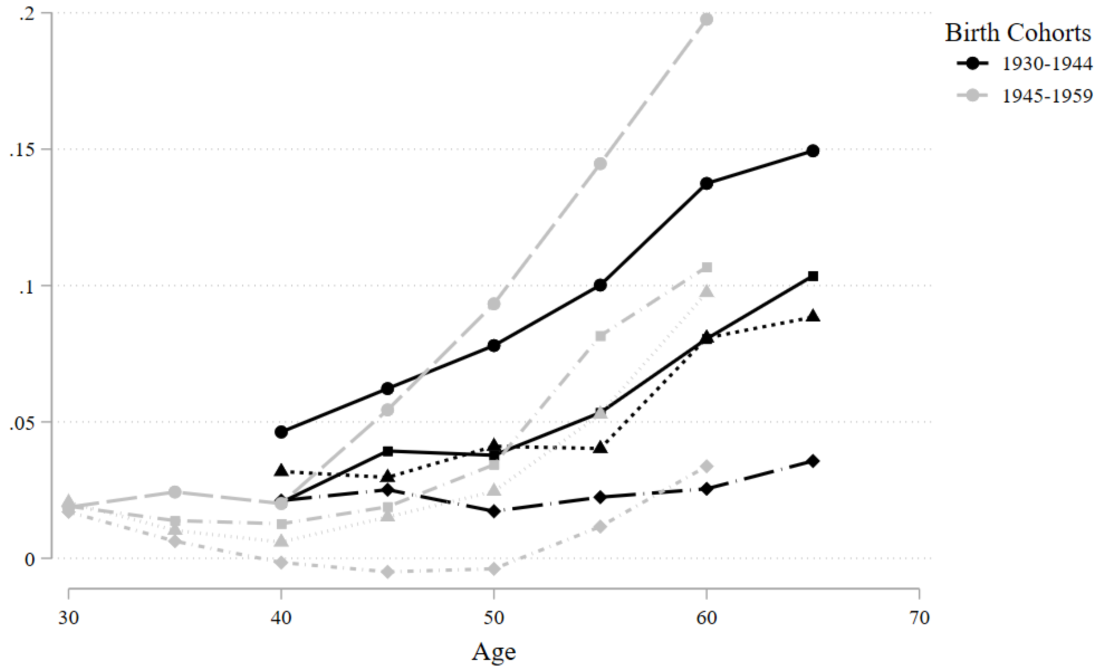


f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for or non-Hispanic, white NHIS respondents born between 1930 and 1959 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 11.

Figure 13. Proportion of Persons Reporting Diabetes by Five-Year Age Group and Education Level Among Respondents (1930-1944 and 1945-1959 Birth Cohorts)

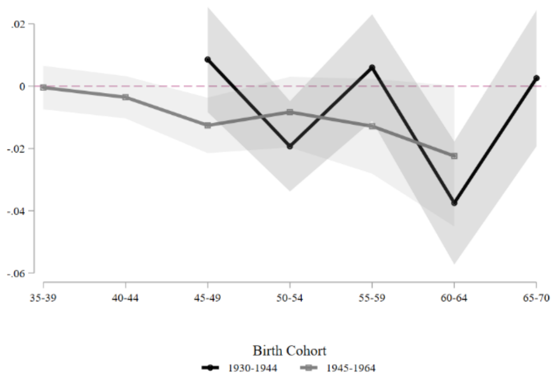


Education: Circle = Less Than HS, Square = HS Degree, Triangle = Some College, Diamond = Bacc Degree or Greater

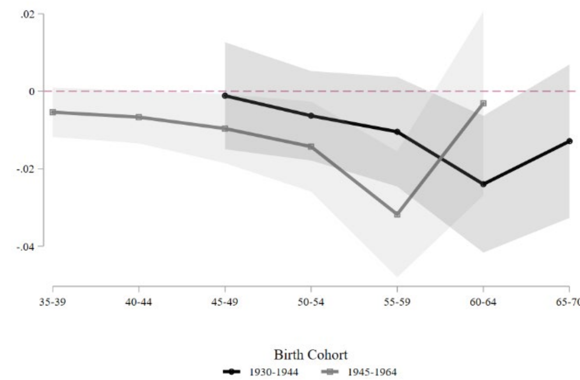
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1959 and surveyed between 1982 and 2016. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Figure 14. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Diabetes by Age for NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)

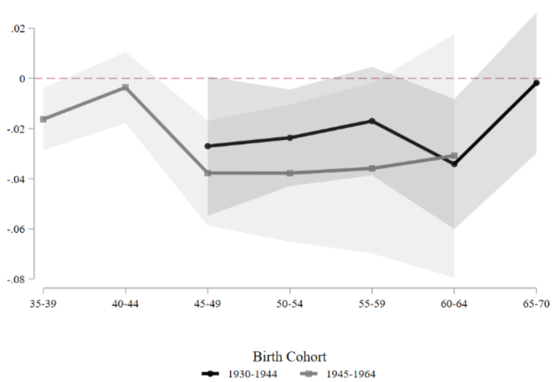
a: BA Degree or Greater – Some College



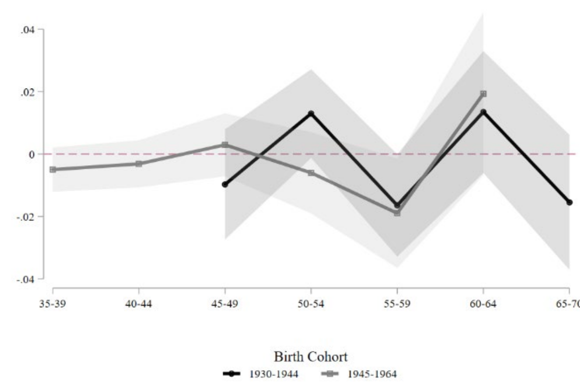
b: BA Degree or Greater – HS Degree



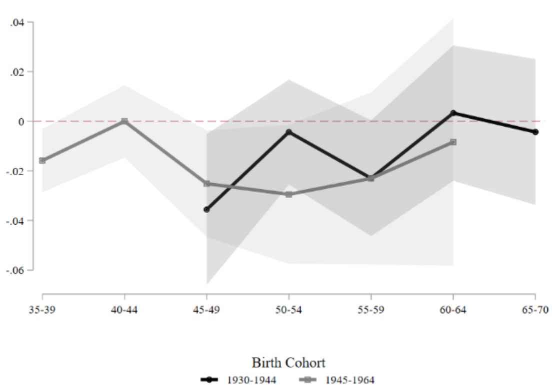
c: BA Degree or Greater – Less Than HS Degree



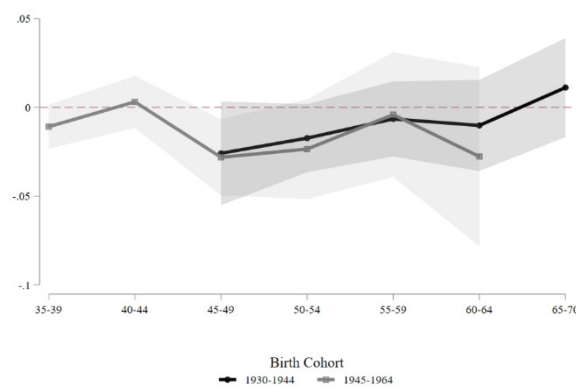
d: Some College – HS Degree



e: Some College – Less Than HS Degree

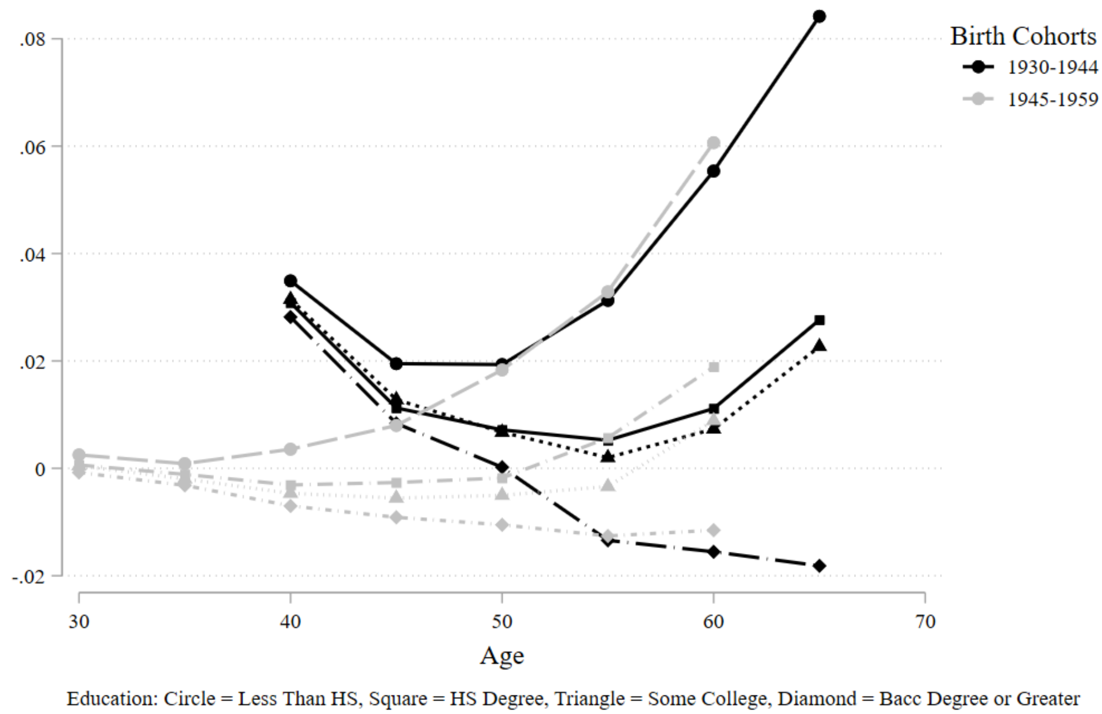


f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for or non-Hispanic, white NHIS respondents born between 1930 and 1959 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 13.

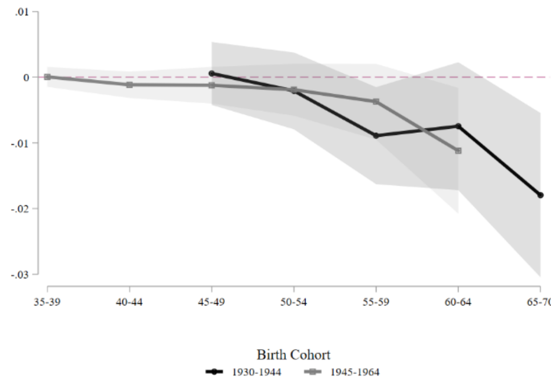
Figure 15. Proportion of Persons Widowed by Five-Year Age Group and Education Level Among NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)



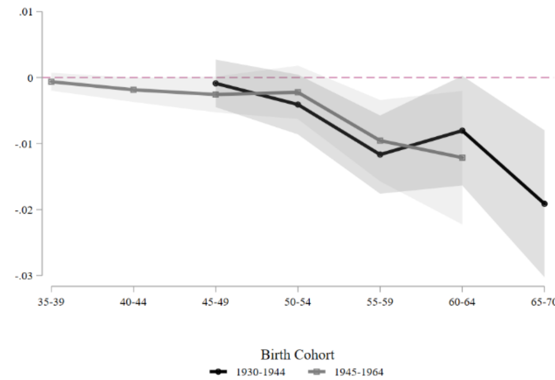
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1959 and surveyed between 1982 and 2016. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Figure 16. Differences-in-Differences Estimates of the Effect of Education on Proportion of Persons Widowed by Age for NHIS Respondents (1930-1944 and 1945-1959 Birth Cohorts)

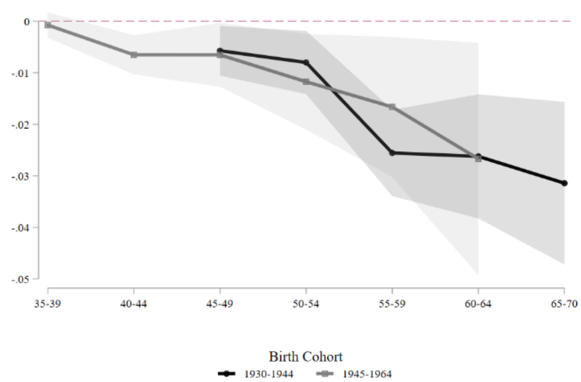
a: BA Degree or Greater – Some College



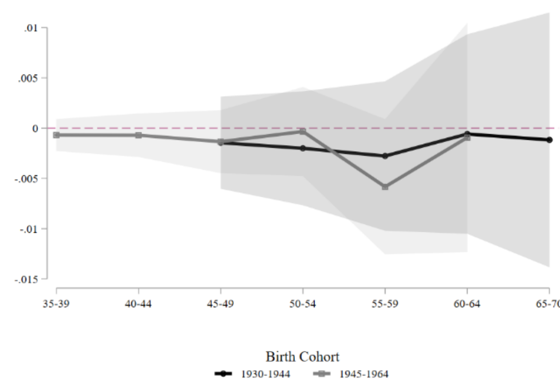
b: BA Degree or Greater – HS Degree



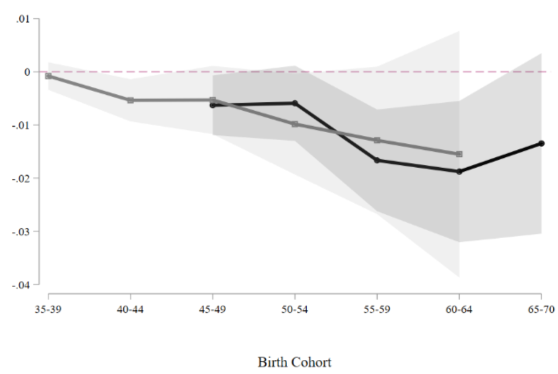
c: BA Degree or Greater – Less Than HS Degree



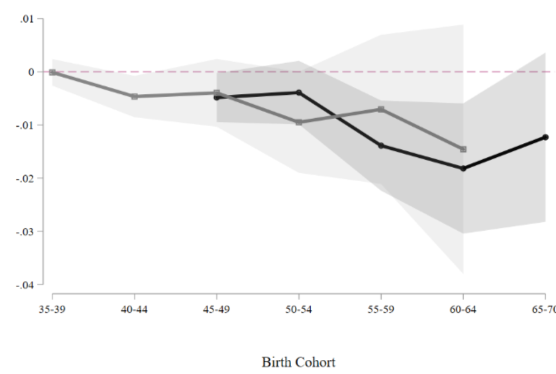
d: Some College – HS Degree



e: Some College – Less Than HS Degree



f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for or non-Hispanic, white NHIS respondents born between 1930 and 1959 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 15.

Table 1. Summary Statistics for CHA Cohorts at Baseline

1927-1934 Birth Cohort	Less than HS	HS Degree	Some College	BA or Greater
Age (year)	37.88	37.85	37.85	37.62
Female	0.35	0.47	0.31	0.13
White	0.75	0.89	0.85	0.93
Black	0.16	0.09	0.12	0.04
Height (inch)	66.65	67.14	68.27	69.59
Baseline Health Status (fraction with):				
Favorable Health (all factors were favorable)	0.02	0.06	0.06	0.06
1+ Health Factors were Elevated but None High	0.13	0.17	0.20	0.26
One Health Factor was High	0.41	0.41	0.43	0.42
2+ Health Factors were High	0.44	0.36	0.31	0.26
Number of Unique Persons (Total = 4,097)	679	1,513	675	1,230
Number of Person-years (Total =133,856)	21,282	49,727	21,685	41,162

1935-1942 Birth Cohort	Less than HS	HS Degree	Some College	BA or Greater
Age (year)	32.72	32.56	32.51	32.30
Female	0.39	0.40	0.29	0.16
White	0.60	0.82	0.76	0.92
Black	0.25	0.16	0.20	0.04
Height (inch)	66.80	67.56	68.21	69.53
Baseline Health Status (fraction with):				
Favorable Health (all factors were favorable)	0.06	0.06	0.09	0.09
1+ Health Factors were Elevated but None High	0.14	0.18	0.20	0.28
One Health Factor was High	0.40	0.43	0.42	0.41
2+ Health Factors were High	0.39	0.31	0.28	0.22
Number of Unique Persons (Total = 4,814)	665	1,594	960	1,595
Number of Person-years (Total =140,667)	18,950	46,230	28,095	47,392

Notes: Birth cohort 1927-1934 includes individuals who were aged 30 to 40 years at baseline and lived until at least age 40 (approximately 0.17% died before age 40). Birth cohort 1935-1942 includes individuals who were aged 30 to 40 years at baseline and lived until at least age 40 (approximately 0.62% died before age 40). Baseline health status is a composite measure of health status and incorporated health risk factors including BMI, blood pressure, whether individual has diabetes, serum cholesterol, and smoking status (see Allen et al. 2017).

Table 2. Summary Statistics for NHIS Non-White Hispanic Respondents by Birth Cohort

	1930-1934	1935-1939	1940-1944	1945-1949	1950-1954	1955-1959	1960-1964	Total
Age	56.71	54.04	51.69	46.96	44.50	42.29	39.97	47.19
	(7.31)	(8.77)	(10.34)	(10.40)	(9.07)	(7.62)	(6.16)	(10.29)
Female	0.519	0.514	0.511	0.507	0.506	0.503	0.503	0.508
Less than HS	0.221	0.178	0.135	0.092	0.070	0.074	0.071	0.109
HS Degree	0.426	0.421	0.386	0.347	0.338	0.354	0.336	0.366
Some College	0.164	0.190	0.218	0.250	0.276	0.277	0.287	0.245
BA or Greater	0.189	0.212	0.261	0.312	0.316	0.295	0.305	0.280
Observations	89,840	101,889	135,463	170,209	152,203	127,989	89,706	867,299

Source: National Health Interview Surveys (1976-2016). Means use NHIS person weights. Observations are actual respondent counts. Standard deviations in parentheses.

Table 3. Descriptive Statistics NLSY79 at Age 40

Variable	Mean (s.d.)
SF-12 Physical Score	52.2 (7.7)
SF_12 Mental Score	53.1 (8.1)
CESD	3.28 (4.1)
Self-reported Health Not Good	0.40
Male	0.47
Non-Hispanic Black	0.30
Hispanic	0.18
High School Degree	0.44
Some College	0.24
BA Degree or Greater	0.19
AFQT Percentile Score	40.8 (28.8)
Rotter Score (4 to 16)	8.7 (2.4)
Self-Esteem Score (7 to 3)	22.3 (4.0)
Mother's Education High School	0.40
Mother's Education Some College	0.10
Mother's Education Bachelor's or more	0.08
Two Parent Family at Age 14	0.70
Mother Only Family at Age 14	0.16
Library Card in Household Growing Up	0.72
Magazine in Household Growing Up	0.58
Newspaper in Household Growing Up	0.76

Notes: The number of observations is 5877. The mean age is 40.6 with a range from 39 to 42.

Table 4. Differences in the Predicted Probability of Survival by Education
CHA Birth Cohorts

CHA 1927-1934 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 70	Live to Age 75
BA - SC	0.019	0.033	0.067	0.089
BA - HS	0.014	0.035	0.043	0.064
BA - LTHS	0.026	0.067	0.134	0.197
SC - HS	-0.004	0.002	-0.024	-0.025
SC - LTHS	0.007	0.034	0.067	0.108
HS - LTHS	0.011	0.032	0.091	0.133
Mean Survival Rate for LTHS	0.947	0.856	0.701	0.576
CHA 1935-1942 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 70	
BA - SC	0.007	0.014	0.044	
BA - HS	0.010	0.034	0.067	
BA - LTHS	0.027	0.045	0.122	
SC - HS	0.003	0.020	0.023	
SC - LTHS	0.020	0.031	0.078	
HS - LTHS	0.017	0.011	0.055	
Mean Survival Rate for LTHS	0.933	0.867	0.749	

Notes: Data are from the CHA study. BA: BA degree or greater. SC: some college. HS: high school degree. LTHS: less than a high school degree. Table presents differences in the probability of surviving to the given age by education level derived from estimates of OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, race dummy variables, height, baseline health (composite measure), the interaction between female and age dummy variables, and the interaction between baseline health and age dummy variables. Predicted values are calculated at the mean of covariates (female, year, race, height, baseline health, interaction terms between female and age, and interaction terms between baseline health and age).

Table 5. Differences in Predicted Probability of Survival by Education
NHIS Birth Cohorts

Panel A: 1945-1949 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 64
BA – SC	0.007	0.007	0.015
BA – HS	-0.003	0.002	0.023
BA – LTHS	0.008	0.034	0.080
SC – HS	-0.010	-0.004	0.008
SC – LTHS	0.001	0.027	0.065
HS – LTHS	0.011	0.032	0.057
Mean Survival Rate for LTHS	0.959	0.898	0.837

Panel B: 1950-1954 Birth Cohort	Live to Age 50	Live to Age 59
BA – SC	0.014	0.014
BA – HS	0.014	0.032
BA – LTHS	0.027	0.077
SC – HS	0.000	0.018
SC – LTHS	0.013	0.063
HS – LTHS	0.013	0.045
Mean Survival Rate for LTHS	0.946	0.874

Notes: Data are from the National Health Interview Surveys from 1986-1992 with linked NDI mortality data through 2011. BA: BA degree or greater. SC: some college. HS: high school degree. LTHS: less than a high school degree. Table presents differences in the probability of surviving to the given age by education level derived from estimates of OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, baseline self-reported health (composite measure), the interaction between female and age dummy variables, and the interaction between baseline self-reported health and age dummy variables. Predicted values are calculated at the mean of covariates (female, survey year, baseline self-reported health, interaction terms between female and age, and interaction terms between baseline self-reported health and age).

Table 6. Estimates of the Association between Education and the SF-12 Physical Health Summary Measure

	Age 40				Age 50				Difference Age 50 – Age 40			
High School	2.05** (0.31)	1.90** (0.32)	1.71** (0.32)	1.44** (0.33)	3.48** (0.41)	3.18** (0.41)	3.03** (0.41)	2.64** (0.43)	1.44** (0.38)	1.27** (0.39)	1.31** (0.39)	1.18* (0.41)
Some College	3.09** (0.34)	2.86** (0.36)	2.56** (0.37)	2.16** (0.39)	4.94** (0.44)	4.34** (0.46)	4.10** (0.47)	3.45** (0.50)	1.81** (0.42)	1.40** (0.44)	1.46** (0.45)	1.20* (0.48)
BA or Greater	4.35** (0.36)	4.07** (0.40)	3.66** (0.40)	3.18** (0.44)	7.52** (0.47)	6.67** (0.51)	6.34** (0.52)	5.45** (0.57)	3.19** (0.44)	2.57** (0.48)	2.66** (0.50)	2.22** (0.54)
Family Background	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Rotter, Self-Esteem	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
AFQT	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Number of Obs.	5877	5877	5877	5877	5852	5852	5852	5852	5826	5826	5826	5826
Mean (sd) Dep. Var. For Omitted Group	49.8 (9.5)				45.3 (12.1)							

Notes: All models control for age fixed effects, gender fixed effects and race/ethnicity fixed effects. Family background factors include: mother education fixed effects (LTHS, HS, some college, BA or more); indicators for whether household had magazines, newspapers or library card while person was growing up; and family structure (two parents, mother only, other). Rotter locus of control scale and self-esteem scale are specified as dummy indicators for quartile of score. AFQT is specified as indicators of quartile of score.

Table 7. Estimates of the Association between Education and the SF-12 Mental Health Summary Measure

	Age 40				Age 50				Difference Age 50 – Age 40			
High School	1.37** (0.33)	1.23** (0.34)	0.96* (0.34)	0.75* (0.35)	1.99** (0.36)	1.95** (0.37)	1.60** (0.37)	1.23** (0.37)	0.68* (0.40)	0.77 (0.41)	0.70 (0.41)	0.56 (0.43)
Some College	2.16** (0.36)	1.98** (0.38)	1.49** (0.39)	1.22* (0.41)	2.82** (0.39)	2.80** (0.41)	2.20** (0.42)	1.70** (0.44)	0.67 (0.44)	0.84 (0.43)	0.72 (0.47)	0.52 (0.50)
BA or Greater	2.20** (0.38)	2.07** (0.42)	1.43** (0.43)	1.21* (0.47)	3.07** (0.41)	3.21** (0.45)	2.39** (0.46)	1.93** (0.50)	0.92* (0.46)	1.19* (0.50)	1.02* (0.52)	0.85 (0.56)
Family Background	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Rotter, Self-Esteem	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
AFQT	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Number of Obs.	5877	5877	5877	5877	5852	5852	5852	5852	5826	5826	5826	5826
Mean (sd) Dep. Var. For Omitted Group	51.6 (9.5)				51.0 (10.2)							

Notes: All models control for age fixed effects, gender fixed effects and race/ethnicity fixed effects. Family background factors include: mother education fixed effects (LTHS, HS, some college, BA or more); indicators for whether household had magazines, newspapers or library card while person was growing up; and family structure (two parents, mother only, other). Rotter locus of control scale and self-esteem scale are specified as dummy indicators for quartile of score. AFQT is specified as indicators of quartile of score.

Table 8. Estimates of the Association between Education and the CESD Scale of Depression and Education

	Age 40				Age 50				Difference Age 50 – Age 40			
High School	-1.13** (0.17)	-0.98** (0.17)	-0.80** (0.17)	-0.55* (0.18)	-1.32** (0.18)	-1.26** (0.18)	-1.09** (0.18)	-0.87** (0.19)	-0.18 (0.19)	-0.27 (0.20)	-0.28 (0.20)	-0.32 (0.21)
Some College	-1.91** (0.18)	-1.71** (0.19)	-1.39** (0.19)	-1.02** (0.21)	-1.97** (0.20)	-1.88** (0.21)	-1.60** (0.21)	-1.25** (0.22)	-0.07 (0.21)	-0.17 (0.22)	-0.21 (0.22)	-0.25 (0.24)
BA or Greater	-2.33** (0.19)	-2.12** (0.21)	-1.69** (0.22)	-1.25** (0.23)	-2.58** (0.21)	-2.50** (0.23)	-2.11** (0.23)	-1.68** (0.25)	-0.23 (0.22)	-0.35 (0.24)	-0.39 (0.25)	-0.42 (0.27)
Family Background	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Rotter, Self-Esteem	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
AFQT	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Number of Obs.	5855	5855	5855	5855	5854	5854	5854	5854	5808	5808	5808	5808
Mean (sd) Dep. Var.	4.7				5.3							
For Omitted Group	(5.1)				(5.2)							

Notes: All models control for age fixed effects, gender fixed effects and race/ethnicity fixed effects. Family background factors include: mother education fixed effects (LTHS, HS, some college, BA or more); indicators for whether household had magazines, newspapers or library card while person was growing up; and family structure (two parents, mother only, other). Rotter locus of control scale and self-esteem scale are specified as dummy indicators for quartile of score. AFQT is specified as indicators of quartile of score.

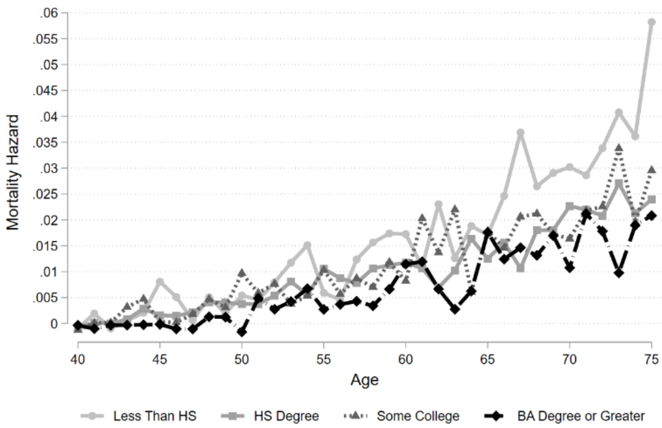
Table 9. Estimates of the Association between Education and Self-reported Poor Health

	Age 40				Age 50				Difference Age 50 – Age 40			
High School	-0.14** (0.02)	-0.12** (0.02)	-0.10** (0.02)	-0.07** (0.02)	-0.12** (0.02)	-0.10** (0.02)	-0.08** (0.02)	-0.06** (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)
Some College	-0.22** (0.02)	-0.20** (0.02)	-0.15** (0.02)	-0.11** (0.02)	-0.21** (0.02)	-0.17** (0.02)	-0.13** (0.02)	-0.10** (0.02)	0.02 (0.02)	0.03 (0.02)	0.02 (0.03)	0.01 (0.03)
BA or Greater	-0.32** (0.02)	-0.30** (0.02)	-0.24** (0.03)	-0.18** (0.03)	-0.35** (0.02)	-0.31** (0.03)	-0.26** (0.03)	-0.20** (0.03)	-0.03 (0.03)	-0.01 (0.03)	-0.01 (0.03)	-0.02 (0.03)
Family Background	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Rotter, Self-Esteem	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
AFQT	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Number of Obs.	5899	5899	5899	5899	5901	5901	5901	5901	5901	5901	5901	5901
Mean Dep. Var. For Omitted Group	0.58	0.58	0.58	0.58	0.68	0.68	0.68	0.68				

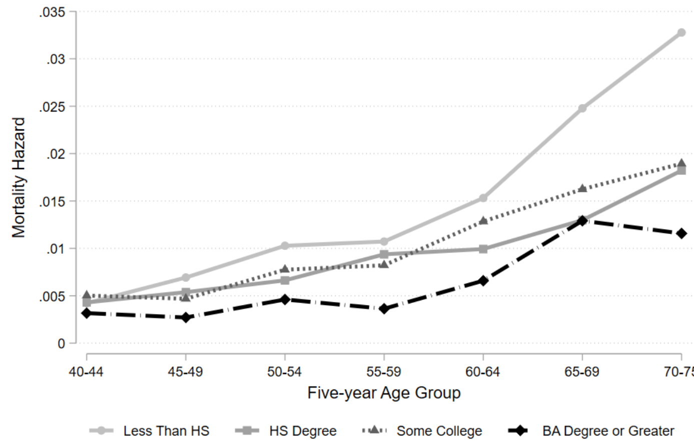
Notes: All models control for age fixed effects, gender fixed effects and race/ethnicity fixed effects. Family background factors include: mother education fixed effects (LTHS, HS, some college, BA or more); indicators for whether household had magazines, newspapers or library card while person was growing up; and family structure (two parents, mother only, other). Rotter locus of control scale and self-esteem scale are specified as dummy indicators for quartile of score. AFQT is specified as indicators of quartile of score.

Appendix Figure 1. Predicted Hazard Rate of Death by Education and Age – CHA Cohorts

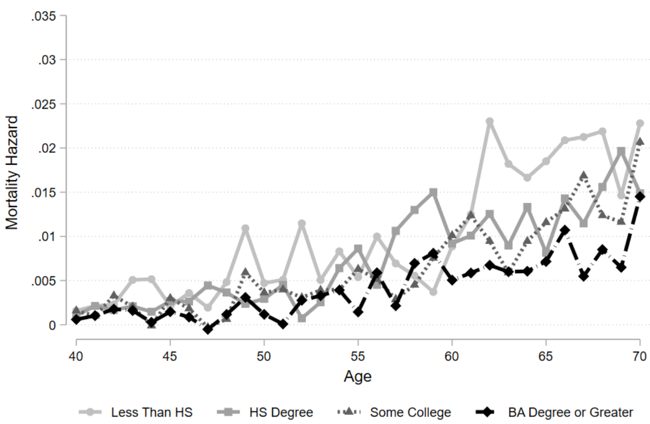
a. 1927-1934 Birth Cohort



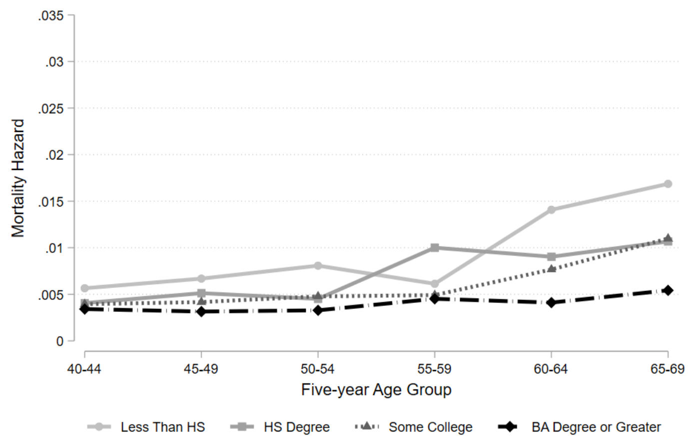
b. 1927-1934 Birth Cohort



c. 1935–1942 Birth Cohort



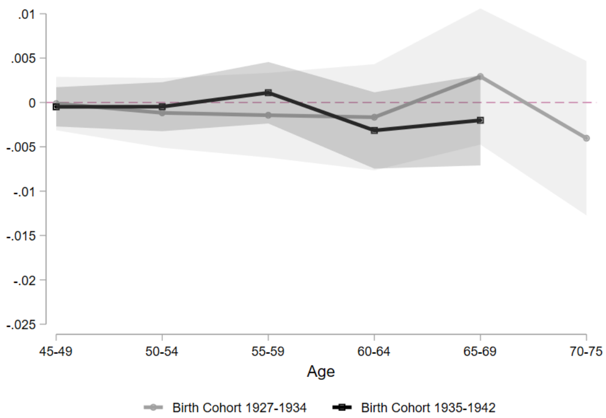
d. 1935–1942 Birth Cohort



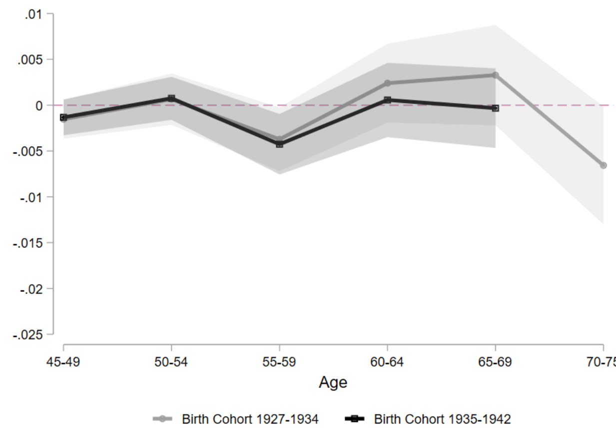
Notes: Figures present single-year and 5-year hazard rates of death by age and education for CHA 1927-1934 birth cohort and 1935-1942 birth cohort. Predicted hazard rate of death are obtained from an OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dumy variable, year dummy variables, and race dummy variables. Predicted values are calculated at the mean of covariates (female, year, and race).

Appendix Figure 2. Difference-in-Differences in the Hazard Rate of Death by Education and Age - CHA Cohorts

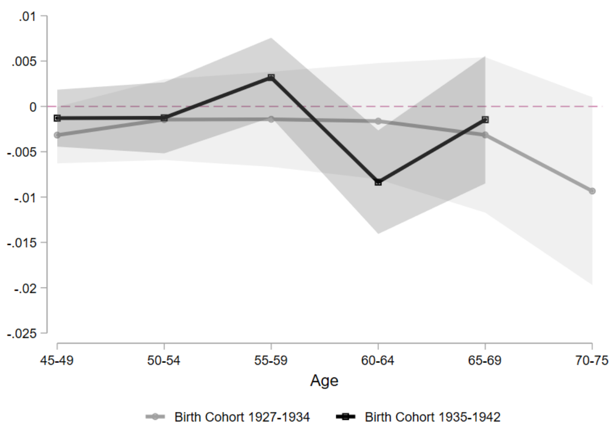
a. BA Degree or Greater – Some College



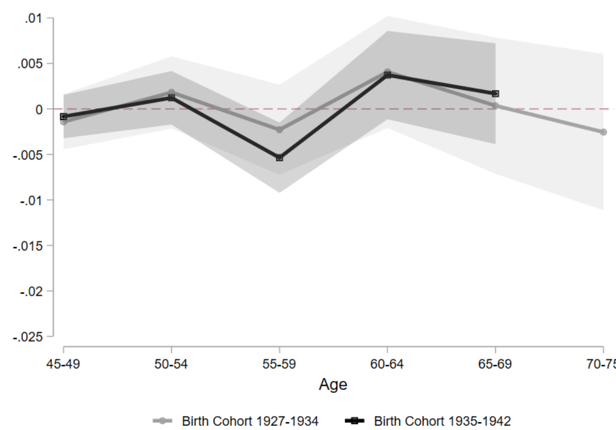
b. BA Degree or Greater – HS Degree



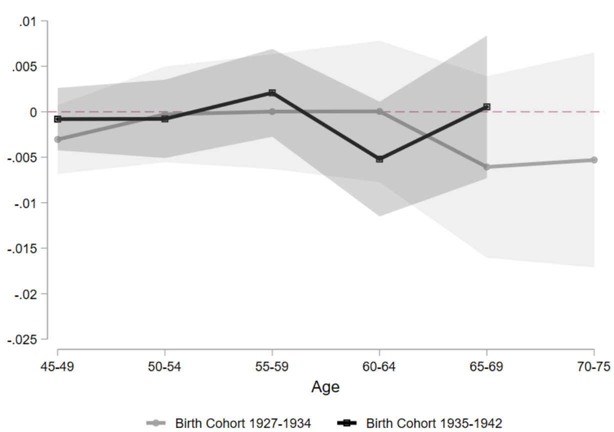
c. BA Degree or Greater – Less Than HS Degree



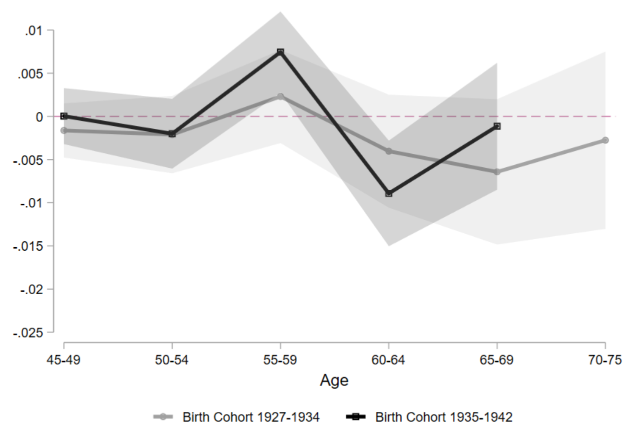
d. Some College – HS Degree



e. Some College – Less Than HS Degree



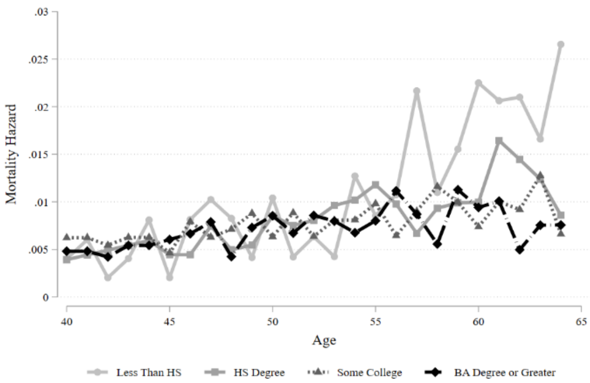
f. HS Degree – Less than HS Degree



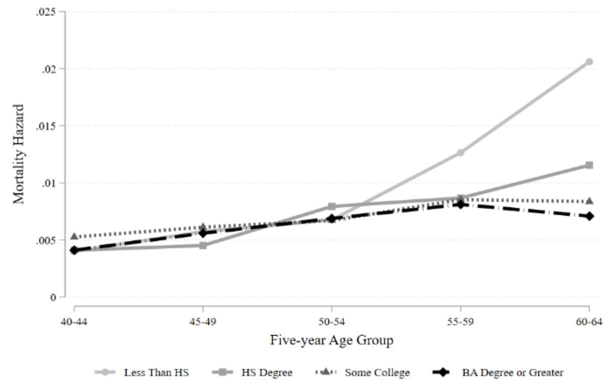
Notes: Figures present differences in the changes in hazard rate of death by age between education groups for CHA 1927-1934 birth cohort and 1935-1942 birth cohort. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 1.

Appendix Figure 3. Predicted Hazard Rate of Death by Education and Age – NHIS Cohorts

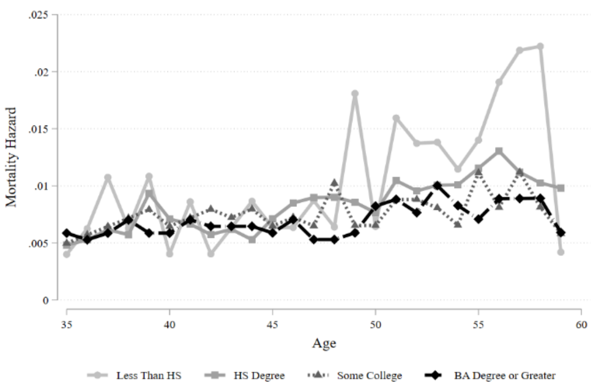
a. 1945-1949 Birth Cohort



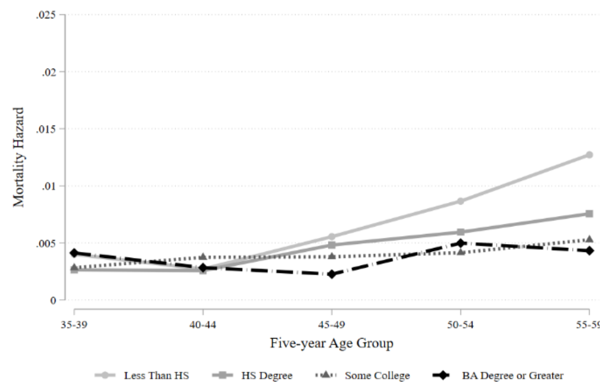
b. 1945-1949 Birth Cohort



c. 1950-1954 Birth Cohort



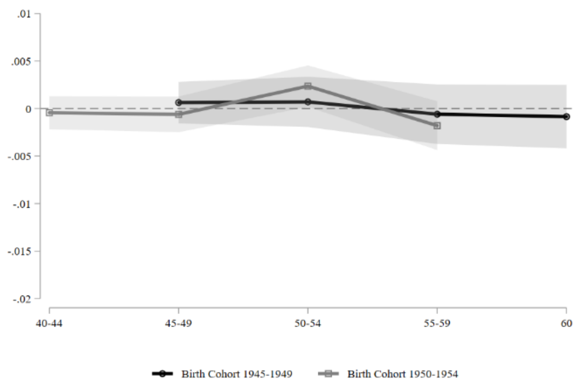
d. 1950-1954 Birth Cohort



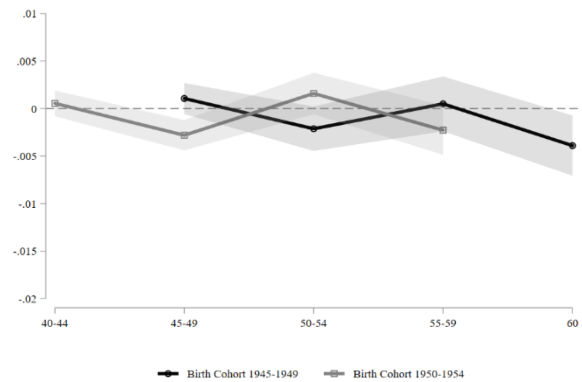
Notes: Figures present single-year and 5-year hazard rates of death by age and education for non-Hispanic white NHIS respondents in indicated birth cohorts who were initially surveyed between 1986 and 1990. Predicted hazard rates of death are obtained from an OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, and baseline self-reported health. Predicted values are calculated at the mean of covariates (female, survey year, baseline self-reported health).

Appendix Figure 4. Difference-in-Differences in the Hazard Rate of Death by Education and Age - NHIS Cohorts

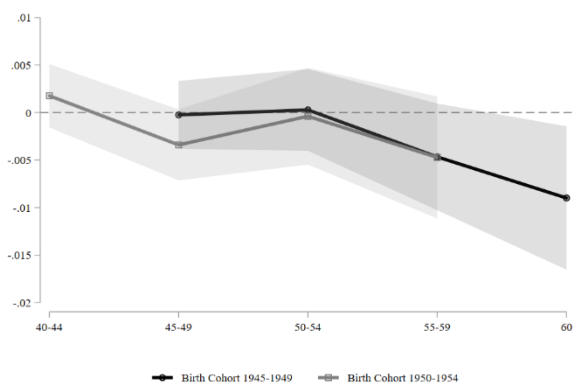
a. BA Degree or Greater – Some College Degree



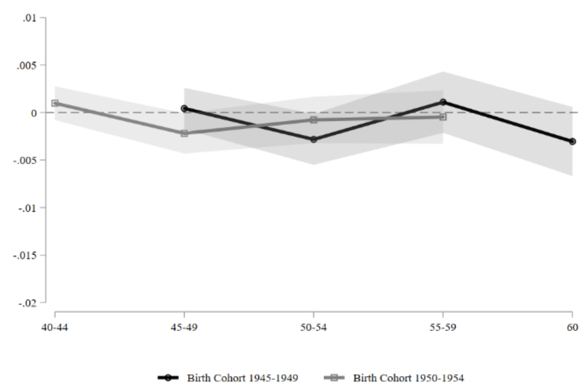
b. BA Degree or Greater – HS



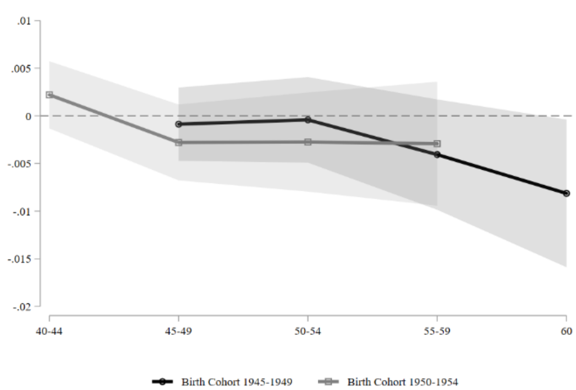
c. BA Degree or Greater – Less than HS Degree



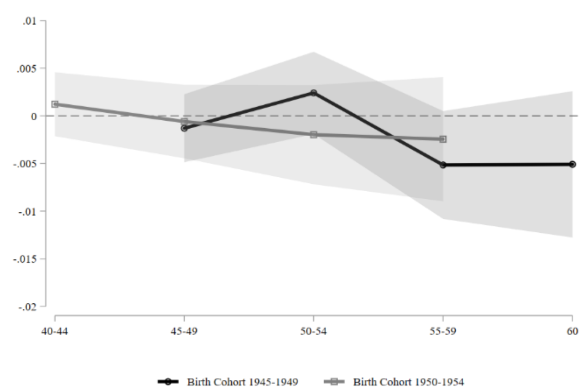
d. Some College – HS Degree



e. Some College – Less Than HS Degree Degree



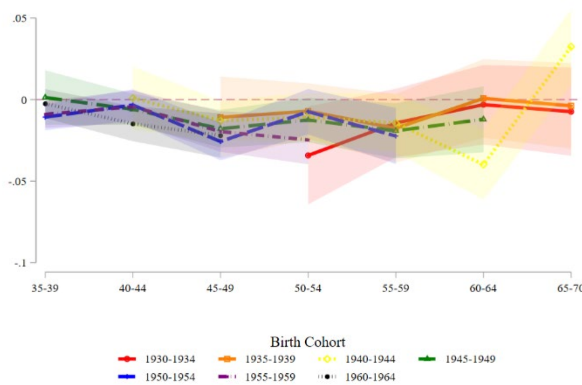
f. HS Degree – Less than HS



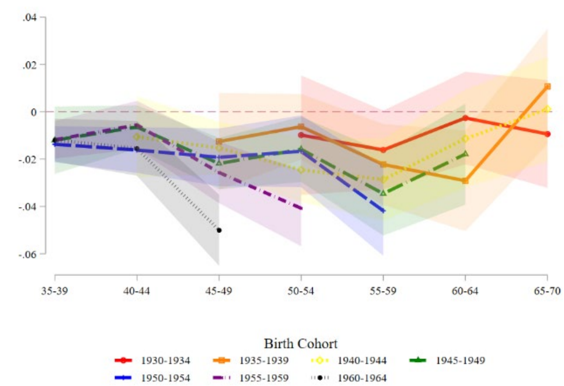
Notes: Figures present differences in the changes in hazard rate of death by age between education groups for NHIS 1945-49 birth cohort and 1950-54 birth cohort. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 3.

Appendix Figure 5. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Poor Health by Age for NHIS Respondents

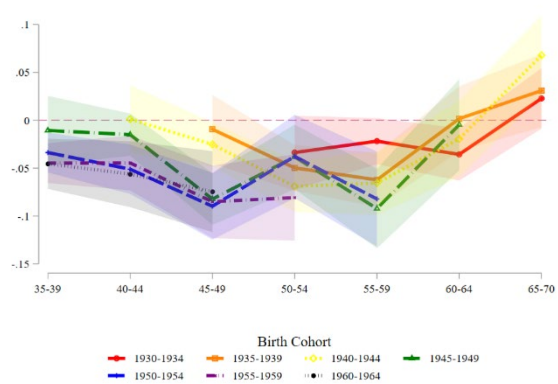
a: BA Degree or Greater – Some College



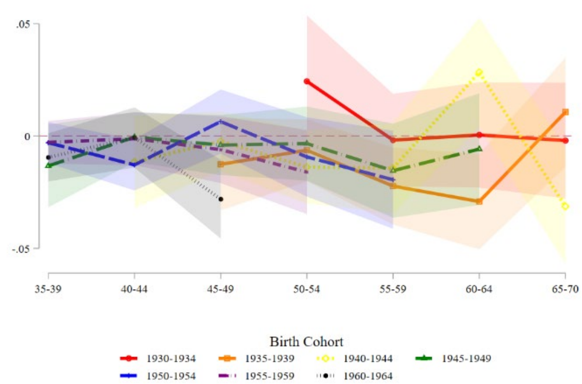
b: BA Degree or Greater – HS Degree



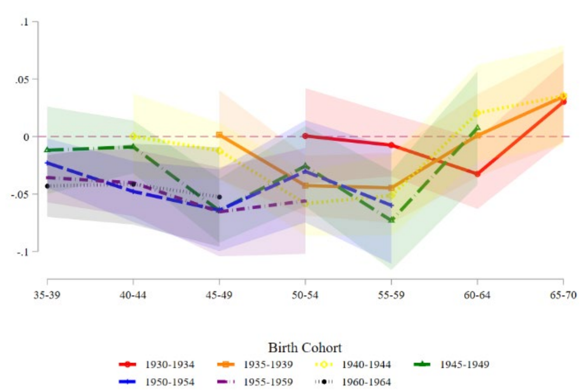
c: BA Degree or Greater – Less Than HS Degree



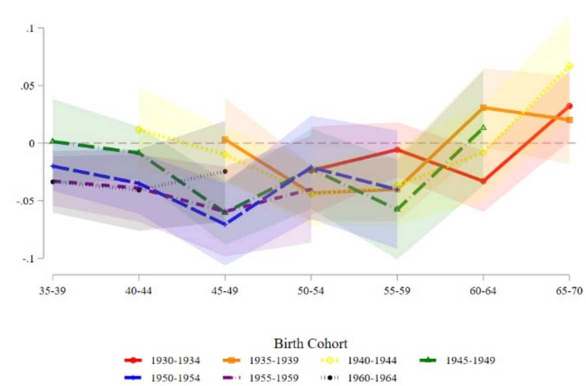
d: Some College – HS Degree



e: Some College – Less Than HS Degree

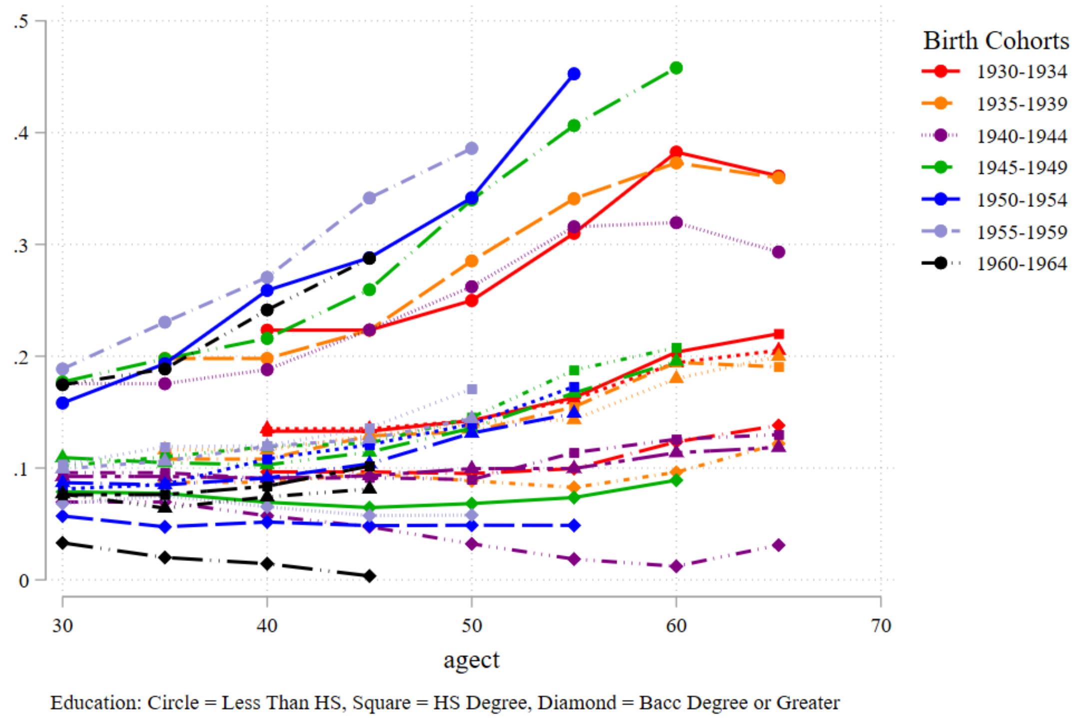


f: HS Degree – Less Than HS Degree



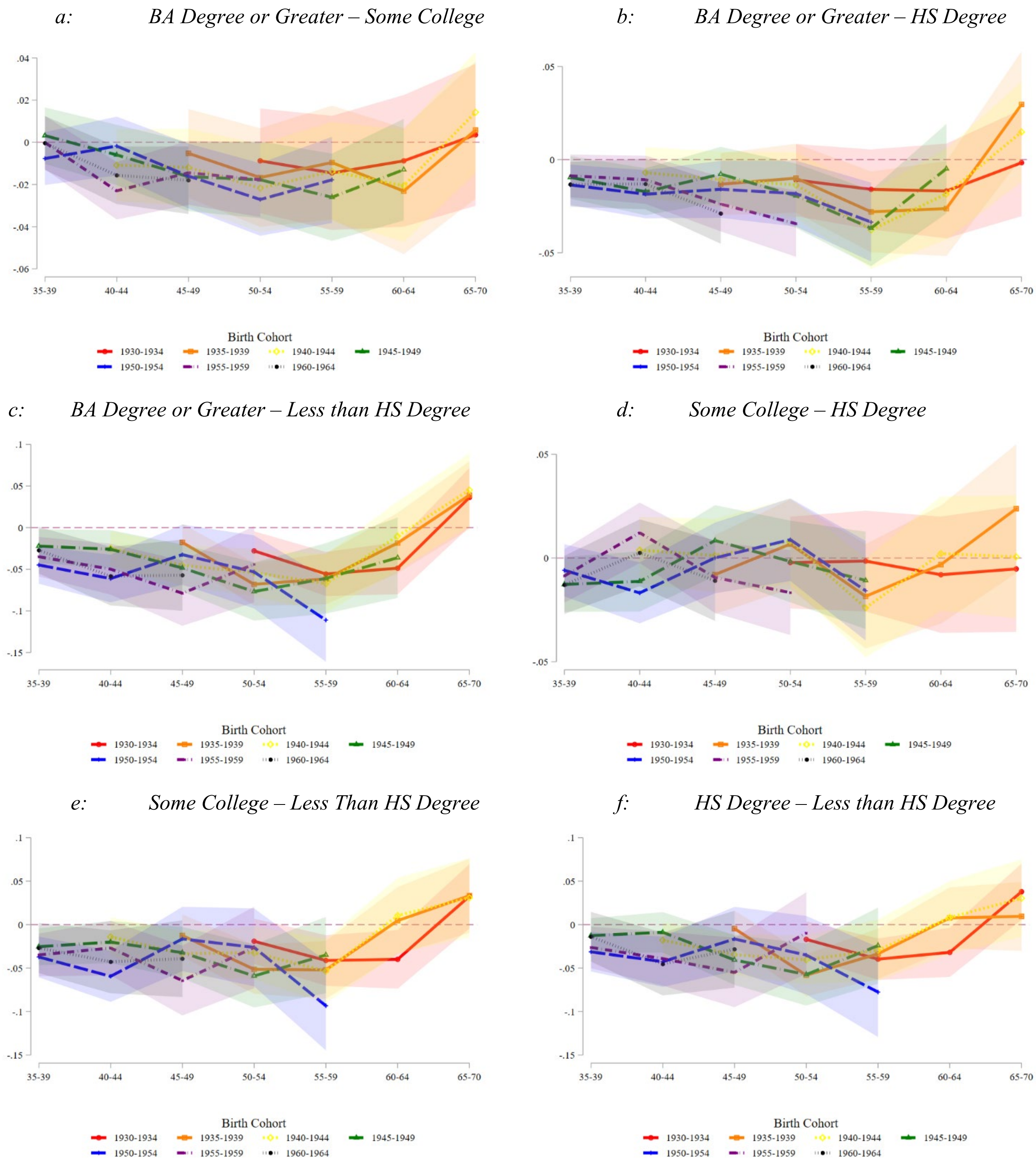
Notes: Figures present differences in the changes in outcome by age between education groups for non-Hispanic, white NHIS respondents born between 1930 and 1964 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Figure 6.

Appendix Figure 6. Proportion of Persons Reporting Any Limitation by Five-Year Age Group and Education Level Among NHIS Respondents Born Between 1930 and 1964



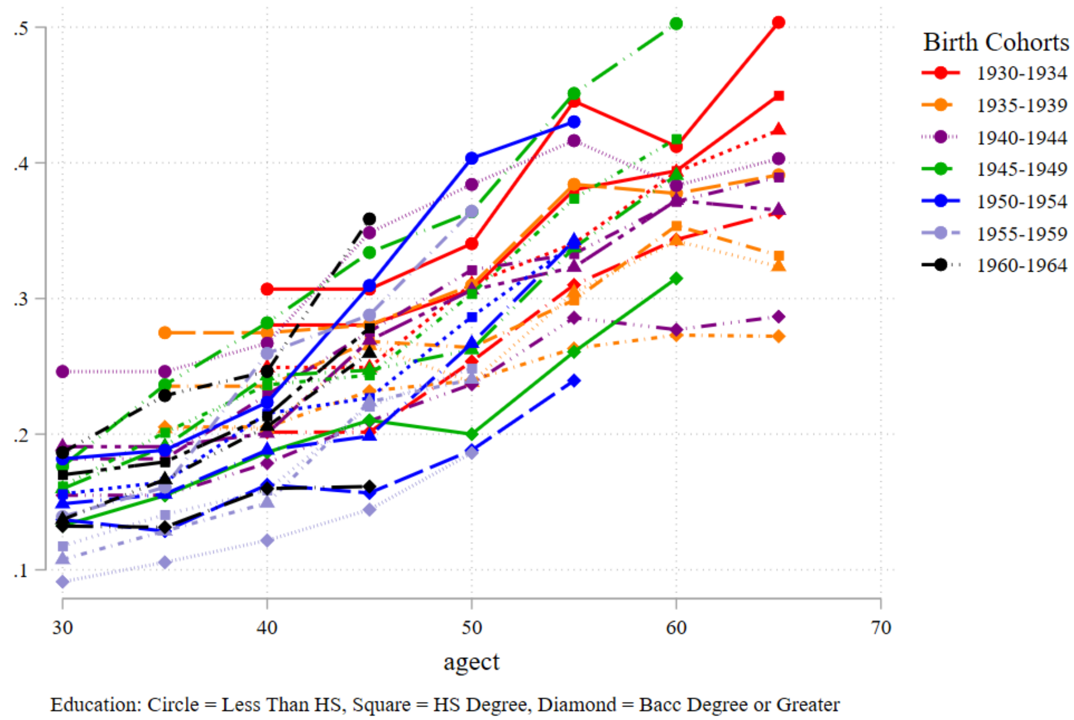
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1964 and surveyed between 1982 and 2016. To aid in interpretation, "Some College" cohorts are excluded from figure. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Appendix Figure 7. Differences-in-Differences Estimates of the Effect of Education on the Probability of Reporting Any Limitation by Age for NHIS Respondents



Notes: Figures present differences in the changes in outcome by age between education groups for non-Hispanic, white NHIS respondents born between 1930 and 1964 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 6.

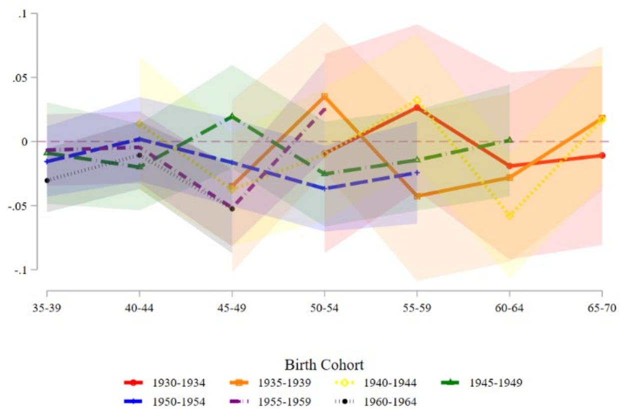
Appendix Figure 8. Proportion of Persons Reporting Ever Being Diagnosed with Hypertension by Five-Year Age Group and Education Level Among NHIS Respondents Born Between 1930 and 1964



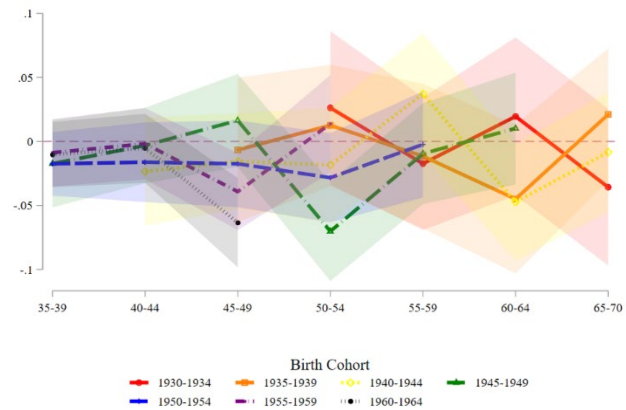
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1964 and surveyed between 1982 and 2016. To aid in interpretation, "Some College" cohorts are excluded from figure. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Appendix Figure 9. Differences-in-Differences Estimates of the Effect of Education on the Probability of Being Diagnosed with Hypertension among NHIS Respondents

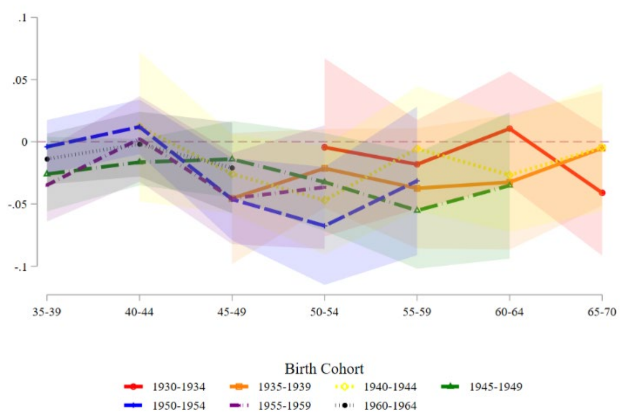
a: BA Degree or Greater – Some College



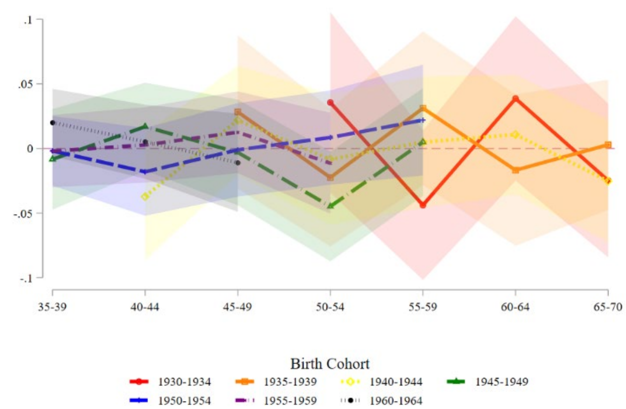
b: BA Degree or Greater – HS Degree



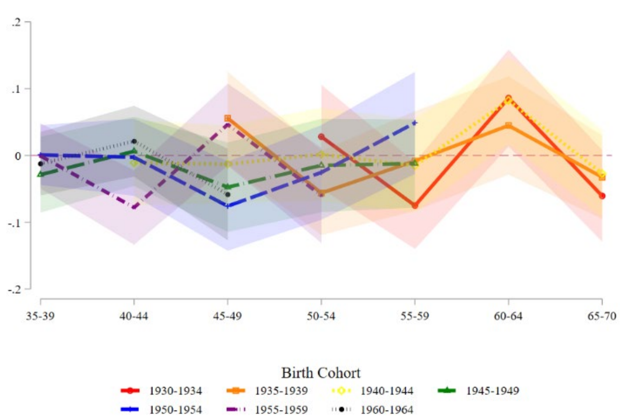
c: BA Degree or Greater – Less than HS Degree



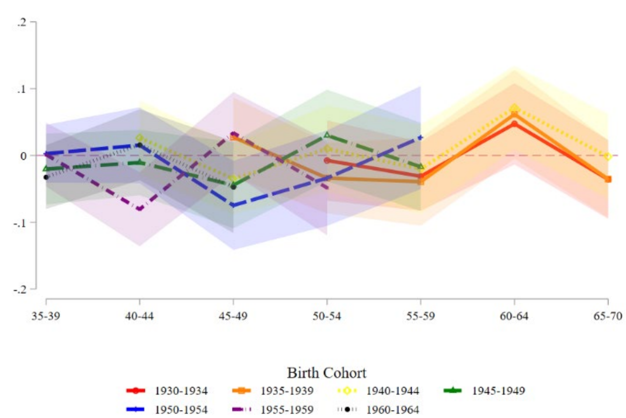
d: Some College – HS Degree



e: Some College – Less Than HS Degree

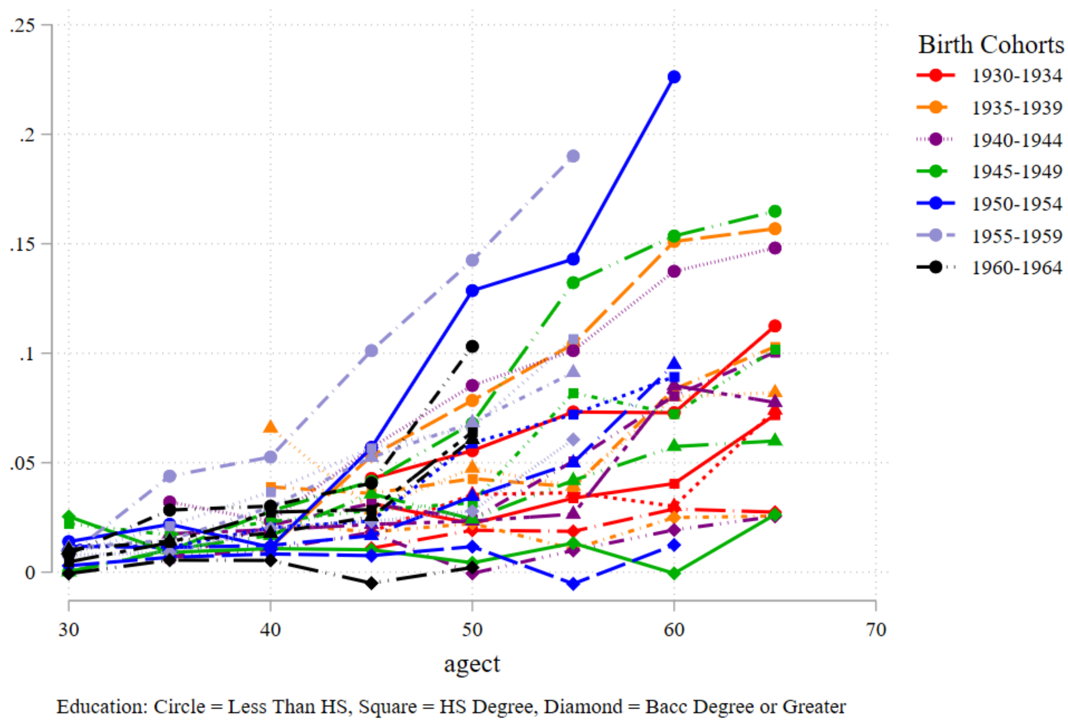


f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for non-Hispanic, white NHIS respondents born between 1930 and 1964 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 8.

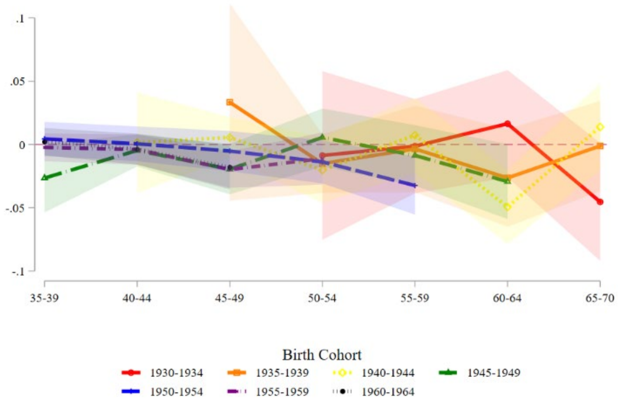
Appendix Figure 10: Proportion of Persons Reporting Ever Being Diagnosed with Diabetes by Five-Year Age Group and Education Level Among NHIS Respondents Born Between 1930 and 1964



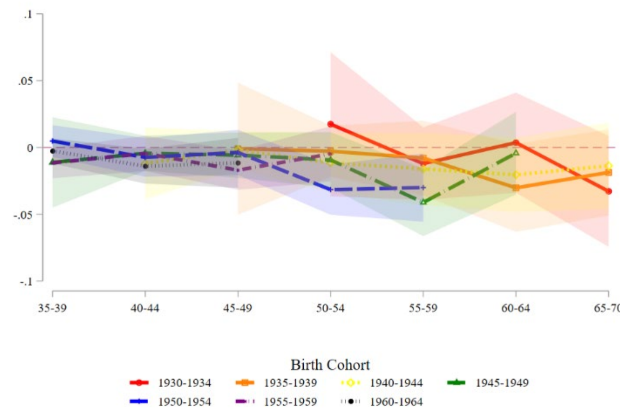
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1964 and surveyed between 1982 and 2016. To aid in interpretation, "Some College" cohorts are excluded from figure. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Appendix Figure 11: Differences-in-Differences Estimates of the Effect of Education on the Probability of Being Diagnosed with Diabetes among Non-Hispanic, White NHIS Respondents

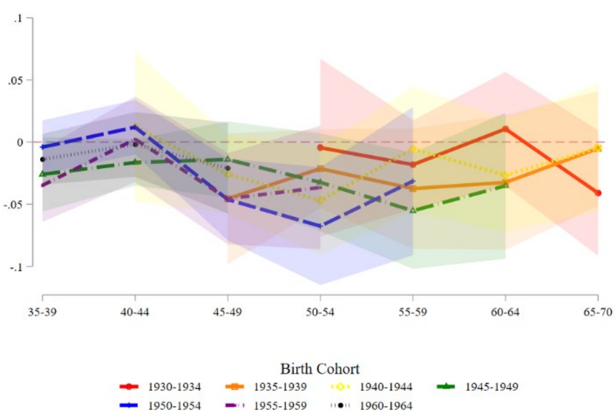
a: BA Degree or Greater – Some College



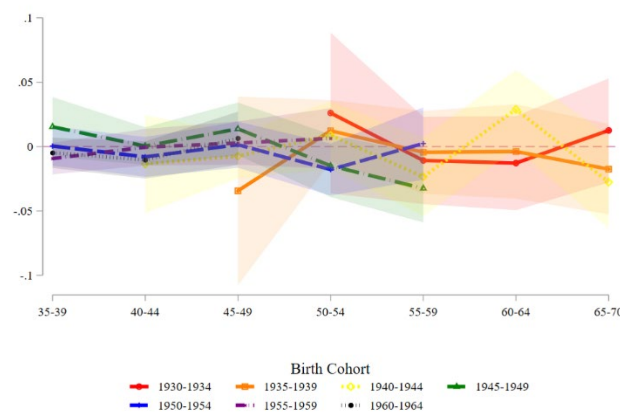
b: BA Degree or Greater – HS Degree



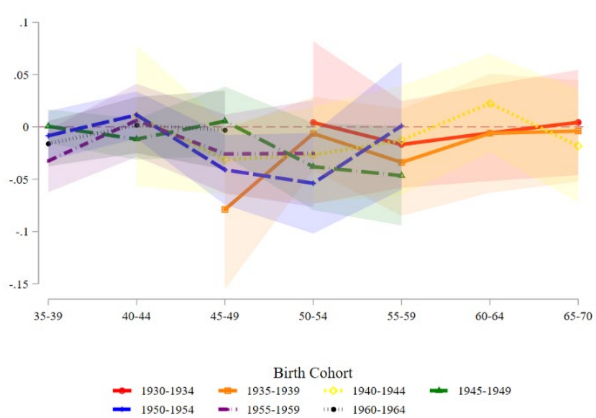
c: BA Degree or Greater – Less than HS Degree



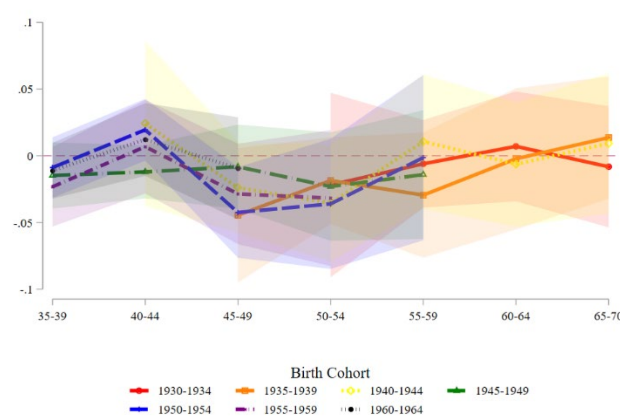
d: Some College – HS Degree



e: Some College – Less Than HS Degree

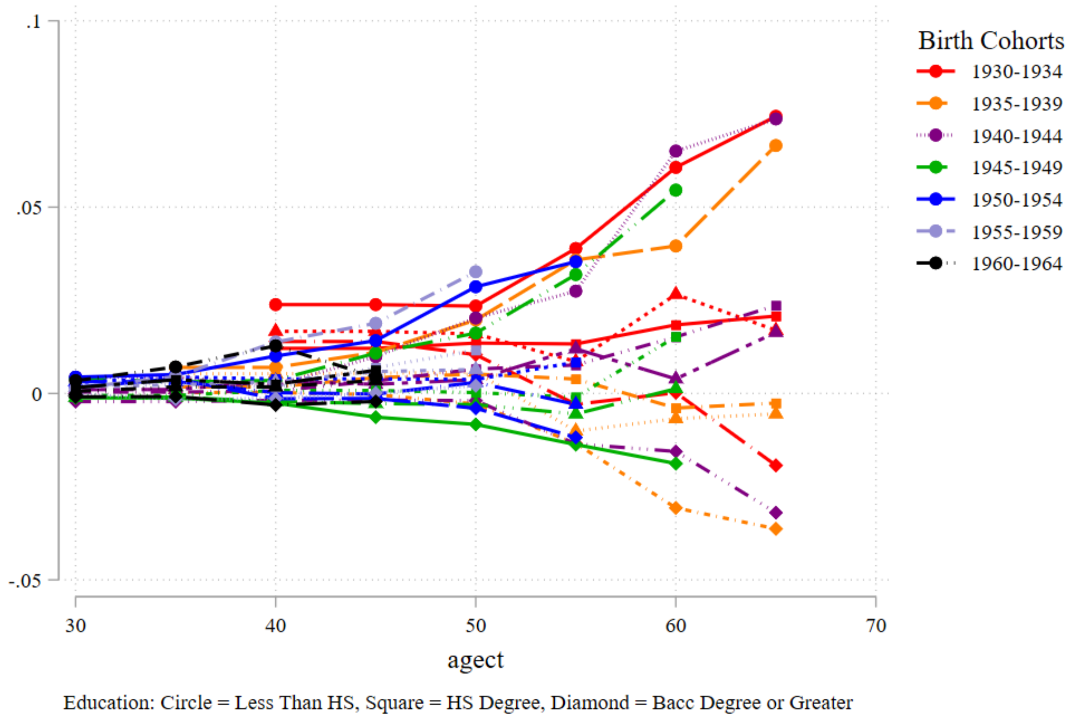


f: HS Degree – Less than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for non-Hispanic, white NHIS respondents born between 1930 and 1964 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 10.

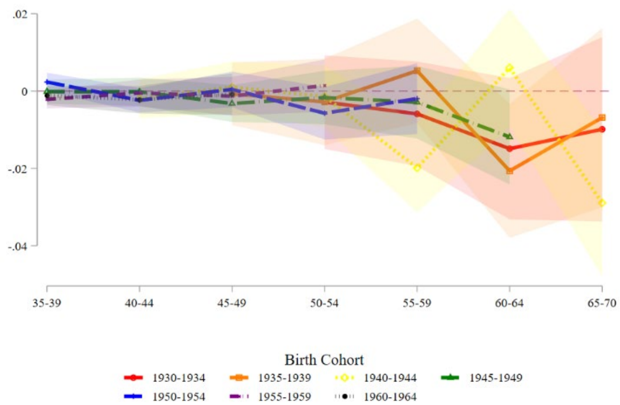
Appendix Figure 12: Proportion of Persons Widowed by Five-Year Age Group and Education Level Among NHIS Respondents Born Between 1930 and 1964



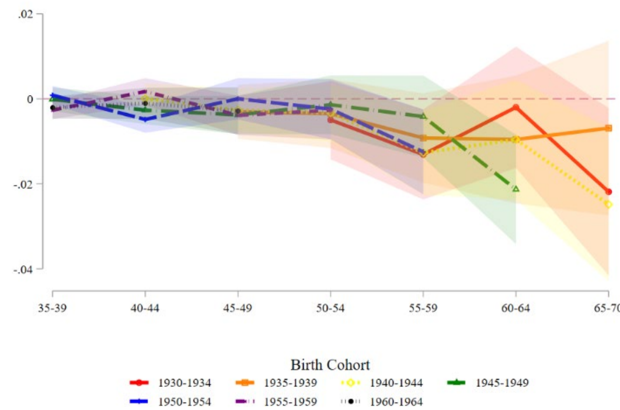
Notes: Sample limited to non-Hispanic white NHIS respondents born between 1930 and 1964 and surveyed between 1982 and 2016. To aid in interpretation, "Some College" cohorts are excluded from figure. Regression-based means are from an OLS regression using NHIS person weights and controlling for gender, gender-by-age interactions, and survey year fixed effects. Heteroskedasticity robust standard errors are estimated.

Appendix Figure 13: Differences-in-Differences Estimates of the Effect of Education on the Probability of Being Widowed among NHIS Respondents

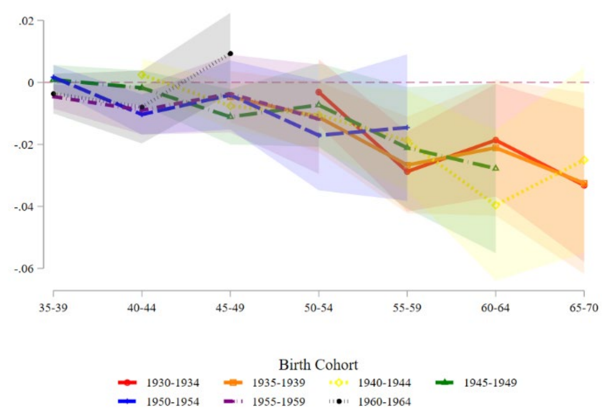
a: BA Degree or Greater – Some College



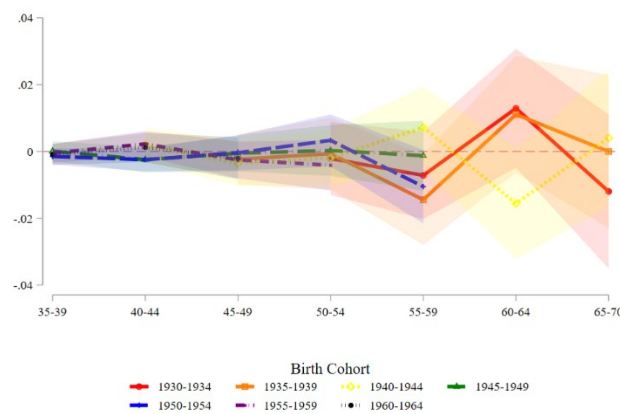
b: BA Degree or Greater – HS Degree



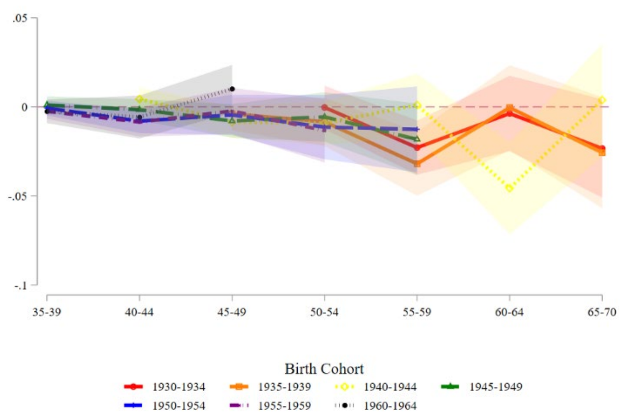
c: BA Degree or Greater – Less than HS Degree



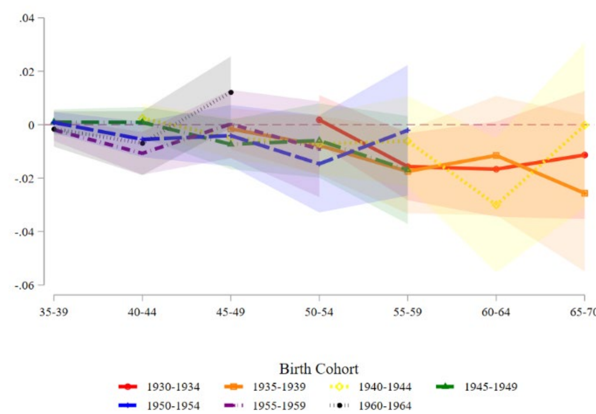
d: Some College – HS Degree



e: Some College – Less Than HS Degree



f: HS Degree – Less Than HS Degree



Notes: Figures present differences in the changes in outcome by age between education groups for non-Hispanic, white NHIS respondents born between 1930 and 1964 who were surveyed between 1982 and 2016. The difference (by education)-in-differences (by age) estimates in the figure are obtained from estimates of the regression models underlying Appendix Figure 12.

Appendix Table 1. Differences in the Predicted Probability of Survival by Education – CHA Cohorts

CHA 1927-1934 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 70	Live to Age 75
BA - SC	0.030	0.050	0.090	0.109
BA - HS	0.024	0.053	0.070	0.083
BA - LTHS	0.032	0.086	0.164	0.217
SC - HS	-0.006	0.002	-0.020	-0.026
SC - LTHS	0.002	0.036	0.074	0.107
HS - LTHS	0.008	0.034	0.094	0.134
Mean Survival Rate for LTHS	0.972	0.868	0.687	0.562
CHA 1935-1942 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 70	
BA - SC	0.010	0.021	0.058	
BA - HS	0.014	0.047	0.085	
BA - LTHS	0.031	0.058	0.143	
SC - HS	0.004	0.026	0.027	
SC - LTHS	0.021	0.037	0.085	
HS - LTHS	0.017	0.011	0.057	
Mean Survival Rate for LTHS	0.957	0.891	0.735	

Notes: Table presents differences in the probability of surviving to the given age by education level derived from regression-based hazard-of-death estimates from a linear probability model. BA: BA degree or greater. SC: some college. HS: high school degree. LTHS: less than a high school degree. Covariates include education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, year dummy variables, race dummy variables.

Appendix Table 2: Differences in Predicted Probability of Survival by Education
NHIS Birth Cohorts

Panel A: 1945-1949 Birth Cohort	Live to Age 50	Live to Age 60	Live to Age 64
BA – SC	0.006	0.007	0.015
BA – HS	-0.005	0.003	0.023
BA – LTHS	0.002	0.033	0.080
SC – HS	-0.011	-0.004	0.008
SC – LTHS	-0.004	0.026	0.065
HS – LTHS	0.007	0.030	0.057
Mean Survival Rate for LTHS	0.973	0.898	0.836

Panel B: 1950-1954 Birth Cohort	Live to Age 50	Live to Age 59
BA – SC	0.012	0.014
BA – HS	0.012	0.032
BA – LTHS	0.022	0.077
SC – HS	0.000	0.018
SC – LTHS	0.009	0.063
HS – LTHS	0.010	0.046
Mean Survival Rate for LTHS	0.958	0.874

Notes: Data are from the National Health Interview Surveys from 1986-1992 with linked NDI mortality data through 2011. BA: BA degree or greater. SC: some college. HS: high school degree. LTHS: less than a high school degree. Table presents differences in the probability of surviving to the given age by education level derived from estimates of OLS regression of an indicator for whether an individual died at age t on education level dummy variables, age dummy variables, the interaction between education and age dummy variables, female dummy variable, baseline self-reported health, and year dummy variables. Predicted values are calculated at the mean of covariates (female, survey year, and baseline self-reported health).