Education and Mortality: Evidence for the Silent Generation from Linked Census and Administrative Data

Ciprian Domnisoru* Anna Malinovskaya† Evan Taylor‡

July 31, 2024

Abstract

We quantify the effect of education on mortality using a linkage of the full count 1940, 2000, and 2010 US census files and the Numident death records file. Our sample is composed of children aged 0-18 in 1940, observed living with their parents, for whom we can construct a rich set of parental and neighborhood characteristics, including parental education and income. We estimate effects of educational attainment in 1940 on survival to 2000, as well as the effects of completed education, observed in 2000, on 10-year survival to 2010. The educational gradients in longevity that we estimate are robust to the inclusion of detailed individual, parental, household, neighborhood and county covariates. Given our full population Census sample, we also explore rich patterns of heterogeneity and examine the effect of mediators of the education-mortality relationship.

Keywords: education, health, mortaliy
JEL Codes: I1, I2, J1

DISCLAIMER: Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2816. (CBDRB-FY24-P2816-R11571, CBDRB-FY24-P2816-R11535 and CBDRB-FY24-P2816-R11325).

*Aalto University School of Business - Helsinki GSE, ciprian.domnisoru@aalto.fi
†Princeton University, Center for Health and Wellbeing, School of Public and International Affairs
‡University of Arizona, Eller College of Management

1 Introduction

Many papers have documented the strong correlation between education and mortality (Kitagawa and Hauser, 1968; Elo and Preston, 1996; Balaj et al., 2024). Evidence shows that the education-mortality gradient in the United States is getting steeper (Meara et al., 2008; Case and Deaton, 2021). While the correlation is well established, there is no consensus on causality and mechanisms. Education might impact behaviors like drinking and smoking (Cutler and Lleras-Muney, 2010; Conti et al., 2014), income (Mincer, 1974), neighborhood, marriage, social networks, and access to health care.

In this paper, we exploit a linkage between the 1940 U.S. Census, 2000 U.S. Census, 2010 U.S. Census and the NUMIDENT file to study the relationship between completed years of schooling and older age mortality. We have four main contributions.

First, we document the relationship between education and mortality in the United States using a large administrative data set. We find that the relationship is highly non-linear, with the largest decline in mortality at 11 to 12 years of schooling.[1]

Second, We find that controlling for detailed parental and neighborhood characteristics has almost no effect on the strength of the relationship. While this does not prove that the impact of education on mortality is causal, it is consistent with a causal relationship. Although we cannot account for differences in health in childhood that might both impact education and later life mortality, we definitively show that the correlation between education and later life health is *not* driven by differences in childhood socioeconomic status as measured by a rich set of parental, neighborhood, and local characteristics that we observe in the 1940 Census.

Third, we show detailed patterns of heterogeneity in the relationship. The impact of education is stronger among non-Hispanic Whites and people born into higher socioeconomic status, both in terms of their parent's education and in terms of their childhood neighborhood. The relationship is slightly stronger for people who lived in rural areas as children.

[1] In the United States, this is completed High School.

Lastly, we use a linkage to the 2000 Census to show how adult income, neighborhood, home ownership, living alone, and disability mediate the relationship between educational attainment and mortality. We find that these mediators account for about 75% of the impact of education on mortality. The relationship between education and mediators such as home ownership, neighborhood income, and disability all show non-linear patterns, with the largest jumps at 12 years of schooling. These results shed more light on the complicated relationship between education and mortality.

2 Literature Review

The most closely related study is Lleras-Muney et al. (2022), which links siblings and uses education data from the 1940 Census, death records and family tree data from familysearch.org. They find that each additional year of school is associated with 0.4 higher life expectancy at age 35. They also find that the relationship between education and mortality is unchanged when comparing siblings or twins, but is stronger for men in high-income places. While their paper focuses on earlier cohorts (main focus is 1906-1915 though they have some data on earlier cohorts), this paper analyzes the experiences of later cohorts (1924-1940). Educational attainment was increasing rapidly during this period of time, so our paper studies cohorts with substantially higher average education.[2] Second, this paper uses a linkage to the 2000 Census to see how the relationship between education and mortality is mediated through income, home ownership and disability.Lleras-Muney (2005) Finally, their paper only studied Whites because non-Whites are less likely to be linked through familysearch.org, whereas this paper studies both Whites and non-Whites. Finally, because we observe children in the 1940 Census, we have more detailed information about childhood socio-economic status, which allows us to conduct heterogeneity analyses at a household level and a geographic level.

[2]About 1/3rd of the 1906-1915 birth cohort graduated from high school compared to around 2/3rd for the 1924-1940 birth cohorts high school.

Another closely related study is Halpern-Manners et al. (2020), which utilizes linkages of 1920 and 1940 Census with the SSA NUMIDENT and Social Security Master Death File. Their main finding is that the relationship between years of education and age of death is modestly attenuated among twins and non-twin siblings. Our paper differs from that study in several ways. First, because they rely on linkages based on names for the 1920 and 1940 Censuses, they limit their study to men only, while we include women in our study. Second, our large sample size allows us to split our sample by many different characteristics to study effect heterogeneity, which they are unable to do because the twin sample is fairly small (only 5,000 individuals). Third, our paper studies the impact of each additional year of education rather than a linear impact of education,[3] and find a non-linear relationship between years of education and mortality, with the largest jump at 11 to 12 years of schooling. Lastly, we use a linkage to the 2000 Census to quantify the impact of mediators such as income, neighborhood and disability in adulthood to explore the potential mechanisms mediating the relationship between education and mortality.

Montez and Hayward (2011) shows that controlling for early life characteristics such as mother's and father's education and childhood socioeconomic status has little impact on the correlation between education and mortality. However, their paper used a much smaller sample (around 20,000) in the Health and Retirement Study and retrospective measures of childhood socioeconomic status.

Other papers have used instruments such as compulsory schooling laws (Lleras-Muney, 2005; Meghir and Palme, 2005; Clark and Royer, 2013; van Kippersluis et al., 2011), Vietnam draft (Buckles et al., 2016) or college expansion (Fletcher and Noghanibehambari, 2024) to identify a causal impact of education on mortality. These studies have found mixed results. This could be because they study different countries, different time periods, or different margins of education. For a comprehensive review see Galama et al. (2018).

Other studies use linked data from the 1940 Census and the Social Security Master Death

[3]The Halpern-Manners paper also breaks down educational attainment into a binary category for individuals who completed high school.

File. For example, Fletcher and Noghanibehambari (2024) studied how college expansions impact mortality and found positive impacts of college education on life expectancy.

Finally, several papers have studied twins to compare children with identical family upbringing and genetics. Behrman et al. (2011) studies twins in Denmark and finds that twins with different levels of education have similar mortality and health. Lundborg et al. (2016) uses Swedish data with a larger sample size (about 50,000) and finds that the relationship between education and mortality is almost unchanged when including a twin fixed-effect. Their paper also controls for early life health differences between twins such as height and birth weight. In the United States, both Lleras-Muney et al. (2022); Halpern-Manners et al. (2020) find that controlling for twin fixed-effects does not substantially reduce the association between education and mortality. The limitations of twin studies are discussed in Bound and Solon (1999).

3 Data

We use a linkage of the 1940 full-count US Census to the 2000 and 2010 Census files and the Social Security Administration NUMIDENT file. Our primary sample, described in Table 2, is composed of US-born children under the age of 16, observed in the 1940 Census living with at least one parent in the contiguous US excluding D.C. This sample is merged with the 2000 and 2010 Census and NUMIDENT using Protected Identification Keys (PIKs). Among the individuals assigned a PIK, we further limit the sample to observations with a unique PIK match. We observe a unique PIK for about 7,807,000 observations, or 64.57 per cent of the baseline sample. Observations with a PIK are slightly younger and more male[4], come from families with more educated parents, and are slightly more likely to be White Non-Hispanic.

Our main mortality measure is survival to 2010 conditional on survival to 2000. Survival to 2010 is an indicator variable created as the intersection of 1) observing a unique link

[4]PIK rates are generally higher for younger cohorts and men. The PIK-ing procedure and match rates for various demographics are detailed in Massey et al. (2018).

for individuals in the 2010 Census *or*, absent the census link, 2) observing a recorded date of death in NUMIDENT after 2010. Conditioning on survival in 2000 is necessary in our study because our main independent variable of interest is educational attainment observed in the 2000 Census. Thus, our main results study mortality at ages 60-85. In one alternative model the main independent variable of interest is educational attainment in 1940 for 15-18 year olds, and we use survival to 2000 as the outcome variable. We use this model to study mortality at younger ages (15-68). Survival to 2000 is imputed as 1) observing a unique link for individuals in the 2000 Census *or* 2) *absent a unique Census 2000 link*, observing a recorded date of death in NUMIDENT after 2000, or a unique match to the 2010 Census.

The main independent variable of interest is educational attainment by 2000, measured in completed years of education. In the 2000 Census long form, respondents were asked to choose a level of attainment from: 'No schooling completed', which is our reference category, 'Nursery school to 4th grade', which we label as '4 years of schooling', '5th grade or 6th grade', which we label '6 years of schooling', and '7th grade or 8th grade', recorded as '8 years of schooling'. We kept distinct categories for 9th, 10th and 11th grade attainment, as in the original 2000 long form questionnaire. For data privacy protection reasons[5] we found it necessary to combine several categories: '12th grade, no diploma' and 'High school graduate' are recorded as '12 years of schooling', 'Some college credit, but less than 1 year' is recorded as '13 years of schooling', 'One or more years of college, no degree' and 'Associated degree' are combined into '14 years of schooling', 'Bachelor's degree' is assigned 16 years, while all higher degrees are bundled into '17 years of schooling'. In all of our analyses, we use fixed effects to measure education rather than a continuous measure.

Five sets of control variables were constructed for the individuals in the main sample based on their personal, parental, household, neighborhood, and county characteristics observed in 1940. *Personal* control variables include information on race, ethnicity, sex, age, and state

[5]We faced a trade-off between providing comparisons across different races, genders, ethnicities, and geographical areas, and having enough meaningful observations to abide by cell count restrictions imposed by the Census Bureau and necessary for disclosure of results. We opted for combining educational categories.

of birth. *Parental* controls include information on mother's education in years of schooling and mother's age in five year intervals (these are replaced with father's characteristics if the mother is absent from the household), whether at least one of the parents is foreign born, and whether only the mother or only the father are present in the household. *Household* controls include information on the number of siblings, whether an individual lives on a farm, whether an individual's family moved in the past five years within the county, state, or from abroad, whether the household was located in a metropolitan area, and the population size of the place of residence.

We construct *neighborhood* characteristics using a feature of the 1940 Census records: the fact that the data was collected by enumerators who traveled in small contiguous geographic areas called enumeration districts and recorded information about households on numbered sheets. This enables us, for each child in our sample, to construct neighborhood measures based on proximity in the Census records.[6] The neighborhood controls include information on the average schooling and income of nearby fathers, average schooling and employment status of nearby mothers, fraction of neighbors living on farms, and fraction of nearby fathers on Emergency Employment, which we use as a proxy for the level of severity of the Great Depression in the individual's neighborhood.[7]

Finally, *county* controls include occupational scores separately for men and women aged 35-54, the fraction of births delivered in hospitals, doctor to population ratio for Whites, infant mortality rate for Whites, average percent of workers employed in agriculture, average density (persons per sq. mile), average percent of dwelling units with a private bathtub or shower, average percent of rural farm dwelling units with running water, average percent of rural farm dwelling units with electric lighting, total major war supply contracts and projects in millions of dollars, the minimal distance to a 4-year college, the minimal distance

[6]In practice, we use the information on the nearest 50 neighbors to define the 'neighborhood', by aggregating information on the closest 25 neighboring households up and down the Census roster line for each individual.

[7]Emergency Employment was a subsidized form of employment under the New Deal, recorded in the 1940 sample as 'At Work, public emergency'.

to a normal school, average number of children born to women aged 21-30, and an indicator variable for whether the county has a city with a population of at least 25,000.

4 Empirical Strategy

We start with a series of linear probability regression models in which S_i^{2010}, an indicator for survival until 2010, is regressed on educational attainment fixed effects in 2000, $\gamma_j ED_j$. In the basic specification, we include a vector X_i of controls for *individual* characteristics that include age in 1940, ethnicity, race, and state of birth. The omitted category is zero years of education in 2000. All controls are fully interacted with gender. The model also includes a vector of county fixed effects, $\theta_c(i)$. We cluster standard errors in all models in this study at the county level in 1940. Additionally, we use survey-provided person weights in all regressions in this study except for those where the outcome is survival to 2000.

$$S_i^{2010} = \sum_j (\gamma_j ED_j) + \beta X_i + \theta_c(i) + u_i \tag{1}$$

Table 1: Summary of main regression models

	Fixed effects	Covariates
Model 1	None	Person
Model 2		+Parent, household
Model 3		+ Neighborhood, County covariates x Female
Model 4	County x Gender FE	Person
Model 5		+Parent, household
Model 6		+ Neighborhood

We augment the basic model by gradually adding covariates. To the baseline set of individual characteristics X_i we add parental and household controls (model 2). Next, we add neighborhood and county-level controls interacted with gender (model 3). Next, we include county-by-gender fixed effects and re-estimate the three models (our county-level controls drop out with the addition of county-by-gender fixed effects). These model variations are summarized in Table 1.

We conduct an extensive set of heterogeneity analyses using the model with the full set of controls at the personal, parental, household, and neighborhood level and county-by-gender fixed effects (after we compare the results from the six models described in Table 1, this model - model 6 - becomes our preferred model and is used in all regressions unless specified otherwise; as we discuss in the Results section, the six models from Table 1 yield similar results). In particular, we estimate the model separately by gender, race and ethnicity, and by levels of parental educational attainment. Leveraging the neighborhood variables, we also run our model separately by state-level terciles of house values, father's income, parental education, and local farm status.

One of the mechanisms through which education may lead to higher rates of survival is its association with higher lifetime income. We estimate models in which we control for the log of total income in 2000 and separately also for social security income in 2000.[8] To further explore the income mechanism, we additionally estimate models in which the log of total income and the log of social security income are the outcome variables.

Exploring additional mediators in the relationship between educational attainment and mortality, we further estimate regressions with educational attainment in 2000 fixed effects as the main independent variable of interest as before but where the outcome variables are home ownership status in 2000, log of block group level household income in 2000, log of block group level home values in 2000, an indicator variable for whether an individual lived alone in 2000, and indicators for whether an individual reported having any difficulty, any physical difficulty, difficulty with performing mental tasks (learning, remembering, or concentrating), or difficulty working at a job or business.[9] Then, we re-estimate the effect of education on

[8]If total or social security income is reported to be zero or negative, we replace the missing values in the log of total or social security income respectively with a zero and include in our regressions a control variable for whether total or social security income was zero or negative.

[9]Census 2000 Long Form asked whether an individual had (1) "blindness, deafness, or a severe vision or hearing impairment"; (2) "a condition that substantially limits one or more basic physical activities such as walking, climbing stairs, reaching, lifting, or carrying"; (3) difficulty with "learning, remembering, or concentrating"; (4) difficulty with "dressing, bathing, or getting around inside the home"; (5) difficulty with "going outside the home alone to shop or visit a doctor's office"; and (6) difficulty with "working at a job or business". We define the any difficulties dummy as an indicator variable taking the value of one if the individual reports having any of these six conditions. The physical difficulty dummy is defined to equal one

survival with these mediators as additional controls.

While our main analysis relies on completed education observed for individuals surviving to 2000, we also present evidence on the impact of educational attainment in 1940 on the probability of survival to 2000. For this purpose, we use a sample of US born 15-18 year olds, observed living with their parents in the 1940 Census in the contiguous US excluding D.C.[10] For those individuals who are no longer enrolled in school, we assign the highest grade attained as their level of educational attainment. Since many teenagers in our sample can be expected to still be enrolled in school, we also use measures of *potential* educational attainment E^{1940}. For every individual still enrolled in school, we impute "on track" educational attainment as the grade currently attained plus two. We estimate a similar specification to equation (1), regressing the *potential* educational attainment imputed in the 1940 Census on survival to 2000. The equation includes the same set of controls X_i and is estimated for the six different models listed in Table 1.

$$S_i^{2000} = \sum_j (\gamma_j E_j^{1940}) + \beta X_i + \theta_c(i) + v_i \tag{2}$$

5 Results

Table 3 shows our estimates of the effect of education completed by 2000 on the 10-year survival probability for the six models we consider in Table 1, for individuals aged 0-15 in 1940, who were aged 60-75 in 2000. The mean of the 10-year survival rate is 0.733. The shape of the education gradient in longevity estimated for this population is also shown in Figure 1 for models that include county-by-gender fixed effects. One striking result is that the gradual addition of covariates does little to change the shape of the education gradient. The stability of the coefficients is an indication that there may be little omitted variable

if the respondent reports having conditions (1), (2), (4), or (5). The mental difficulties dummy is defined to equal one if the respondent reports having condition (3) and the work related difficulties dummy is defined to equal one if the respondent reports having condition (6).

[10] And having a unique PIK match in Census 2000.

bias (Oster (2019)). The gradient is fairly linear up to 10 years of schooling, and exhibits a flattening at 11 years of schooling, and a jump at 12 years of schooling, associated with high school completion. This pattern is consistent with the labor market effects of high school dropout and degree attainment, which have been observed to lead to nonlinear income effects (see, for example, Table 10). The gradient exhibits a smaller slope for individuals with some college or an associates degree, but picks back up again for individuals with a bachelor's degree and those with higher levels of educational achievement.

The gradient in Table 3 is estimated for the entire sample. The race-by-gender fixed effects show the lower overall survival rates of the Black and American Indian population, and the higher survival of ethnic Hispanics. Additionally, we show, for Model 6, the negative effect of living with only a mother in 1940, the positive effects of having a foreign-born parent, and the positive effects of living on a farm in 1940.[11] Overall, these results indicate that the effects of educational attainment on survival to 2010, conditional on survival to 2000, are strong and robust. In terms of magnitude, we find that, for example, moving from 8 years of schooling to 16 years of schooling leads to a 0.178 increase in the ten survival rate. A back-of-the-envelope calculation suggests that is roughly equivalent to 2.67 years of life expectancy at 2000, or about .33 years of life expectancy per year of school.[12] This gradient is similar to results found by Lleras-Muney et al. (2022), though smaller than results found by Meara et al. (2008).[13]

We are also able to examine the effect of the amount of *actual* and *potential* grade attainment in 1940 by 15-18 year old teenagers on their probability of survival to 2000. We show the effects of *actual* grade attainment for the entire population in Table 4, for

[11]While all our models include a rich set of controls, we do not report coefficients for every single control variable in this paper in order to minimize disclosure risks and the administrative burden associated with having the Census Bureau conduct a disclosure review of a large number of coefficients.

[12]The average person who does not live to 2010 will have 5 years of life expectancy at 2000. The average person that survives to 2010 would have about 10 more years of life expectancy based on social security life tables.

[13]Meara et al. (2008) find life expectancy for 25 year olds with at least some college is 5-8 years higher than those without any college. Our measure of life expectancy is at an older age, so it would not account for gaps in mortality at younger ages.

White teenagers in Table 5, and for Black teenagers in Table 6, and the effects of *potential* grade attainment for the full sample in Table A2, for White teenagers in Table A3, and for Black teenagers in Table A4. Similar to the gradient in completed education, the gradient in potential education remains largely unaffected by the gradual addition of covariates. Figures 2, 3, and 4 illustrate that the gradient for Whites appears U-shaped, however very few individuals in our sample have less than five years of schooling completed, so those are likely exceptional cases.[14] The slope past five years of potential schooling is linear and the magnitudes are considerable. For example, being on track to attain 12 years or more of schooling (which indicates either students had attained this level by 1940, or had attained 10 grades and were still enrolled in 1940) is associated with about a .15 percentage point increase in survival to 2000, which decreases only slightly when controlling for parental, household, and neighborhood characteristics.

For Black children, two patterns stand out. First, the gradient is linear even at lower levels of educational attainment. While having less than five years of schooling was an exception for White children, likely related to medical conditions or other exemptions from compulsory schooling laws, about 23 percent of Black 15 to 18 year olds in the public IPUMS 100% sample had attained at most fourth grade. It therefore appears that for Black children, most of whom were educated in the underfunded segregated schooling system, attaining even low levels of schooling led to gains in the probability of surviving to 2000. The second pattern that stands out for Black teenagers is that the magnitudes of the effects of educational attainment are lower. For example, being on track to complete ninth grade leads to a .1 increase in the survival probability for Whites, but only a 0.5 increase for Black teenagers. The lower magnitude may reflect that in the labor market, because of systemic discrimination, access to occupations with lower risk or occupational stress was not equally open to White and Black graduates with the same levels of educational attainment. As such, Black and White workers with similar educational levels were likely exposed to different work

[14] In the IPUMS public use sample, less than 3.7% of White individuals aged 15-18 have attained less than five years of schooling by age 15.

conditions and occupational hazards. This may also reflect the higher measurement error in the number of years of schooling for Black children, who were more likely to be educated in one-room schools, where the recording of grade information was less precise.

5.1 Heterogeneity analysis

In Table 7, we contrast the education-mortality gradient for 10-year survival from 2000 to 2010 across genders, races, ethnicities, and levels of parental education in 1940. The different gradients by gender are illustrated in Figure 6. The survival gains for women are generally higher than for men, with the exception of 16 years of schooling, where they equal estimates for men, and 17 or more years of schooling, where the gradient for men is actually higher than for women, although the confidence intervals overlap. The steeper gradient for men has been widely documented. Montez et al. (2009) discuss potential mechanisms behind this finding, pointing to gender differences in smoking patterns, given a strong education-by-gender interaction with mortality rates in lung cancer and chronic respiratory diseases.

Differences by race and ethnicity in the education gradient for 10-year survival are illustrated in Figure 7 for White Non-Hispanics, Blacks, and White Hispanics. The gradient is generally higher for White Non-Hispanics, except at lower levels of education, where confidence intervals overlap. We also explore heterogeneity by parental education background (Table 7). The education-mortality gradient is steeper for children with better-educated mothers (or fathers if the mother is not in the household). Unsurprisingly, children born into households with higher educated mothers also had higher survival rates.

While our results are very similar across the six models that gradually expand the set of controls, it is still possible that unobservable environmental and socio-economic factors might bias our estimates. To explore this further, we estimate the education-mortality gradient across different areas. In Table 8 we split our sample into state-level terciles of house values at the enumeration district level, father's income at the neighborhood level, and parent's education at the neighborhood level. Results are similar across these groups,

indicating that the shape of the education-mortality gradient is not fundamentally driven by the neighborhoods in which children grew up. In Table 9 we also compare the shape of the gradient for children growing up on farms and those from non-farm households, in rural vs. urban areas, and across terciles of the fraction of households in the 1940 neighborhood that were farms. Here again, the coefficients overlap across areas, with some differences in the lower tail of the educational distribution, where lower levels of educational attainment appear to have a statistically significant impact on the survival rate in more rural areas, but not in urban areas. This pattern is likely driven by selection, as very few individuals in urban areas acquired only four, six, or eight years of schooling, whereas such levels of attainment were more common in rural areas.[15]

5.2 Mediators

Income is an important possible mediator of the education-mortality relationship, as it can facilitate access to higher-quality care or goods and services that promote health.[16] Indeed, as shown in Table 10, there is a clear association between higher levels of education and total income as well as social security income. For social security income, the relationship is weaker, as not all individuals in our sample are collecting social security yet, and the maximal amounts of social security benefits are capped, creating ceiling effects for individuals with high levels of educational attainment. When we control for total and social security income, the educational gradient decreases only slightly in magnitude, and, in fact, confidence intervals with and without controls for income overlap. The total income effect is statistically significant. A 10 percentage point increase in income is associated with a 0.33 percentage point increase in the 10-year survival rate, roughly the same increase as moving from 14 years of schooling to 16 years.

[15]Using data from the public IPUMS 1940 100% sample, we find that among White 15-18 year olds living in urban areas, only 17 percent had attained eighth grade or less. In rural areas, that number stood at 42.3 percent.

[16]On the other hand, income may also allow for purchases of unhealthy goods such as alcohol and cigarettes. For more on the effect of income on health see Lleras-Muney (2022).

Beyond income, we consider a series of other mediators that can be observed in the 2000 Census long form. In Table 11 we examine the impact of educational attainment on homeownership, the median household income and median housing values at the block level of the individual's house in 2000, living alone, and reporting having physical or mental difficulties. In Table 12, we estimate the education-mortality gradient including these mediators as controls. Homeowership is associated with about a six percentage point higher probability of survival to 2010, and educational attainment is strongly correlated with homeownership. Compared to individuals with less than four years of schooling, college graduates are 20 percentage points more likely to own a home, and individuals with nine years of schooling are 9.4 percentage points more likely. Living in a Census block where neighborhood income or house values are 10 percentage points higher is associated with a 0.13 and 0.14 percentage point higher survival rates, respectively, even after controlling for individual income and educational attainment. This pattern is could be related to the quality of local public amenities and public health resources. In turn, education is highly associated with living in areas with higher housing and household income values. For example, college graduates live in areas where house values are 45 percent higher than in the areas where those with less than four years of schooling live. Relative to the same reference group, high school graduates live in census block groups where house values are only 15 percent higher. Living alone is associated with a lower survival probability (estimates in Table 12 are around -2.4 to -4.1 percentage points). Conversely, individuals with higher levels of educational attainment are less likely to be observed living alone in the 2000 Census. Turning to physical and mental difficulties, they are generally associated with lower survival rates. Individuals with higher educational attainment are also less likely to be reporting mental or physical difficulties.

Once we control for all the mediators examined in Table 11, we find that education continues to have a statistically significant effect on the 10-year survival rate, for 12 years of schooling and above. Relative to the initial magnitudes examined in Table 3, mediators (total income, homeownership, living alone, presence of physical and mental difficulties,

neighborhood house and household income values) explain about 75 percent of the education coefficient for high school graduates and 70 percent of the coefficient for bachelor's degree holders. For people with lower than 12 years of schooling, the mediators explain almost the entire relationship between education and mortality. Interpretation of these results is complicated. Many of these mediators (such as income) *impact* health, but are also *measures* of health (i.e. people that are healthier in their working adult life are likely to have higher income in their 60s and 70s.) However, it is clear from these results that income, disability, neighborhood, and living alone all play important roles in the impact of education on mortality.

6 Conclusion

We examine the education-mortality gradient in the 10-year survival rate for cohorts born 1922-1940, commonly referred to as "The Silent Generation." Advances in linkages between US Census files (Massey et al. (2018)) allow us to observe the old age survival patterns of children recorded living with their parents in the 1940 full Census sample. We find the relationship between educational attainment and mortality to be remarkably stable to a rich set of controls for the characteristics of individuals' parents and childhood environmental circumstances. As we are observing individuals in the 2000 Census long form, we are also able to present evidence on the potential mediators of the education-mortality relationship. About 70-75% of the effect of education on the 10-year survival rate is explained by the effect of education on lifetime income, home-ownership, likelihood of living alone, mental and physical difficulties and Census block characteristics. In future work, likely facilitated by further advances in Census file linkages, we hope to explore the role of other mediators of the education-mortality relationship, such as occupational choice, marriage, divorce and fertility.

References

Balaj, M., C. A. Henson, A. Aronsson, A. Aravkin, K. Beck, C. Degail, L. Donadello, K. Eikemo, J. Friedman, A. Giouleka, et al. (2024). Effects of education on adult mortality: a global systematic review and meta-analysis. The Lancet Public Health.

Behrman, J. R., H.-P. Kohler, V. M. Jensen, D. Pedersen, I. Petersen, P. Bingley, and K. Chistensen (2011). Does more schooling reduce hospitalization and delay mortality? new evidence based on danish twins. Demography 48(4), 1347–1375.

Bound, J. and G. Solon (1999). Double trouble: on the value of twins-based estimation on the return to schooling. Economics of Education Review 18(2), 169–182.

Buckles, K., A. Hagemann, O. Malamud, M. Morrill, and A. Wozniak (2016). The effect of college education on mortality. Journal of Health Economics 50, 99–114.

Case, A. and A. Deaton (2021). Life expectancy in adulthood is falling for those without a ba degree, but as educational gaps have widened, racial gaps have narrowed. Proceedings of the National Academy of Sciences 118(11).

Clark, D. and H. Royer (2013). The effect of education on adult mortality and health: Evidence from britain. American Economic Review 103(6), 2087–2120.

Conti, G., J. Heckman, and S. Urzua (2014). The education-health gradient. American Economic Review: Papers and Proceedings 100(2), 234–238.

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

Elo, I. T. and S. H. Preston (1996). Educational differentials in mortality: United states, 1979-1985. Social Science Medicine 42(1), 47–57.

Fletcher, J. and H. Noghanibehambari (2024). he effects of education on mortality: Evidence using college expansions. Health Economics 33(3).

Galama, T. J., A. Lleras-Muney, and H. van Kippersluis (2018, January). The effect of education on health and mortality: A review of experimental and quasi-experimental evidence. Working Paper 24225, National Bureau of Economic Research.

Halpern-Manners, A., J. Helgertz, J. R. Warren, and E. Roberts (2020). The effects of education on mortality: Evidence from linked u.s. census and administrative mortality data. Demgraphy 57(4), 1513–1541.

Kitagawa, E. and P. Hauser (1968). Education differentials in mortality by cause of death: United states, 1960. Demography 5(1), 318–353.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the united states. Review of Economic Studies 72, 189–221.

Lleras-Muney, A. (2022). Education and income gradients in longevity: The role of policy. NBER Working Paper Series (29694).

Lleras-Muney, A., J. Price, and D. Yue (2022). The association between educational attainment and longevity using individual-level data from the 1940 census. Journal of Health Economics 84, 102649.

Lundborg, P., C. H. Lyttkens, and P. Nystedt (2016, 07). The Effect of Schooling on Mortality: New Evidence From 50,000 Swedish Twins. Demography 53(4), 1135–1168.

Massey, C. G., K. R. Genadek, J. T. Alexander, T. K. Gardner, and A. O'Hara (2018). Linking the 1940 u.s. census with modern data. Historical Methods: A Journal of Quantitative and Interdisciplinary History 51(4), 246–257.

Meara, E. R., S. Richards, and D. M. Cutler (2008). The gap gets bigger: Changes in mortality and life expectancy, by education, 1981-2000. Health Affairs 27(2), 350–360.

Meghir, C. and M. Palme (2005). Educational reform, ability, and family background. American Economic Review 95(1), 414–424.

Mincer, J. (1974). Schooling, experience, and earnings. human behavior & social institutions no. 2.

Montez, J. K. and M. D. Hayward (2011). Early Life Conditions and Later Life Mortality, pp. 187–206. Dordrecht: Springer Netherlands.

Montez, J. K., M. D. Hayward, D. C. Brown, and R. A. Hummer (2009, 05). Why Is the Educational Gradient of Mortality Steeper for Men? The Journals of Gerontology: Series B 64B(5), 625–634.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37(2), 187–204.

van Kippersluis, H., O. O'Donnell, and E. van Doorslaer (2011). Long-run returns to education. Journal of Human Resources 46(4), 695 – 721.

7 Tables and Figures

Table 2: Summary statistics

	US-born children under 16, living with at least one parent in 1940							
			Subsample assigned a unique protected identification key (PIK)					
					Subsample surviving to 2000			
							Subsample surviving to 2010	
	N	Mean	N	Mean	N	Mean	N	Mean
Age	12,090,000	7.604	7,807,000	7.424	5,816,000	6.912	4,353,000	6.342
American Indian	12,090,000	0.002662	7,807,000	0.001825	5,816,000	0.00147	4,353,000	0.0013
Asian	12,090,000	0.001397	7,807,000	0.001137	5,816,000	0.001171	4,353,000	0.00126
Black	12,090,000	0.09268	7,807,000	0.06162	5,816,000	0.04918	4,353,000	0.04435
White Hispanic	12,090,000	0.01749	7,807,000	0.01459	5,816,000	0.01436	4,353,000	0.01465
White Non-Hispanic	12,090,000	0.8858	7,807,000	0.9208	5,816,000	0.9338	4,353,000	0.9384
Female	12,090,000	0.4917	7,807,000	0.4812	5,816,000	0.5163	4,353,000	0.5401
Only father in household	12,090,000	0.02047	7,807,000	0.01463	5,816,000	0.01276	4,353,000	0.01133
Only mother in household	12,090,000	0.09391	7,807,000	0.06706	5,816,000	0.06037	4,353,000	0.05577
Foreign-born parent	12,090,000	0.1217	7,807,000	0.131	5,816,000	0.132	4,353,000	0.1323
Father's years of schooling	10,890,000	8.743	7,241,000	9.044	5,434,000	9.218	4,087,000	9.373
Mother's years of schooling	11,800,000	9.079	7,665,000	9.372	5,721,000	9.536	4,288,000	9.677
Mother employed	11,840,000	0.1252	7,693,000	0.1108	5,742,000	0.1048	4,304,000	0.1004
Father employed	10,950,000	0.9191	7,284,000	0.9242	5,465,000	0.9281	4,110,000	0.931
Average schooling, nearby fathers	12,090,000	8.484	7,807,000	8.709	5,816,000	8.78	4,353,000	8.826
Average schooling, nearby mothers	12,090,000	8.837	7,807,000	9.026	5,816,000	9.086	4,353,000	9.123
Average income, nearby fathers	12,090,000	1178	7,807,000	1239	5,816,000	1,256	4,353,000	1,268
Average employment rate, nearby mothers	12,090,000	0.1331	7,807,000	0.1297	5,816,000	0.128	4,353,000	0.1274
Fraction of neighbors living on farms	12,090,000	0.2345	7,807,000	0.2112	5,816,000	0.2095	4,353,000	0.2076
Fraction fathers on emergency employment	12,090,000	0.05789	7,807,000	0.05622	5,816,000	0.05525	4,353,000	0.05442

Number of observations rounded in accordance with the Census Bureau rules for disclosure.

Table 3: Effect of education on survival to 2010, conditional on survival to 2000

	Dependent variable: Survival to 2010					
	Without county FEs			With county by gender FEs		
	person covs	**pers, parent and HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	2	3	4	5	6
4 yrs. of schooling	0.019	0.019	0.019	0.017	0.018	0.017
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
6 yrs. of schooling	0.046***	0.046***	0.046***	0.045***	0.045***	0.045***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
8 yrs. of schooling	0.073***	0.071***	0.072***	0.071***	0.070***	0.070***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
9 yrs. of schooling	0.080***	0.080***	0.081***	0.079***	0.080***	0.080***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
10 yrs. of schooling	0.094***	0.095***	0.096***	0.094***	0.095***	0.095***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
11 yrs. of schooling	0.096***	0.098***	0.098***	0.097***	0.098***	0.098***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
12 yrs. of schooling	0.160***	0.161***	0.162***	0.161***	0.161***	0.161***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
13 yrs. of schooling	0.187***	0.189***	0.189***	0.188***	0.189***	0.189***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
14 yrs. of schooling	0.186***	0.190***	0.190***	0.188***	0.190***	0.190***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
16 yrs. of schooling	0.227***	0.229***	0.229***	0.228***	0.229***	0.229***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
17 yrs. of schooling	0.246***	0.248***	0.248***	0.248***	0.248***	0.248***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Am. Indian x Female	0.006	-0.005	-0.003	0.004	-0.003	-0.007
	(0.029)	(0.029)	(0.029)	(0.028)	(0.028)	(0.028)
Am. Indian	-0.080***	-0.079***	-0.075***	-0.091***	-0.088***	-0.084***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)
Asian x Female	0.017	0.020	0.021	0.024	0.023	0.021
	(0.026)	(0.026)	(0.027)	(0.027)	(0.027)	(0.027)
Asian	0.034	0.021	0.020	0.031	0.019	0.021
	(0.024)	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)
Black x Female	0.014**	0.005	-0.001	0.008	0.002	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Black	-0.051***	-0.042***	-0.038***	-0.047***	-0.040***	-0.036***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)

Continued on next page

	Dependent variable: Survival to 2010					
	Without county FEs			With county by gender FEs		
	person covs	**pers, parent and HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	2	3	4	5	6
	Continued from previous page					
Hispanic x Female	0.013	0.009	0.010	0.015	0.011	0.009
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Hispanic	0.029***	0.025**	0.025**	0.027***	0.022**	0.023**
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Female	0.048***	0.072***	0.121***			
	(0.003)	(0.017)	(0.029)			
Only mother present						-0.020***
						(0.004)
Only mother present x Female						0.008
						(0.004)
Only father present						-0.012
						(0.007)
Only father present x Female						0.001
						(0.009)
Foreign-born parent						0.022***
						(0.003)
Foreign-born parent x Female=1						0.005
						(0.004)
Living on farm in 1940						0.010**
						(0.004)
Living on farm in 1940 x Female						0.003
						(0.005)
Observations	974,000	974,000	974,000	974,000	974,000	974,000
Adjusted R-squared	0.063	0.064	0.064	0.065	0.066	0.066
Y-mean	0.733	0.733	0.733	0.733	0.733	0.733

Sample is composed of US-born individuals who were under 16 years old and leaving with at least one parent in 1940, and who were assigned a unique PIK.

See Table 1 for details on the sets of covariates across specifications.

Number of observations was rounded in accordance with the Census Bureau rules for disclosure.

Standard errors were clustered at the 1940 county level.

** p<0.01; *** p<0.001

Table 4: Effect of years of schooling on survival to 2000, individuals aged 15-18 in 1940, all races

	Dependent variable: Survival to 2000[a]					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
1 year of schooling	0.045***	0.045***	0.043***	0.046***	0.045**	0.044**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
2 yrs. of schooling	0.014	0.014	0.013	0.015*	0.015*	0.014*
	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
3 yrs. of schooling	0.010	0.012	0.011	0.012	0.013*	0.012
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
4 yrs. of schooling	-0.003	-0.001	-0.001	0.002	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
5 yrs. of schooling	-0.011	-0.007	-0.006	-0.009	-0.006	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
6 yrs. of schooling	0.004	0.008	0.010	0.006	0.009	0.011
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
7 yrs. of schooling	0.024***	0.026***	0.028***	0.025***	0.027***	0.029***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
8 yrs. of schooling	0.061***	0.060***	0.061***	0.061***	0.059***	0.060***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
9 yrs. of schooling	0.093***	0.092***	0.093***	0.095***	0.092***	0.093***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
10 yrs. of schooling	0.116***	0.112***	0.113***	0.118***	0.113***	0.113***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
11 yrs. of schooling	0.142***	0.137***	0.136***	0.144***	0.137***	0.137***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
12+ yrs. of schooling	0.165***	0.157***	0.156***	0.167***	0.158***	0.158***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	2,386,000	2,386,000	2,386,000	2,386,000	2,386,000	2,386,003
Adjusted R-squared	0.054	0.057	0.058	0.058	0.059	0.059
Y-mean	0.549	0.549	0.549	0.549	0.549	0.549

Sample is composed of US-born individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
* p<0.01; *** p<0.001

Table 5: Effect of years of schooling on survival to 2000, Whites aged 15-18 in 1940

	Dependent variable: Merged2000ordiedafter					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
1 year of schooling	0.065***	0.066***	0.065***	0.067***	0.066***	0.066***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
2 yrs. of schooling	0.025**	0.026**	0.026**	0.026**	0.027**	0.027**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
3 yrs. of schooling	0.010	0.014	0.014	0.012	0.015*	0.015*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
4 yrs. of schooling	-0.014*	-0.008	-0.007	-0.011	-0.006	-0.006
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
5 yrs. of schooling	-0.023**	-0.017*	-0.016*	-0.021**	-0.016*	-0.015*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
6 yrs. of schooling	-0.001	0.003	0.005	0.000	0.004	0.006
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
7 yrs. of schooling	0.026***	0.028***	0.029***	0.027***	0.029***	0.030***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
8 yrs. of schooling	0.067***	0.064***	0.065***	0.066***	0.063***	0.064***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
9 yrs. of schooling	0.100***	0.097***	0.097***	0.100***	0.097***	0.097***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
10 yrs. of schooling	0.123***	0.118***	0.117***	0.124***	0.119***	0.118***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
11 yrs. of schooling	0.150***	0.144***	0.142***	0.151***	0.144***	0.143***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
12+ yrs. of schooling	0.173***	0.165***	0.163***	0.175***	0.166***	0.164***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
Observations	2,238,000	2,238,000	2,238,000	2,238,000	2,238,000	2,238,000
Adjusted R-squared	0.048	0.051	0.052	0.049	0.052	0.053
Y-mean	0.559	0.559	0.559	0.559	0.559	0.559

Sample is composed of US-born White individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
* p<0.01; *** p<0.001

Table 6: Effect of years of schooling on survival to 2000, Blacks aged 15-18 in 1940

	Dependent variable: Merged2000ordiedafter					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
1 year of schooling	-0.021	-0.026	-0.026	-0.023	-0.026	-0.026
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
2 yrs. of schooling	-0.011	-0.016	-0.016	-0.012	-0.015	-0.015
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
3 yrs. of schooling	0.005	-0.000	0.000	0.004	-0.000	0.000
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
4 yrs. of schooling	0.007	0.003	0.003	0.006	0.003	0.003
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
5 yrs. of schooling	0.016	0.012	0.012	0.018	0.014	0.014
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
6 yrs. of schooling	0.029**	0.027*	0.028*	0.033**	0.029**	0.030**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
7 yrs. of schooling	0.029**	0.029**	0.030**	0.035**	0.031**	0.032**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)
8 yrs. of schooling	0.044***	0.050***	0.050***	0.053***	0.050***	0.051***
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
9 yrs. of schooling	0.053***	0.059***	0.059***	0.063***	0.059***	0.060***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
10 yrs. of schooling	0.070***	0.076***	0.076***	0.083***	0.078***	0.078***
	(0.011)	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)
11 yrs. of schooling	0.086***	0.090***	0.090***	0.100***	0.092***	0.091***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
12+ yrs. of schooling	0.106***	0.106***	0.106***	0.119***	0.107***	0.107***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Observations	140,000	140,000	140,000	140,000	140,000	140,000
Adjusted R-squared	0.037	0.040	0.041	0.038	0.041	0.041
Y-mean	0.381	0.381	0.381	0.381	0.381	0.381

Sample is composed of US-born Black individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table 7: Effect of education on survival to 2010, conditional on survival to 2000, by select characteristics

						Parental education				
	Male	Female	Black	White Hispanic	White Non-hispanic	13-17 yrs. schooling	12 yrs. schooling	9-11 yrs. schooling	8 yrs. schooling	Less than 8 yrs. schooling
Sample number	7a	7b	7c	7d	7e	7f	7g	7h	7i	7j
4 yrs. of schooling	0.010	0.025	-0.012	0.041	0.020	0.021	0.022	0.029	-0.038	0.024
	(0.017)	(0.022)	(0.035)	(0.039)	(0.016)	(0.076)	(0.050)	(0.038)	(0.032)	(0.019)
6 yrs. of schooling	0.038**	0.054***	0.036	0.089**	0.045***	-0.017	0.056	0.027	0.020	0.052***
	(0.014)	(0.016)	(0.025)	(0.031)	(0.012)	(0.057)	(0.037)	(0.028)	(0.024)	(0.015)
8 yrs. of schooling	0.060***	0.086***	0.049*	0.057	0.084***	0.070	0.078**	0.074***	0.056**	0.069***
	(0.012)	(0.014)	(0.024)	(0.033)	(0.011)	(0.040)	(0.028)	(0.022)	(0.020)	(0.013)
9 yrs. of schooling	0.067***	0.099***	0.018	0.011	0.101***	0.061	0.095***	0.093***	0.080***	0.067***
	(0.012)	(0.014)	(0.024)	(0.034)	(0.011)	(0.040)	(0.028)	(0.021)	(0.020)	(0.014)
10 yrs. of schooling	0.089***	0.108***	0.049*	0.127***	0.112***	0.081*	0.110***	0.111***	0.083***	0.093***
	(0.012)	(0.014)	(0.024)	(0.035)	(0.011)	(0.037)	(0.027)	(0.021)	(0.020)	(0.013)
11 yrs. of schooling	0.082***	0.118***	0.058*	0.083**	0.115***	0.124**	0.118***	0.122***	0.083***	0.084***
	(0.012)	(0.013)	(0.023)	(0.030)	(0.011)	(0.039)	(0.027)	(0.021)	(0.020)	(0.014)
12 yrs. of schooling	0.144***	0.183***	0.105***	0.117***	0.180***	0.171***	0.191***	0.181***	0.153***	0.143***
	(0.011)	(0.013)	(0.021)	(0.025)	(0.010)	(0.034)	(0.026)	(0.020)	(0.019)	(0.012)
13 yrs. of schooling	0.178***	0.207***	0.134***	0.132***	0.208***	0.203***	0.216***	0.207***	0.184***	0.172***
	(0.011)	(0.013)	(0.023)	(0.032)	(0.010)	(0.034)	(0.026)	(0.021)	(0.020)	(0.014)
14 yrs. of schooling	0.178***	0.207***	0.149***	0.156***	0.207***	0.212***	0.216***	0.212***	0.179***	0.162***
	(0.011)	(0.013)	(0.022)	(0.026)	(0.010)	(0.034)	(0.025)	(0.020)	(0.019)	(0.013)
16 yrs. of schooling	0.230***	0.232***	0.184***	0.184***	0.247***	0.255***	0.252***	0.251***	0.217***	0.193***
	(0.011)	(0.013)	(0.022)	(0.030)	(0.010)	(0.034)	(0.026)	(0.021)	(0.019)	(0.014)
17 yrs. of schooling	0.254***	0.240***	0.199***	0.185***	0.266***	0.275***	0.274***	0.261***	0.231***	0.221***
	(0.011)	(0.013)	(0.023)	(0.030)	(0.010)	(0.034)	(0.026)	(0.021)	(0.019)	(0.014)
Observations	472,000	502,000	41,000	12,500	918,000	158,000	233,000	233,000	215,000	135,000
Adjusted R-squared	0.071	0.048	0.064	0.055	0.065	0.064	0.056	0.062	0.069	0.069
Y-mean	0.697	0.767	0.666	0.747	0.736	0.775	0.768	0.728	0.705	0.673

Sample is composed of US-born individuals under the age of 16, observed in 1940 living with at least a parent, who were assigned a unique PIK and linked to the 2000 Census or observed in Numident with a death record after 2000.
Model 6 from Table 1 is used in all regressions. Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table 8: Effect of education on survival to 2010, conditional on survival to 2000, by place of residence in 1940 (part I)

	Dep. Var.: Survival to 2010, conditional on survival to 2000								
	Model: with county by gender fixed effects								
	Terciles of house values (1940) at the ED level:			Terciles of fathers' income in the 1940 neighborhood:			Terciles of parents' years of education in the 1940 neighborhood:		
	1	2	3	1	2	3	1	2	3
Sample	8a	8b	8c	8d	8e	8f	8h	8i	8j
4 yrs. of schooling	-0.002	0.056*	0.021	0.011	0.009	0.072*	0.012	0.012	0.046
	(0.019)	(0.023)	(0.030)	(0.018)	(0.023)	(0.033)	(0.017)	(0.026)	(0.035)
6 yrs. of schooling	0.042**	0.051**	0.041	0.056***	0.040*	0.021	0.047***	0.052*	0.019
	(0.015)	(0.018)	(0.024)	(0.014)	(0.020)	(0.027)	(0.014)	(0.020)	(0.028)
8 yrs. of schooling	0.069***	0.079***	0.048*	0.080***	0.060***	0.048*	0.065***	0.084***	0.045*
	(0.013)	(0.015)	(0.019)	(0.012)	(0.016)	(0.022)	(0.012)	(0.016)	(0.021)
9 yrs. of schooling	0.070***	0.090***	0.077***	0.084***	0.074***	0.074***	0.078***	0.086***	0.072***
	(0.013)	(0.015)	(0.019)	(0.013)	(0.017)	(0.021)	(0.012)	(0.017)	(0.021)
10 yrs. of schooling	0.089***	0.108***	0.086***	0.106***	0.080***	0.098***	0.091***	0.107***	0.091***
	(0.013)	(0.014)	(0.019)	(0.012)	(0.016)	(0.023)	(0.012)	(0.016)	(0.021)
11 yrs. of schooling	0.080***	0.123***	0.091***	0.092***	0.098***	0.107***	0.093***	0.107***	0.098***
	(0.014)	(0.015)	(0.020)	(0.013)	(0.016)	(0.022)	(0.012)	(0.016)	(0.021)
12 yrs. of schooling	0.146***	0.181***	0.155***	0.162***	0.158***	0.166***	0.150***	0.175***	0.167***
	(0.012)	(0.014)	(0.018)	(0.012)	(0.015)	(0.021)	(0.011)	(0.016)	(0.020)
13 yrs. of schooling	0.175***	0.207***	0.185***	0.187***	0.186***	0.195***	0.175***	0.204***	0.197***
	(0.013)	(0.014)	(0.018)	(0.012)	(0.015)	(0.021)	(0.012)	(0.016)	(0.020)
14 yrs. of schooling	0.177***	0.206***	0.186***	0.185***	0.184***	0.200***	0.178***	0.200***	0.200***
	(0.012)	(0.014)	(0.018)	(0.012)	(0.015)	(0.021)	(0.011)	(0.016)	(0.020)
16 yrs. of schooling	0.208***	0.247***	0.228***	0.217***	0.225***	0.241***	0.208***	0.245***	0.239***
	(0.012)	(0.014)	(0.018)	(0.012)	(0.015)	(0.021)	(0.011)	(0.016)	(0.020)
17 yrs. of schooling	0.230***	0.265***	0.246***	0.240***	0.240***	0.260***	0.230***	0.256***	0.261***
	(0.013)	(0.014)	(0.019)	(0.012)	(0.015)	(0.021)	(0.012)	(0.016)	(0.020)
Observations	320,000	320,000	319,000	325000	325,000	325,000	325,000	325,000	325,000
Adjusted R-squared	0.071	0.068	0.063	0.071	0.068	0.065	0.068	0.07	0.066
Y-mean	0.724	0.731	0.742	0.724	0.728	0.746	0.723	0.730	0.747

Sample is composed of US-born individuals aged 0-15 in 1940, with at least one parent present, who were assigned a unique PIK and observed in the 2000 Census or in the 2010 Census or in NUMIDENT with death date later than 2000.
Model 6 from Table 1 is used in all regressions.Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table 9: Effect of education on survival to 2010, conditional on survival to 2000, by place of residence in 1940 (part II)

	Dep. Var.: Survival to 2010, conditional on survival to 2000						
	Model: with county by gender fixed effects						
	Own farm status in 1940:		Rural/urban status in 1940:		Terciles of farm status in the 1940 neighborhood:		
	non-farm	farm	rural	urban	1	2	3
Sample	8k	8l	8m	8n	8o	8p	8q
4 yrs. of schooling	0.015	0.022	0.028	0.004	0.0022	0.040	0.027
	(0.018)	(0.021)	(0.016)	(0.024)	(0.021)	(0.034)	(0.021)
6 yrs. of schooling	0.033*	0.061***	0.054***	0.031	0.038*	0.035	0.057***
	(0.013)	(0.017)	(0.013)	(0.018)	(0.016)	(0.027)	(0.016)
8 yrs. of schooling	0.049***	0.094***	0.089***	0.028	0.043**	0.073**	0.095***
	(0.011)	(0.015)	(0.011)	(0.015)	(0.013)	(0.024)	(0.014)
9 yrs. of schooling	0.070***	0.094***	0.095***	0.054***	0.068***	0.088***	0.094***
	(0.011)	(0.015)	(0.012)	(0.015)	(0.013)	(0.024)	(0.015)
10 yrs. of schooling	0.089***	0.106***	0.105***	0.079***	0.089***	0.091***	0.107***
	(0.011)	(0.015)	(0.011)	(0.014)	(0.012)	(0.025)	(0.014)
11 yrs. of schooling	0.093***	0.106***	0.107***	0.0817***	0.0918***	0.101***	0.106***
	(0.011)	(0.015)	(0.011)	(0.014)	(0.013)	(0.024)	(0.015)
12 yrs. of schooling	0.156***	0.169***	0.170***	0.145***	0.155***	0.155***	0.172***
	(0.010)	(0.014)	(0.011)	(0.013)	(0.012)	(0.022)	(0.013)
13 yrs. of schooling	0.186***	0.192***	0.196***	0.175***	0.183***	0.191***	0.198***
	(0.011)	(0.014)	(0.011)	(0.013)	(0.012)	(0.023)	(0.014)
14 yrs. of schooling	0.187***	0.187***	0.195***	0.176***	0.186***	0.183***	0.194***
	(0.010)	(0.014)	(0.011)	(0.013)	(0.012)	(0.023)	(0.014)
16 yrs. of schooling	0.227***	0.221***	0.232***	0.216***	0.226***	0.215***	0.231***
	(0.011)	(0.014)	(0.011)	(0.014)	(0.012)	(0.023)	(0.014)
17 yrs. of schooling	0.246***	0.235***	0.246***	0.238***	0.246***	0.249***	0.239***
	(0.011)	(0.014)	(0.011)	(0.014)	(0.012)	(0.023)	(0.014)
Observations	745,000	228,000	475,000	499,000	611,000	96,000	267,000
Adjusted R-squared	0.065	0.079	0.071	0.064	0.064	0.09	0.074
Y-mean	0.735	0.726	0.730	0.735	0.736	0.725	0.729

Sample is composed of US-born individuals aged 0-15 in 1940, with at least one parent present, who were assigned a unique PIK and observed in the 2000 Census or in the 2010 Census or in NUMIDENT with death date later than 2000. Model 6 from Table 1 is used in all regressions. Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table 10: Effect of education on survival to 2010, conditional on surviving to 2000, controlling for measures of income observed in 2000 Census

	Dep. Var.: Survival to 2010					
	Specification:				Dep. Var:	Dep. Var.:
	Non-missing income	Controlling for log of total income	Non-missing social security income	Controlling for log of SS income	Log total income	Log SS income
4 yrs. of schooling	0.016	0.018	0.016	0.017	0.174*	0.080
	(0.014)	(0.013)	(0.014)	(0.014)	(0.071)	(0.105)
6 yrs. of schooling	0.044***	0.043***	0.044***	0.045***	0.262***	0.281***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.060)	(0.085)
8 yrs. of schooling	0.0696***	0.063***	0.070***	0.070***	0.464***	0.470***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.052)	(0.070)
9 yrs. of schooling	D	0.071***	D	0.080***	0.548***	0.559***
	D	(0.009)	D	(0.009)	(0.054)	(0.073)
10 yrs. of schooling	0.095***	0.085***	0.095***	0.095***	0.589***	0.662***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.053)	(0.071)
11 yrs. of schooling	D	0.087***	D	0.098***	0.649***	0.595***
	D	(0.009)	D	(0.009)	(0.053)	(0.070)
12 yrs. of schooling	0.161***	0.144***	0.161***	0.161***	0.886***	0.611***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.050)	(0.068)
13 yrs. of schooling	0.189***	0.165***	0.189***	0.188***	1.255***	0.583***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.052)	(0.072)
14 yrs. of schooling	0.189***	0.165***	0.189***	0.188***	1.263***	0.453***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.051)	(0.068)
16 yrs. of schooling	0.229***	0.194***	0.229***	0.227***	1.595***	0.147*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.050)	(0.069)
17 yrs. of schooling	0.247***	0.204***	0.247***	0.245***	2.020***	-0.248***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.051)	(0.072)
Log total income		0.033***				
		(0.001)				
Log SS income				0.008***		
				(0.001)		
Observations	974,000	974,000	974,000	974,000	974000	974,000
Adjusted R-squared	0.066	0.07	0.066	0.066	0.138	0.406
Y-mean	0.733	0.733	0.733	0.733	9.439	5.826

Sample is composed of US-born individuals aged 0-15 in 1940, with at least one parent present, who were assigned a unique PIK and observed in the 2000 Census or in the 2010 Census or in NUMIDENT with death date later than 2000.
Model 6 from Table 1 is used in all regressions. Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001
We assign to individuals with zero or negative total income or social security income the loginctot value of zero and the logssinc value of zero, respectively. Thus, loginctot and logssinc will only contain missing values if total income or social security income information is missing, respectively.

Table 11: Effect of education on mediators of mortality observed in 2000 Census

	Dependent variable:							
	Homeowner status	Log of block group median household income	Log of block group median housing value	Any mental or physical difficulties?	Physical difficulties?	Mental difficulties?	Work related difficulties?	Lives alone dummy
Sample	10a	10b	10c	10d	10e	10f	10g	10h
4 yrs. of schooling	0.018	0.011	-0.005	-0.042**	-0.051***	-0.070***	-0.003	0.010
	(0.012)	(0.012)	(0.014)	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)
6 yrs. of schooling	0.041***	0.024*	0.011	-0.084***	-0.095***	-0.143***	-0.066***	-0.003
	(0.010)	(0.009)	(0.012)	(0.010)	(0.011)	(0.010)	(0.010)	(0.009)
8 yrs. of schooling	0.086***	0.071***	0.048***	-0.141***	-0.145***	-0.217***	-0.102***	-0.027***
	(0.009)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
9 yrs. of schooling	0.094***	0.070***	0.053***	-0.165***	-0.167***	-0.242***	-0.120***	-0.029***
	(0.009)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
10 yrs. of schooling	0.105***	0.091***	0.072***	-0.198***	-0.199***	-0.261***	-0.131***	-0.045***
	(0.009)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
11 yrs. of schooling	0.117***	0.100***	0.076***	-0.210***	-0.214***	-0.269***	-0.134***	-0.056***
	(0.008)	(0.008)	(0.011)	(0.010)	(0.010)	(0.009)	(0.010)	(0.008)
12 yrs. of schooling	0.170***	0.171***	0.157***	-0.307***	-0.305***	-0.296***	-0.174***	-0.082***
	(0.008)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
13 yrs. of schooling	0.181***	0.213***	0.216***	-0.320***	-0.316***	-0.305***	-0.183***	-0.060***
	(0.008)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
14 yrs. of schooling	0.179***	0.254***	0.289***	-0.339***	-0.332***	-0.306***	-0.186***	-0.065***
	(0.008)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.007)
16 yrs. of schooling	0.202***	0.370***	0.456***	-0.417***	-0.404***	-0.318***	-0.221***	-0.083***
	(0.008)	(0.008)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
17 yrs. of schooling	0.204***	0.385***	0.498***	-0.427***	-0.411***	-0.320***	-0.227***	-0.058***
	(0.008)	(0.009)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Observations	964,000	928,000	927,000	974,000	974,000	974,000	974,000	964,000
Adjusted R-squared	0.03	0.213	0.29	0.063	0.062	0.036	0.021	0.041
Y-mean	0.88	10.71	11.66	0.304	0.260	0.050	0.133	0.231

Sample is composed of US-born individuals aged 0-15 in 1940, with at least one parent present, who are assigned a unique PIK and observed in the 2000 Census or in the 2010 Census or in NUMIDENT with death date later than 2000.
Model 6 from Table 1 is used in all regressions. Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001
See the Data section for the definitions of outcome variables.

Table 12: Effect of education on survival to 2010, conditional on survival to 2000, controlling for mediators observed in the 2000 Census

	Dependent variable: Survival to 2010		
Log total income	0.019***		
	(0.001)		
Log household income		0.015***	
		(0.001)	
Log social security income			0.005***
			(0.001)
Homeowner	0.059***	0.059***	0.063***
	(0.002)	(0.002)	(0.002)
Log of median hh. inc. at block group level	0.013***	0.013***	0.018***
	(0.002)	(0.002)	(0.002)
Log of median house values at block group level	0.014***	0.014***	0.016***
	(0.001)	(0.001)	(0.001)
Lives alone	-0.041***	-0.024***	-0.034***
	(0.001)	(0.001)	(0.001)
Any difficulty	0.027***	0.029***	0.028***
	(0.003)	(0.003)	(0.003)
Physical difficulty	-0.163***	-0.166***	-0.167***
	(0.003)	(0.003)	(0.003)
Difficulty in Ability	-0.125***	-0.127***	-0.128***
to Perform Mental Tasks	(0.003)	(0.003)	(0.003)
Difficulty in Ability	-0.059***	-0.059***	-0.059***
to Work At a Job or Business	(0.002)	(0.002)	(0.002)
4 yrs. of schooling	-0.000	-0.000	-0.001
	(0.014)	(0.014)	(0.014)
6 yrs. of schooling	0.006	0.007	0.006
	(0.011)	(0.011)	(0.011)
8 yrs. of schooling	0.008	0.010	0.010
	(0.010)	(0.010)	(0.010)
9 yrs. of schooling	0.004	0.006	0.007
	(0.010)	(0.010)	(0.010)
10 yrs. of schooling	0.008	0.010	0.011
	(0.009)	(0.009)	(0.009)
11 yrs. of schooling	0.007	0.009	0.010
	(0.010)	(0.010)	(0.010)
12 yrs. of schooling	0.041***	0.044***	0.047***
	(0.009)	(0.009)	(0.009)
13 yrs. of schooling	0.059***	0.064***	0.068***
	(0.009)	(0.009)	(0.009)
14 yrs. of schooling	0.054***	0.059***	0.065***
	(0.009)	(0.009)	(0.009)
16 yrs. of schooling	0.069***	0.076***	0.084***
	(0.009)	(0.009)	(0.009)
17 yrs. of schooling	0.081***	0.091***	0.101***
	(0.009)	(0.009)	(0.009)
Observations	927,000	927,000	927,000
Adjusted R-squared	0.106	0.105	0.105
Y-mean	0.740	0.740	0.740

Sample is composed of US-born individuals aged 0-15 in 1940, with at least one parent present, who were assigned a unique PIK and observed in the 2000 Census or in the 2010 Census or in NUMIDENT with death date later than 2000.
Model 6 from Table 1 is used in all regressions. Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
See the Data section for details on how the mediator variables were defined.
** p<0.01; *** p<0.001
We assign to individuals with zero or negative total income or social security income the loginctot value of zero and the logssinc value of zero, respectively. Thus, loginctot and logssinc will only contain missing values if total income or social security income information is missing, respectively.

Figures

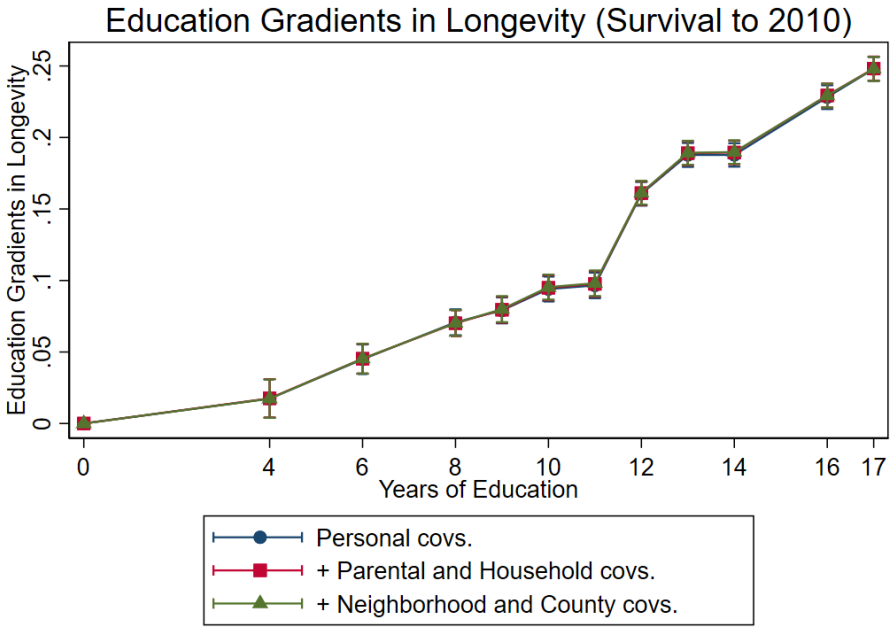


Figure 1: Education Gradients in Longevity (Survival to 2010) - estimates from Table 3

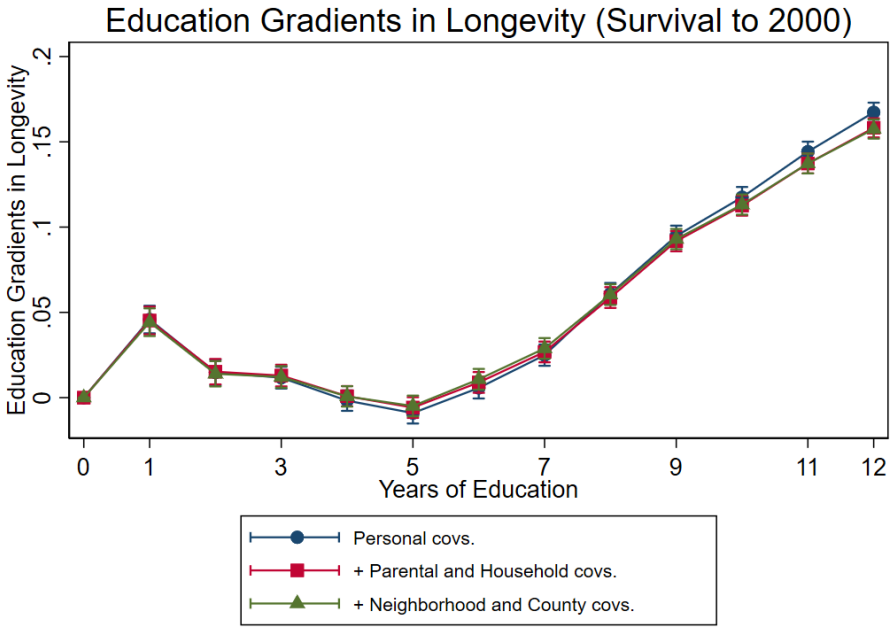


Figure 2: Education Gradients in Longevity (Survival to 2000) - estimates from Table 4

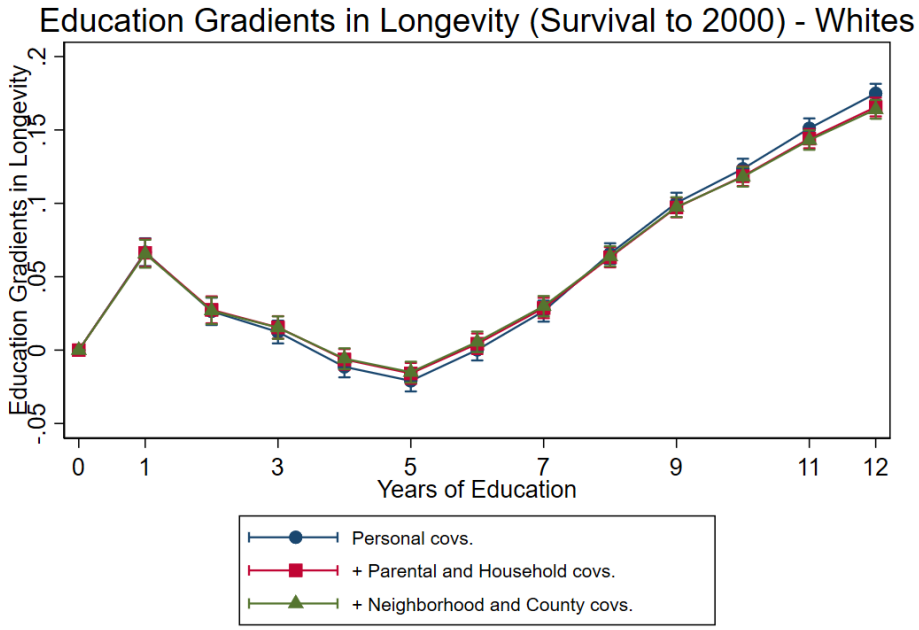


Figure 3: Education Gradients in Longevity (Survival to 2000), Whites - estimates from Table 5

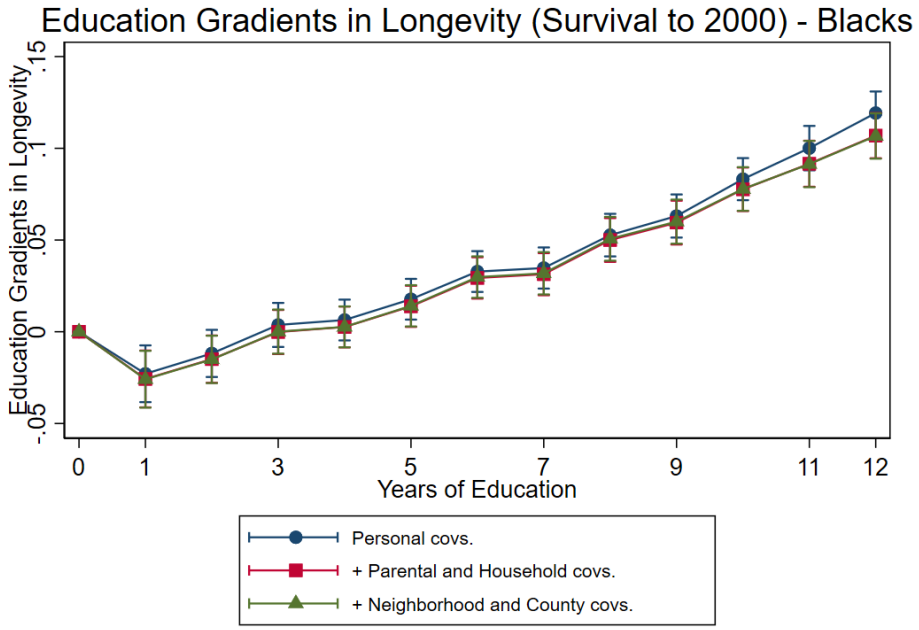


Figure 4: Education Gradients in Longevity (Survival to 2000), Blacks - estimates from Table 6

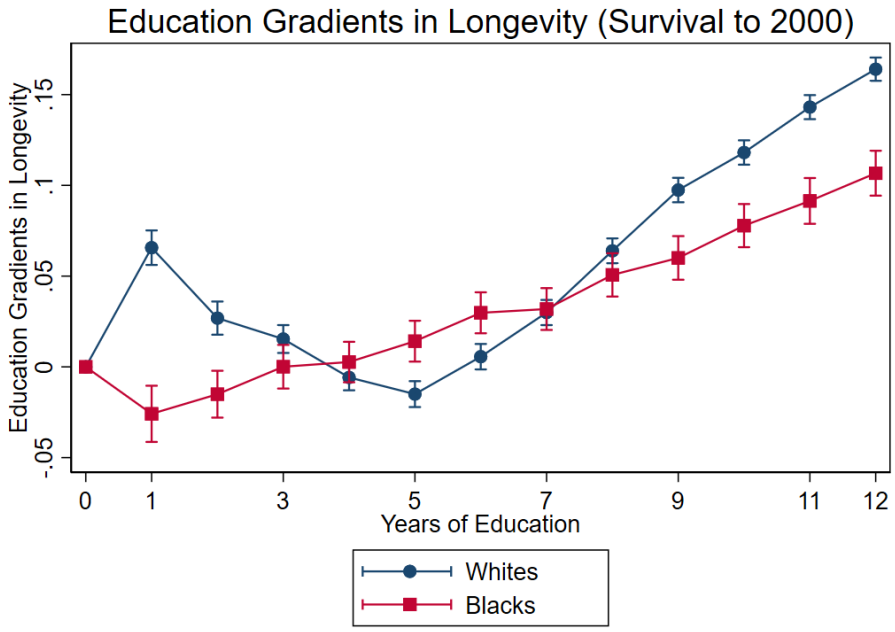


Figure 5: Education Gradients in Longevity (Survival to 2000) - estimates from Tables 5 and 6

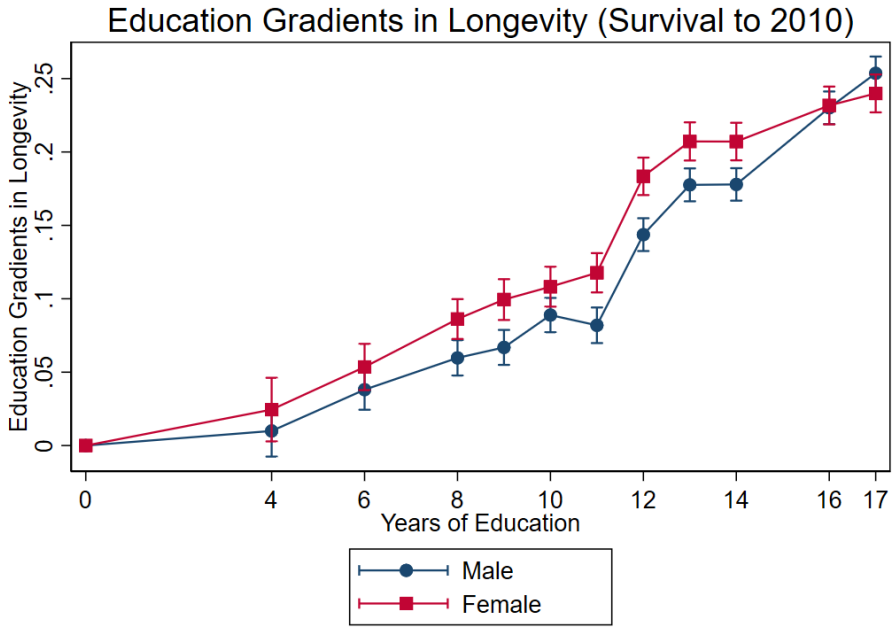


Figure 6: Education Gradients in Longevity (Survival to 2010) - estimates from Table **??**

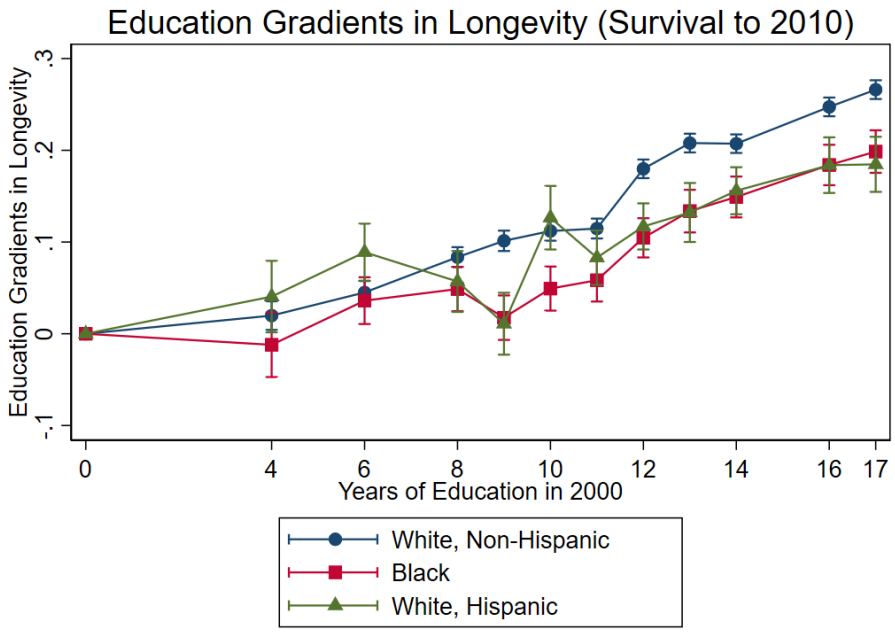


Figure 7: Education Gradients in Longevity (Survival to 2010) - estimates from Table **??**

Appendix A. Additional tables and figures

Table A1: County-Level Variables, with Sources

County-level variable	Source
Average occupational score, men aged 35–54 Average occupational score, women aged 35–54 Average number of children, women aged 21–30 Indicator, city with population 25,000+ in county Doctor/population ratio for Whites	1940 IPUMS, 100 percent sample
Minimal distance to a normal school	Author's calculations, using locations listed in The Factors Operating in the Location of State Normal Schools by Harry C. Humphreys (Columbia University, 1923)
Minimal distance to a 4-year college	Author's calculations, using data for institutions founded before 1940, College Scorecard, US Department of Education
Fraction of White births delivered in hospital Infant mortality for Whites, 1941	Data provided by Michael Haines
Population density (persons per square mile), 1940 % of rural farm dwelling units with electric lighting, 1940 % of rural farm dwelling units with running water, 1940 % of workers employed in agriculture, 1940 % of dwelling units with private bathtub or shower, 1940 Sum of major war supply contracts and projects ($)	County and City Data Book[a]

[a]US Bureau of the Census, County and City Data Book, Consolidated File: County Data, 1947–1977. Distributed by Inter-University Consortium for Political and Social Research, 2012-09-18. https://doi.org/10.3886/ICPSR07736.v2

Table A2: Effect of *potential* years of schooling on survival to 2000, for individuals aged 15-18 in 1940, all races

	Dependent variable: Merged2000ordiedafter					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
On track to 3rd grade	-0.000	0.002	0.001	0.0009	0.003	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
On track to 4th grade	-0.035***	-0.031***	-0.030***	-0.035***	-0.030***	-0.030***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
On track to 5th grade	-0.007	-0.006	-0.005	-0.007	-0.006	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
On track to 6th grade	0.009*	0.009*	0.010**	0.009*	0.009*	0.010**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
On track to 7th grade	0.015***	0.013***	0.013***	0.014***	0.012***	0.013***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
On track to 8th grade	0.034***	0.029***	0.029***	0.033***	0.029***	0.029***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
On track to 9th grade	0.032***	0.027***	0.026***	0.032***	0.027***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
On track to 10th grade	0.016***	0.013***	0.012***	0.016***	0.012***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
On track to 11th grade	0.025***	0.022***	0.021***	0.025***	0.022***	0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
On track to 12+ grades	0.008***	0.006***	0.006***	0.008***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2,386,000	2,386,000	2,386,000	2,386,000	2,386,000	2,386,000
Adjusted R-squared	0.05	0.055	0.055	0.052	0.055	0.056
Y-mean	0.549	0.549	0.549	0.549	0.549	0.549

Sample is composed of US-born individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table A3: Effect of *potential* years of schooling on survival to 2000, for individuals aged 15-18 in 1940, Whites

	Dependent variable: Merged2000ordiedafter					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
On track to 3rd grade	0.011	0.014*	0.015*	0.013	0.016*	0.016*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
On track to 4th grade	-0.065***	-0.057***	-0.055***	-0.064***	-0.056***	-0.055***
	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
On track to 5th grade	-0.020**	-0.018**	-0.018**	-0.019**	-0.018**	-0.018**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
On track to 6th grade	0.021***	0.020***	0.020***	0.020***	0.019***	0.019***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
On track to 7th grade	0.024***	0.020***	0.020***	0.023***	0.019***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
On track to 8th grade	0.038***	0.033***	0.032***	0.037***	0.032***	0.031***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
On track to 9th grade	0.035***	0.029***	0.028***	0.034***	0.029***	0.028***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
On track to 10th grade	0.017***	0.013***	0.013***	0.016***	0.013***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
On track to 11th grade	0.026***	0.023***	0.022***	0.026***	0.023***	0.023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
On track to 12+ grades	0.008***	0.006***	0.006***	0.008***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2,238,000	2,238,000	2,238,000	2,238,000	2,238,000	2,238,000
Adjusted R-squared	0.044	0.048	0.049	0.046	0.049	0.05
Y-mean	0.559	0.559	0.559	0.559	0.559	0.559

Sample is composed of US-born White individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001

Table A4: Effect of *potential* years of schooling on survival to 2000, for individuals aged 15-18 in 1940, Blacks

	Dependent variable: Merged2000ordiedafter					
	without county FEs			**with county x gender FEs**		
	Model:			**Model:**		
	person covs	**pers, parent HH covs**	**pers, parent, HH, neigh and county covs**	**person covs**	**pers, parent, HH covs**	**pers, parent, HH, and neigh covs**
	1	**2**	**3**	**4**	**5**	**6**
On track to 3rd grade	-0.008	-0.012	-0.012	-0.008	-0.011	-0.011
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
On track to 4th grade	0.010	0.010	0.010	0.009	0.009	0.009
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
On track to 5th grade	0.016*	0.016*	0.017*	0.017*	0.017*	0.017*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
On track to 6th grade	-0.001	0.000	0.000	0.000	0.000	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
On track to 7th grade	0.005	0.005	0.005	0.006	0.006	0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
On track to 8th grade	0.015**	0.018***	0.018***	0.017**	0.017**	0.017**
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
On track to 9th grade	0.003	0.002	0.002	0.004	0.002	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
On track to 10th grade	0.011	0.012*	0.012*	0.014*	0.012*	0.012*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
On track to 11th grade	0.007	0.006	0.005	0.008	0.006	0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
On track to 12+ grades	0.012	0.010	0.010	0.012	0.009	0.009
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	140,000	140,000	140,000	140,000	140,000	140,000
Adjusted R-squared	0.035	0.039	0.039	0.036	0.039	0.039
Y-mean	0.381	0.381	0.381	0.381	0.381	0.381

Sample is composed of US-born Black individuals aged 15-18, living with at least one parent in 1940, who were assigned a unique PIK.
See Table 1 for details on the sets of covariates across specifications.
Standard errors were clustered at the 1940 county level.
Number of observations was rounded in accordance with the Census Bureau rules for disclosure.
** p<0.01; *** p<0.001