The Effects of Education on Mortality: Evidence Using College Expansions*

Jason Fletcher†

Hamid Noghanibehambari‡

Abstract

This paper explores the long-run health benefits of education for longevity. Using mortality data from the Social Security Administration (1988-2005) linked to geographic locations in the 1940-census data, we exploit changes in college availability across cohorts in local areas. Our treatment on the treated calculations suggest increases in longevity between 1-1.6 years. Some further analyses suggest the results are not driven by pre-tends, endogenous migration, and other time-varying local confounders. This paper adds to the literature on the health and social benefits of education.

Keywords: Education, Health, Mortality, College Education

JEL Codes: I23, I26, H51, H75, D62

---

* The authors claim no conflict of interest. We thank Enrico Moretti for generously providing the data on college openings. We also thank Joshua Goldstein, Adriana Lleras-Muney, participants of Health-Aging-Place working group at the University of Wisconsin-Madison, CenSoc Working Group at the University of California-Berkeley, seminar participants of the Midwest Economic Association conference, participants of the American Society of Health Economists Conference, and seminar participants of the Data-Intensive Research Conference for their comments and suggestions. The authors would like to acknowledge financial support from NIA grant R01AG060109 and the Center for Demography of Health and Aging (CDHA) at the University of Wisconsin-Madison under NIA core grant P30 AG17266.

† La Follette School of Public Affairs, University of Wisconsin-Madison, 1225 Observatory Drive, Madison, WI 53706-1211, USA

Email: jason.fletcher@wisc.edu

‡ **Corresponding Author**, Center for Demography of Health and Aging, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706

Email: noghanibeham@wisc.edu, Phone: +1-806-620-1812

1. Introduction

It is well documented that the returns to education go beyond labor market outcomes. Education not only has spillover effects for peers, colleagues, and other family members (Fruehwirth, 2014; Jaffe et al., 2006; Martins and Jin, 2008) but also shapes long-run outcomes such as health at older ages (Albouy and Lequien, 2009; Mazzonna, 2014)[4]. A recent review by Galama et al. (2018) classified the literature that establishes a causal path between education and health, health behavior, and mortality into three general categories: 1) studies that employ randomized controlled trials such as the Perry Preschool Program and Abecedarian which primarily provides early childhood education intervention (Campbell et al., 2014; Conti et al., 2016; Heckman et al., 2013); 2) within-twin variation in years of schooling as the primary education shock (Behrman et al., 2011; Lundborg et al., 2016; Madsen et al., 2010); 3) quasi-experimental settings with primary focus on school reforms which increase minimum school leaving age and compulsory schooling (Black et al., 2015; Clark & Royer, 2013; Fischer et al., 2013; Fletcher, 2015; Gathmann et al., 2015; Lager & Torssander, 2012; Lleras-Muney, 2005; Mazumder, 2008; Meghir et al., 2018; van Kippersluis et al., 2011a).

With the exception of twin/sibling studies, the literature concentrates on interventions and improvements at early ages or during K-12 education focusing on years of schooling as the primary explanatory variable and has offered mixed evidence. Some studies point to large positive effects for life expectancy and gains in mortality while other studies do not find any suggestive evidence and in some cases even opposite-signed coefficients (Albouy & Lequien, 2009; Behrman et al., 2011; Buckles et al., 2016; Campbell et al., 2014; Conti et al., 2016; Gathmann et al., 2015;

[4] Many earlier studies within this literature examined cross-sectional associations of education and mortality without fully addressing the common endowments' influence in driving both education and health (Kravdal, 2008; Meara et al., 2017; Ross et al., 2012; Zajacova, 2006).

Lundborg et al., 2016; Madsen et al., 2010; Meghir et al., 2018). This paper aims to extend this literature and provide new insight into the long-term effects of education by quantifying the benefits of higher education on old-age mortality in the US.

We investigate whether changes in the number of available colleges in the local area when individuals are 17 years old could influence their age at death. We find robust evidence that the availability of 4-year colleges at the age of high school completion can influence their age at death. The reduced form effects are economically meaningful. For example, our most parametrized specification suggests that an additional college opening in own or neighboring counties raises the age of death by 0.13 months. This is equivalent to roughly 5 percent of white-nonwhite differences in old age mortality. Because college openings have modest direct effects on educational attainments, the treatment on the treated calculations suggest large effects, between 1-1.6 years of added life for older age individuals.

We implement a wide range of robustness exercises to control for family unobserved heterogeneity, assess for alternative standard error adjustments, consider alternative specifications, functional forms, and various measures of college access. In addition, we explore the possibility of endogeneity of college access, as people with specific demographic characteristics may migrate to areas that have experienced college expansions. In addition, we consider placebo tests that assign college expansions to the year individuals turn age 25, 30, and 35, well after the age that college availability could have an effect on individual's education attainment.

Our paper contributes to the literature on the social and health benefits of education in two ways. First, this is the first study to link the construction of new colleges on education and later-life mortality. Second, while similar papers have used longitudinal data with limited observations, our new longitudinal dataset provides millions of observations which significantly adds power to

our statistical tests. In addition, the increased sample size enables a wide range of heterogeneity analysis by cohort, place, and demographic.

The rest of the paper is organized as follows. Section 2 gives a review of related literature. Section 3 introduces the data sources and sample construction. In section 4, we discuss the econometric method and potential endogeneity concerns. Section 5 offers the main results and discusses additional analysis. Concluding remarks are provided section 6.

2. Literature Review

Education has spillover effects across a wide range of outcomes which can operate as a set of channels for health at older age, including longevity (Oreopoulos & Salvanes, 2011). While there is a relatively large literature on education and mortality, most studies focus on years of education in general and exploit compulsory schooling laws or similar reforms as the shock to education and find mixed evidence (Albouy and Lequien, 2009; Buckles et al., 2016; Conti et al., 2010; Everett et al., 2013; Jemal et al., 2008; Kalediene and Petrauskiene, 2005; Kippersluis et al., 2011; Kravdal, 2008; Lager and Torssander, 2012; Lynch, 2003; Meara et al., 2017; Ross et al., 2012; Zajacova, 2006). In her seminal study, Lleras-Muney (2005) takes advantage of changes in compulsory schooling (child labor and compulsory attendance laws) as the instrument for education and builds a synthetic cohort from decennial censuses to measure 10-year mortality rates. She finds that an additional year of education leads to a 6.3 percentage point reduction in 10-year death rates. Mazumder (2008b) shows that these findings are not robust to adding state trends. Black et al. (2015b) use Vital Statistics death records to investigate the effects of compulsory schooling laws on mortality and find that the gains in mortality can be explained by cohort and state fixed effects. Fletcher (2015) employs survey data (AARP Diet and Health Study) and instruments education with compulsory schooling laws and finds positive health effects and

large gains in mortality. His calculations provide similar effects to the findings of Lleras-Muney (2005) but they are statistically insignificant.

Lleras-Muney et al. (2020) examines the association between education and old age longevity for cohorts born between 1906-1915 in the US. They link social security death records with 1940 full-count census and find that an additional year of education is associated with 0.4 higher age at death. In a similar study, Halpern-Manners et al. (2020) apply linking techniques to merge 1920 and 1940 full-count census covering males born between 1910-1920 with social security administrative mortality data and implement a twin-fixed-effect strategy to study the effect of education on mortality. They find that an additional year of schooling raises age at death by 0.3 years. These effects are slightly smaller than OLS estimates which suggests that common endowments such as genetic factors only modestly influence the education-mortality association in their sample. Savelyev et al. (2022) employ data from the Minnesota Twin Registry and explore the effect of education on longevity and other health outcomes. They find that an additional year of education reduces the probability of death within 20 years of the survey by about 2.3 percentage points.

Clark and Royer (2013) apply a regression discontinuity design strategy to take advantage of schooling reforms in the UK and find that the reforms were successful in raising educational attainments and wages while they had no discernible effect on health and mortality. In a similar study, Meghir et al. (2018) use the phase-in of a series of compulsory schooling reforms across different municipalities in Sweden as an educational promotion policy and find that although the reforms improved educational outcomes they did not have any statistical effect on mortality, health care utilization, and hospitalization. Using two French compulsory schooling reforms that increased school leaving age to 14 and 16, Albouy and Lequien (2009) find small and statistically

insignificant evidence that education lowers mortality rates of affected cohorts. In contrast to these findings, Gathmann et al. (2015) exploit the compulsory schooling reforms across 15 European countries and find small but statistically significant effects of education on mortality among men with no impact on female mortality rates. Kippersluis et al. (2011) exploit a compulsory schooling reform in Netherland which significantly increased schooling to assess its impact on longevity. They find that an additional year of education reduces the likelihood of death before age 89 by 6 percent relative to the mean, conditional on survival up to age 81.

A small literature has used quasi-experimental and sibling methods to examine the effects of college on older age health. Buckles et al. (2016) employ the deferment of the Vietnam War draft for college students as the source of identification to explore the effect of college education on cumulative mortality. They find that a 1 percentage point increase in college completion rate among men reduces their mortality rate by 0.95 fewer death per 1,000. Their results suggest that raises in income and health insurance are plausible mechanisms of impact. Savelyev (2020) employs the Terman Life-Cycle Study of Children with High Ability data that contains a sample of individuals with high-IQ and follows them throughout their life cycle. He finds that among high-IQ individuals having a college education significantly increases longevity.

Cowan & Tefft (2020) use Census and American Community Survey data to explore the effect of college accessibility (number of colleges per capita) on adult outcomes. They find positive effects on education, employment, and self-reported health. However, they fail to detect a significant improvement for old-age health outcomes and mental health outcomes[5]. Kamhöfer et al. (2019) explore education and health effects of new college openings in Germany. They also

[5] The authors use state-level variation in college access rather than our focus on county level variation in college openings.

find positive and significant local average effects on education, wages, and cognitive ability. However, they argue that these findings are partly driven by selection based on unobservables and that the effects on those with the lowest desire to attend college is zero.

Fletcher & Frisvold (2014) compare siblings in the Wisconsin Longitudinal Study to show college attendance is related to old age preventive care decisions and college quality is related to better weight outcomes in old age, respectively.[6] Bautista et al. (2020) use the discontinuity in college attendance as a result of the 1973 military coup in Chile, which significantly reduced access to college for those reaching college age, and show that the reduction in access and enrollment was associated with higher age-adjusted mortality later in life.

Attending college can affect later-life health and old-age mortality through several channels. The primary channel is labor market outcomes as the large and old literature of returns to education has documented (Acemoglu and Angrist, 2000; Angrist and Keueger, 1991; Card, 1999, 2001; Cellini and Chaudhary, 2014; Dickson and Harmon, 2011; Leigh and Gill, 1997; Long, 2010; Psacharopoulos, 1985). For instance, Cellini and Chaudhary (2014) show that individuals who attend a for-profit college compared to their counterfactual case have 10 percent higher earnings. Income-induced raises in welfare improves other mediatory outcomes such as a healthier living environment, better nutrition, better health insurance, and improved health behavior which in turn have strong effects on mortality and life expectancy (Backlund et al., 1999; Gonzalez and Quast, 2013; Jamison et al., 2007; Lefèbvre et al., 2018; Lindahl, 2005; Muller, 2002; Snyder and Evans, 2006). Lacroix et al. (2019b) investigate the health impacts of college education in Canada. They find that college-educated individuals live 4.1 additional years

[6] Fletcher & Frisvold (2011) compare siblings in Add Health to show that college selectivity is associated with lower likelihood of tobacco and marijuana use in young adulthood.

(conditional on being alive at age 51), have 27.3 percent lower lifetime hospital stays, they have lower rates of diabetes and stroke, and they have higher survival rates conditional on having a health condition.

College-educated individuals utilize healthier behaviors which leave them in trajectories that increase their longevity. For instance, Walque (2007) shows that college-educated people are less likely to smoke, and among those who do smoke, they are more likely to quit. An old and established literature documents the mortality consequences of smoking (Doll and Hill, 1956; Fenelon and Preston, 2012; Preston et al., 2010). In addition, college education facilitates individuals with more white-collar jobs that are comparably safer than jobs in manufacturing, mining, and agriculture. Several studies show that mortality rates vary across individuals in different occupations (Johnson et al., 1999; Luy et al., 2011). Higher educated individuals are less likely to have adverse mental health conditions, too. There is a small literature that documents the negative association between education and mental health issues such as excess anxiety and depression (Chevalier et al., 2004; Cornaglia et al., 2015). Poor mental health, in turn, leads to higher mortality and lower longevity (Cuijpers and Smit, 2002; Schulz et al., 2002; Wulsin et al., 1999). Another channel is social spillovers in education. The benefits of education roll down to health and specifically mortality outcomes of spouse (Jaffe et al., 2006; Spoerri et al., 2014), other family members (Kravdal, 2008), and colleagues in the workplace (Martins and Jin, 2008).

3. Data Source

The primary source of data comes from the Censoc project outlined in Goldstein et al. (2021).[7] It uses death records from social security administration for individuals who die in old

[7] The name is extracted from the beginnings of ***cen***sus and ***soc***ial security administration.

age and implement data-linkage techniques to link with full-count 1940 census records. We use the Censoc-Numident (hereafter Numident) dataset from this project which has linked 7.9 million individuals to the 1940 census who died between the years 1988-2005. We merge this data with the 1940 census extracted from Ruggles et al. (2020). There are three primary advantages of this linkage that help our identification strategy. First, the resulting sample size consists of millions of individuals, which is considerably larger than available longitudinal datasets and enables much more powerful statistical tests. Second, the county identifier in the 1940 census provides detailed granularity that can be used to match with the county-level college dataset, unlike state-of-birth indicators available in the Health and Retirement Study, Decennial Census, or American Community Survey data, among others. Third, we can observe other family member's socioeconomic characteristics such as parental education that we exploit in our robustness and heterogeneity analyses.

The data on county-level college counts are obtained from Currie and Moretti (2003b). It reports the total number (as well as some disaggregated categories) of 4-year and 2-year colleges at each county from 1940-onward. Figure 1 shows the geographic distribution of college inventory at 1940 (top panel) and changes in college counts (bottom panel). About 45 and 33 states had at least one county with a 4-year and 2-year college opening in our sample.[8] About 80 percent of counties experienced only one expansion. The counties with the higher number of openings include Los Angeles, CA with five new 4-year colleges and six new 2-year colleges, Nassau, NY with four new 4-year colleges, and Allegheny, PA with four new 2-year colleges. We merge this

[8] The states with no 4-year college expansions include: Delaware, DC, Montana, Utah, Wisconsin, and Wyoming. The states with no 2-year college expansions include: Alaska, Arizona, Arkansas, Delaware, DC, Idaho, Iowa, Montana, Nebraska, Nevada, New Jersey, Rode Island, South Dakota, Tennessee, Vermont, Virginia, West Virginia, and Wisconsin.

with our Numident dataset based on the county of residence and the year individuals turn age 17. This leaves us with 3,849,792 observations from cohorts who were born between the years 1923-1940 and died at ages 47-82 between the years 1988-2005.

Summary statistics of the final sample are reported in Table 1. To have a better picture of the demographic composition of this sample, we compare with that of the full-count 1940 census. The sample underrepresents females (0.43 vs 0.49), underrepresents first-generation immigrants (0.3 vs 8.9), and overrepresents second-generation immigrants (0.14 vs 0.09) while the share of whites, blacks, and Hispanics are quite similar to the 1940 census. The average age-at-death is 68.8 years or equivalently 831.4 months. Since the months of longevity provides a more accurate measure, we use age-at-death in months as the primary outcome. Figure 2 shows the geographic distribution of age-at-death by county of residence in 1940.

4. Econometric framework

The identification strategy compares the age of death of individuals who, at the age of 17, resided in counties that experienced a college expansion to those who resided in counties with no college openings, after a college opening compared to before the expansion. We operationalize this difference-in-difference model using the following regression:

$$DA_{icb} = \alpha_0 + \alpha_1 Coll_{c,b+17}^{4year} + \alpha_2 Coll_{c,b+17}^{2year} + \alpha_3 X_i + \zeta_c + \gamma_b + \varepsilon_{icb} \quad (1)$$

Where the outcome (DA) is the death age of individual i from birth cohort b who, at the age of 17, resided in county c. Since college openings could have spillover effects for college attendance decisions of not only the residents of the county but also the residents of neighboring counties, we also create a second measure that aggregates our measure of college expansion to the college counts of own and neighboring counties, conditional on being in the same state. Therefore,

the parameter $Coll$ represents the total number of 4-year and 2-year colleges at the own and within-state neighboring counties where the individual resided at age 17 (b+17).[9] In matrix X, we include some individual covariates including indicators for race/ethnicity, gender, first-generation immigrant, and second-generation immigrant. The county fixed effects and birth cohort fixed effects are represented by parameters ζ and γ, respectively. ε is a disturbance term. In some additional specifications, we also include a county-by-year linear trend. We cluster the standard errors at the county level.

4.1. Concerns over Endogeneity

The parallel trend assumption behind the identification strategy -that the outcomes of those residing in the vicinity of a college opening would have followed the same path and been influenced by the same factors as the outcomes of those not exposed to a new college opening- may be violated for three primary reasons. First, people may migrate to a county that had a college opening either for attending college or for other reasons, such as improved economic conditions, which resulted in college expansion in the first place. If individuals who chose to migrate have characteristics that are correlated with their health in the long run, the OLS results of equation 1 are biased as a result of this self-selection. For example, if whites migrate more than other races, because of their ability and affordability to move or their willingness to be more educated, there will be a sample selection problem and the results overestimate the true effects as whites have higher longevity and better health endowment for other reasons that cannot be fully captured by race dummies. To explore this potential source of endogeneity, we use a series of observable characteristics as the outcome of equation 1. The results are reported in Table 2 for a specification

[9] Appendix A shows that the results are quite robust and similar when we replace our measure of county by own county number of colleges or with own county and all neighboring counties regardless of the fact that they are within the same state or not.

with county and birth cohort fixed effects (panel A) as well as a specification that includes a county-specific trend in birth cohort (panel B). The results offer a mixed and inconsistent pattern of migration. Panel A suggests some association between 2-year college expansion and increases in the share of whites and Hispanics and a reduction in the share of immigrants. There is also an increase in individuals with higher parental education (columns 7-8of panel A) as a response to an increase in 4-year colleges. Since education is a leading factor in family socioeconomic status, the reduction in the family's Duncan Socioeconomic Index (column 9, panel A) reveals an inconsistency. However, all the effects become small in magnitude and statistically insignificant when we add county trends. For instance, the marginal effect of 4-year college on father's schooling drops by 94 percent and suggests an insignificant effect of 0.08 percent change as a result of a new 4-year college opening. Since inclusion of county-birth-cohort trend rule out the discussed endogeneity issues and help balancing the sample, our preferred identification strategy includes, in addition to county and cohort fixed effects, a linear county trend.

Second, changes in the number of colleges in an area could be correlated with changes in local and state-level regulations and legislations that impact, through education or other channels, the longevity of individuals. Moreover, college expansion could reflect other contemporaneous changes in the local economic and non-economic environment that, in turn, has long-term effects for residents. For example, city expansions could lead to additional college openings as well as improving health care access, better job opportunities, and new hospital construction, with plausibly positive long-term effects, or degrading environmental quality by raising pollution, with potentially negative cumulative effects for longevity. These omitted variables result in upward-biased and downward-biased OLS coefficients for the former case and latter case, respectively. The problem of searching for such confounding factors is the scarcity of county-level data over

the period of the study (1940-1957). However, there are limited vital statistics information at the county-level which we could exploit as a proxy for general heath trends and health care access. In so doing, we use county-level birth and death data from Bailey et al. (2016) and implement regressions that include county fixed effects, year fixed effects, and county-specific linear trend. We regress four outcomes of interest on our measure of 4-year and 2-year college. These outcomes include infant mortality rate, total mortality rate, fertility rate, and share of births attended in the hospital. The results are reported in Table 3. We do not observe any significant association between college openings and contemporaneous changes in health outcomes (proxied by death rates), endogenous changes in fertility, and access to health care (proxied by share of births delivered in hospital). The marginal effects are statistically insignificant and economically small.

Third, the results could be driven by pre-trend changes in health levels that are revealed in age-at-death and cannot be absorbed by birth cohort and county fixed effects. To examine this pre-trend problem, we implement an event study analysis in which the event time is the opening of a new college. This strategy compares the outcomes of individuals aged 17 in different years relative to a college opening in their own and neighboring counties, conditional on fixed effects and covariates.[10] The results, shown in two panels of Figure 4 for 4-year and 2-year colleges, are not consistent with important pre-trends. Compared to unexposed cohorts in counties with no expansion (event time=zero), cohorts who turn age 18-above at the time of college opening (both

[10] The event-study regressions are similar to the full specification of equation 1. Specifically, we use a regression of the following form: $DA_{icb} = \alpha + \sum_{k=-1}^{\underline{T}} \xi_k 1(t_{c,b+17} - t_{cb}^* = k) + \sum_{k=+1}^{\overline{T}} \eta_k 1(t_{c,b+17} - t_{cb}^* = k) + \beta X_i + \zeta_c \times T_b + \gamma_b + \varepsilon_{icb}$. In this formulation, the coefficients ξ represent the average difference in the outcome between cohorts who turned 17 k periods before a new college opening and cohorts who turned 17 at the time of a new opening (pre-treatment coefficients). Likewise, the coefficients η represent the average difference in the outcome between cohorts who turned 17 k periods after a new college opening and cohorts who turned 17 at the time of a new opening (post-treatment coefficients). We include county fixed effects interacted with a linear cohort trend ($\zeta \times T$), birth year fixed effects (λ), and individual covariates. We cluster standard errors at the county level.

panels) reveal no differences in their age-at-death. The event-time coefficients are small in magnitude and not statistically different from zero. For 4-year colleges, the coefficients of exposed cohorts start to rise in magnitude and become statistically significant for cohorts who turn age 16-below at the time of college expansion. The overall difference-in-difference estimate that compares post-treatment to pre-treatment cohorts is 0.7 (se=0.15). However, we observe no post-treatment difference in the case of exposure to 2-year colleges. Important to our empirical strategy is the fact that we do not detect any pre-trend in outcomes for unaffected cohorts for both sets of college expansions.

As an additional step we implement a series of placebo tests by assigning the number of colleges to individuals at ages later than 17, specifically, ages 25, 30, and 35. If healthier individuals, who would have otherwise higher longevity, move to counties with college expansions for reasons related to overall trends of the county (e.g. better job opportunities) when they age 25, then the college availability in the new environment should be strongly associated with their health outcomes later in life. The results (shown in Table 4) do not provide evidence for this issue. The effects are quite small (relative to the main results) and in most cases statistically insignificant. Interestingly, the signs are reversed in specifications that include a county-specific trend for assignment at ages 25 (column 3) and 30 (column 6).

Another concern is sorting individuals based on the probability of linkage between Numident death records and the 1940 census. This becomes problematic for our identification strategy if the linkage is correlated with higher/lower number of colleges in local area of residence. To explore this source of selection, we start with the original population of individuals born between the years 1923-1940 and observed in the full-count 1940 census (roughly 40M observations). We link this with the Numident death records in order to create a new dummy

variable that indicates successful merging between the two datasets. We then merge this with county-level college data and regress the successful merging dummy variable on measures of 4-year and 2-year college opening using the same empirical method as discussed in equation 1. The results are reported in Table 5. In the full specification of column 3 where we include fixed effects, trends, and individual controls, there is no statistically significant association between our measures of college opening and the probability of merging between census and Numident. Moreover, the average effects on the outcome is close to zero, implying roughly 0.4 and 0.5 percent change as a result of a 4-year and a 2-year college opening.

One may also argue that the window of observation in social security administration death record is narrow and does not include those who die earlier or later. Looking at the Vital Statistics cause-specific death record data from 1959-2017, roughly 39 percent of deaths to birth cohorts of 1923-1940 occur between 1988-2005 (Numident window). Comparing to deaths outside of this window (1959-1987, 2006-2017), Numident records are 1.3 percentage points less likely to be white, 4.6 percentage points less likely to be female, and 3.4 percentage points more likely to be black. We will show that, in our sample, the effects are stronger for 4-year colleges and that the effects are not statistically different among different races and ethnicity but more pronounced among males. Therefore, one possible concern in extending the results in this paper to the whole population is the overrepresentation of males for which the effects are stronger.

In another attempt to explore possible issues with the window selection of Numident, we replicate the main results using Censoc-DMF data which links death records of males who die between 1975 to 2005 with the 1940 census. This allows us to explore the effects of deaths that occurred up to 12 years before the start of Numident data. We apply the same sample selections and implement the same econometric method as in the main analysis in the text. These results are

reported in Appendix B. The effects of the death window of 1988-2005 (similar to Numident) are virtually the same as the main results of section 5.1. However, the effects are smaller in magnitude when we look at all years covered in DMF. This may suggest that the effects of education on mortality are better detected at older ages.

5. Results

5.1. Main Results

Figure 3 shows the density distribution of age-at-death for counties with no colleges (never treated, in red) and counties with at least one college opening (treated, in green). For observations below the median of age at death, we observe a higher density for those in no college opening counties. Likewise, for those in counties with at least an opening, the longevity density is higher in higher values of the outcome. Specifically, cohorts exposed to college expansions have 3.6 months higher age-at-death (the raw difference in mortality in counties with at least one opening compared with those in counties with no college opening).

To account for confounding factors, we turn our focus from this visual difference to the difference-in-difference strategy of equation 1. The main results are reported in Table 6 for specifications that include county and birth year fixed effects (column 1), individual covariates (column 2), and county-specific linear trend in birth cohort (column 3). The marginal effects of 4-year colleges are positive and very similar in columns 1 and 2 and drops by 35 percent when we add county-trends. It implies that an additional increase in the number of 4-year colleges is associated with roughly 0.13 months higher longevity. While this result is a small effect, we can put it into perspective by comparing it with the marginal effects of other covariates, specifically, females, blacks, and other races. This effect is equivalent to 5 percent of the difference in age-at-death between people of other races and whites, 3.4 percent difference in black-white gap in age-

at-death, and 2 percent of the difference in age-at-death of females versus males. Over the sample period, there are 729 counties with no 4-year colleges. Adding one 4-year college to these counties (an increase of 3.3 colleges) would increase the age-at-death of the population by 0.4 months, which closes the life expectancy gap between the US and other OECD countries by 2.4 percent.[11] The effect of 2-year college expansion is mixed, negative for columns 1-2 and positive for column 3. However, they are small compared to 4-year college marginal effects and also statistically insignificant. To have a better comparison of the effects of 4-year versus 2-year colleges, we report the elasticities at the end of each column. The elasticity comparison reveals that the effect of the same changes in 4-year college expansion is much larger than 2-year college expansion, by a factor of 26 (column 2) to 9 (column 3).

The results reported here are intention-to-treat effects as the college expansion leads to college education of only a fraction of the population. Although we explore the direct link between college expansion and college education in section 5.3 and try to convert the marginal effects into treatment-on-treated effects in section 5.4, we should note that these effects provide a minimum benefit of college openings on long-run mortality outcomes.

Moreover, not all the positive effects of college opening operate through increases in own education.[12] For instance, it could improve the education of the spouse which results in gains in mortality for both husband and wife (Jaffe et al., 2006; Spoerri et al., 2014). Therefore, college openings could affect mortality by increases in the education of other current and future family

[11] Based on OECD health status reports, life expectancy of USA in 2008 was 77.9 and the average OECD countries was 79.3.

[12] This is the primary reason that we prefer a reduced-form analysis and avoid applying 2SLS-IV tests as our main analysis method here. Since college opening could potentially operate through non-own-education channel to improve mortality outcomes, the exclusion restriction assumption will be violated.

members rather than one's own education. Another aspect is the improvements in education of coworkers which in turn has possible health spillover for own mortality outcomes later in life (Martins and Jin, 2008).

5.2. Robustness Checks

In Table 7, we execute a wide range of robustness checks. One important concern is the omission of family background and, specifically, parental characteristics. However, the results are quite robust when we add mother-father schooling and family's socioeconomic index (column 2), residential house value and father's wage (column 3), and mother fixed effects (column 4). The sibling strategy reported in column 4 -which compares the age-at-death of different siblings to the same mother who aged 17 in different years relative to a college expansion- suggests that after accounting for unobserved time-invariant family characteristics the effects could be even larger in magnitude than the main results (0.19 versus 0.13).

While in the main analysis we cluster standard errors at the county of residence level, we show the results when we use Huber-White robust standard errors (column 5) and alternative levels of clustering including birth year, state of residence, and county by birth year level (columns 6-8). In columns 9-13, we add various additional controls and fixed effects. Adding a polynomial function of county population (county of residence when the individual turns age 17) only slightly change the marginal effects (column 9). The coefficients are similar when we add birth-month-by-birth-year fixed effects to account for seasonality in birth outcomes that may appear in later-life health as well as death-month fixed effects to account for seasonality in mortality (columns 10 and 11).

While we extensively discuss and address the potential migration issues in section 4.1, we add two additional tests to further explore this concern. First, we exclude from the sample

observations whose state of birth is different than state of residence in 1940 under the assumption that between state migration is motivated, among other factors, availability and opening of colleges. However, the marginal effect of column 12 is virtually identical to the main effects and are consistent with our assumption regarding selective between-state migrations. Second, since individuals may migrate for college education when they are 17-18 years old, including these cohorts could introduce bias in our estimations. To explore this, we restrict the sample to observations whose age in 1940 is less than 15. The marginal effects for this subsample is reported in column 13 and is slightly larger than those in the main results. This comparison implies that probable migrations of older cohorts may bias the effects downward.

As mentioned in Currie and Moretti (2003a), while the Integrated Postsecondary Education Data System (IPEDS) is one primary source of college data, some colleges are not reported in this database.[13] Column 14 shows that eliminating those colleges that are not in the IPEDS listing leads to quite similar effects (0.125 versus 0.132). Moreover, not all colleges in our sample grant a degree. Column 15 suggests that focusing only on colleges that grant at least an associate degree produces similar coefficients. Decomposing the colleges by public-private status reveals somewhat heterogeneous effects. The 4-year college coefficient increases by 20 percent when we look at public colleges and remain similar to the main results for private colleges (column 16 and 17).

Additional tests are related to functional form checks. In column 18, we report a semi-log specification where the outcome is log of age at death. To capture the nonlinearities in the effects, we replace the outcome with binary variables that equal one if age at death is greater than 60 and

[13] Roughly 1.2 percent of 4-year colleges and 23 percent of 2-year colleges in our data are not reported and their information is not collected by IPEDS listing.

70 years. These results are reported in columns 19 and 20. We observe a positive and significant effect of 4-year college opening on the probability of aging beyond 60 and 70. . Finally, in column 21, we change the unit of the outcome from age in months to age in years as there could potentially be some seasonality factors associated with death not fully captured by death month fixed effects (column 11,Table 7) that have differential effects on mortality by education. The magnitude of 4-year college points to virtually the same effect as the main results.[14]

5.3. First Stage Effects

The results so far suggest a consistent and robust reduced-form effect of college openings on later-life mortality. The next step is to explore whether or not college openings increased educational outcomes and if so, to quantify the first stage effects of college openings on education. The Numident data does not report the education or other labor market outcomes of the deceased. One possible approach to estimate the first stage of college openings on educational attainment is to turn to 1960 5% Decennial Census, as it has detailed information on completed education as well as income, most of the 1923-1940 cohorts have completed their education by 1960. The main disadvantage of this data is that the county identifier is not available in the public use version and instead, it reports Public Use Microdata Area (PUMA). In addition, IPUMS de-identifies counties based on other geographic variables (including PUMA) and details about location characteristics. In the 1960 census, there are 1,344 PUMAs, and IPUMS de-identifies 435 counties. A PUMA is primarily defined based on the population of an area. In low populated areas (usually rural areas), a PUMA contains several counties while in high populated areas (usually urban areas) a county contains several PUMAs. Therefore, we aggregate college counts at areas where PUMAs cover several counties and use county-level college counts for areas where several PUMAs are included

[14] In Appendix E, we implement additional tests on subsamples based on region of residence, birth cohort, and county characteristics.

in a large county, conditioning on the fact that the county is de-identified by IPUMS.[15] Therefore, the variation in college expansions across areas is based on a combination of PUMA and county. Using this method, we can merge college data with the 1960 sample and have a match rate of about 97 percent. We drop respondents below age 22 as the primary outcome of interest is a college education. To mitigate migration issues, we implement two additional sample selection criteria. First, we use the information of state of residence 5 years ago (available in the 1960 census) to exclude those individuals whose current state of residence is different from their residence 5 years earlier. Second, we restrict the sample to respondents below age 30 in the year 1960. This leaves us with 341,834 observations. We implement regressions introduced in equation 1 while adding a PUMA-county fixed effect and trend. The results are reported in Table 8 for different outcomes in columns. An additional 4-year college raises the probability of having any college education by over 1 percentage point (column 1), equivalent to an increase of roughly 5.3 percent from the mean of the outcome. The effect of 2-year college is smaller in magnitude and imprecisely estimated. To compare these two coefficients, we focus on the elasticities (reported at the end of columns). A 10 percent increase in the number of 4-year and 2-year colleges raises any college education by 2.3 and 0.04 percent, respectively. The larger effect of 4-year versus 2-year colleges can also be observed at the elasticities for the years of schooling as the outcome (column 5), where the marginal effect of 4-year college is as large as three times the insignificant marginal effect of 2-year college. This pattern is also observed on log of wage and salary income (column 7) and log of total income (column 8). The effects of 4-year college expansion is statistically significant and relatively larger than the insignificant effects of 2-year college opening. An additional 4-year college is associated with approximately 3.8 percent rise in total income (significant at 5 percent

[15] For instance, Los Angeles County consists of 45 PUMAs in 1960 census.

level) while the 2-year college effects are quite small and statistically insignificant. One possible reason for small and imprecise estimations of 2-year college is that we are observing these individuals at younger ages than the age range over which the labor market returns of college education would appear. The overall picture reveals the positive effects of college opening on college education and possibly labor market outcomes with larger effects for 4-year college expansions.[16]

These estimated first-stage effects are partly in line with the results of Cowan & Tefft (2020) and Currie & Moretti (2003). Cowan & Tefft (2020) employ state-level measures of college access over the years 1980-2015 and find that college access (per capita) reduces high school graduation but increases college attendance. However, their first-stage results are primarily driven by 2-year public colleges. The first stage results of Currie & Moretti (2003) suggest that both 2-year and 4-year college expansions increase years of schooling with largest effects for 4-year college opening. Their data source is Natality birth records over the years 1970-1999. There are two reasons that the first stage results of both papers are somewhat different than ours. First, both papers employ different years of data and for different cohorts. Second, while the focus of both papers are in concentration and accessibility of colleges (measured by per capita values), our paper focuses on college openings.

While the results of Table 8 point to improvements in education and income of exposed cohorts as likely pathways, another potential mechanism is changes in the composition of occupations in the county following college openings. For instance, if a 4-year college opening is

[16] While the effects found here are consistent with our main findings, we are aware that the results could be distorted by migration issues as there are evidence of the effects of college education on migration (Malamud & Wozniak, 2012).

accompanied by the establishment of new firms with better-paid jobs and occupations, then part of the effects on longevity could operate through occupational composition changes. In Appendix G, we explore this channel using decennial census data over the years 1940-1960. Our results fail to provide robust and consistent evidence that there are discernible changes in the composition of occupations following college expansions.

5.4. A Discussion on the Magnitude of the Main Results

The results so far suggest that college expansions have long-run longevity gains and that these gains are primarily driven by 4-year college expansions. These point estimates and larger effects of 4-year college expansions in Table 6 are also in line with the first stage results of Table 8. They suggest that college expansions (and specifically 4-year college openings) increase college education and possibly total income, and through these channels, they positively affect the longevity of individuals. If this story is true, we can combine the first stage results (column 1, Table 8) with the main results (column 3, Table 6) and convert the Intent-to-Treat effects into Treatment-on-Treated effects. A back-of-the-envelope calculation suggests that an additional 4-year college raises the age-at-death of those who would have otherwise not attended college by about 12 months.[17] This is smaller than the effects of college education on life expectancy reported by Lacroix et al. (2019b) who found that college education raises longevity at age 51 by about 49 months. This effect is comparable to the findings of Halpern-Manners et al. (2020) who implement

[17] This is based on the outcome in column 1 of Table 8: attending at least one year of college. This does not include only those who have a college degree but all those who have ever attended any college. This is a better measure for our interpretation as college opening could operate through other channels except education-income to improve old-age health such as peer effects and access to better health-related information and improved critical thinking skills (Cutler & Lleras-Muney, 2006).

a twin-strategy and show that each additional year of schooling is associated with roughly 4 months higher age at death.[18]

To add to the TOT effect calculation, we also implement a two-sample two-stage least-square regression by using the information of completed education from the 1960 census and the information on mortality from 1940-census-Numident sample. We aggregate the college counts at the PUMA-county level following the same method as in the first stage results (section 5.3). We apply a two-sample SLS method that includes birth cohort fixed effects, PUMA-county fixed effects, PUMA-county trend, and individual controls. The results are reported in Appendix C. In the fully specified model, having at least a college degree is associated with roughly 20 months higher age-at-death. We also note that measurement error in the first stage analysis as well as the two-sample 2SLS framework (which uses PUMA-county rather than county linkages) would likely attenuate our coefficients and therefore would suggest smaller treatment on the treated effects.

5.5. Heterogeneity of the Results: Heterogeneity by Gender-Race/Ethnicity.

The similar patterns in the main results (Table 6) and the findings of the first stage effects (Table 8) suggest that improvements in education and income are likely channels of impact. To better connect the first stage effects and the reduced-form results, we explore how the effects vary across observable characteristics. If our assumption regarding the mechanism channel is correct, then we would observe larger mortality reductions in subpopulations that experience larger improvements in education-income outcomes. Therefore, we explore (and compare) the heterogeneity by gender, race, and ethnicity for both mortality and education-income outcomes.

[18] We should note that our results measure Local Average Treatment Effects (LATE) and establish the college-opening-induced marginal effects which by construction is different than what Halpern-Manners et al. (2020) estimate.

In Table 9, we explore how the effects vary over demographic characteristics by interacting with college counts a dummy for female, other races, black, and Hispanic (columns 1-4, in order). The mortality effect of 4-year and 2-year colleges is 0.03 and 0.06 months larger for males than females. The marginal effects of 4-year and 2-year colleges are larger among people of other races/ethnicities (Native Americans, Chinese, Japanese, and 'other' Asian-Pacific Islanders versus whites and blacks) by 0.08 and 0.03 months, respectively, although both coefficients are imprecisely estimated. Among blacks, the mortality gains from 2-year college opening are significantly larger than other races, by 0.12 additional months. On the other hand, the mortality gains from 2-year colleges are more pronounced among non-Hispanics than Hispanic people. We observe similar heterogeneity patterns in the first stage results when we interact the gender/race dummies with measures of college expansion and replace the outcome with various measures of college education using the 1960 census data. These results are reported in Appendix D. The effects of 4-year college expansion are more pronounced among males for all measures of college education. The marginal effects of both 4-year and 2-year college openings are slightly (and insignificantly) larger among other races in comparison with blacks and whites. This consistent pattern between the first stage effects and reduced-form effects also holds when we look at differential effects among blacks versus non-blacks. The 4-year college effects are larger among non-blacks while the effects of 2-year colleges are larger among blacks (compare with the results in Table 9).

Heterogeneity in 2×2 Difference-in-Difference (DD) Estimate. The empirical strategy of equation 1 operates as a difference-in-difference (DD) model that compares the outcome of cohorts with higher versus lower exposure to college expansions (treatment). However, the OLS estimation of the DD estimator in a two-way fixed effect framework, that the treated group receives

the treatment at different points of time (in comparison to a pre-post and treatment-control DD estimation), compares the outcome of all combinations of two-by-two treatment-control/pre-post pairs. The least-square coefficient is finally a weighted average of all these comparisons with the weights in the proportion of how long the pair had received the treatment and also the variance of the treatment (Goodman-Bacon, 2021). For instance, the OLS compares the outcomes of those cohorts who lived in counties that experienced an expansion to those in counties with no colleges at all (treatment vs never treated), it compares the outcomes of those who experienced a college expansion to those that had an expansion earlier and no new expansion after that (later treatment vs earlier control), and those who had an expansion earlier versus those who had later (earlier treatment vs later control). To explore this pairwise heterogeneity of the OLS-produced DD estimation, we implement bacon-decomposition which shows the coefficients of the pairwise DD estimates for 4-year colleges on age-at-death against their respective weights. In so doing, we collapse the data at the county by birth cohort level and assigned a dummy for treatment based on exposure to a college opening.[19] The results are reported in Figure 5. While there is heterogeneity in the pairwise DD coefficients, all three types of comparisons reveal positive effects. The largest weight is assigned to the comparisons between those with a college expansion versus those with no colleges in their county (weight=0.92, DD-coefficient=0.26).[20] Comparing earlier expansions as the treatment to later expansions as the control group (weight=0.04, DD-coefficient=0.41) or comparing later expansions as treatment and earlier expansions as control (weight=0.04, DD-coefficient=0.26) reveal average DD coefficients that are quite close to the treated-vs-never-treated

[19] To implement this analysis, we assigned a dummy for treatment that takes a value of one if the county experienced a new opening from year t t to year $t-1$ and zero otherwise.

[20] In Appendix F, we replicate the event-study analysis for a subsample that includes treated versus never-treated observations as the bacon-decomposition puts a larger weight on this subset. The results are quite similar to the pattern observed in Figure 4.

comparison as well as the overall DD estimation (=0.266). Therefore, though the OLS effects of DD estimation reveals various effects that are sometimes negative, there is no evidence to suggest that the main results are driven by one specific type of comparison.

As an additional check for the robustness of the effects to other DD estimations, we implement an event study analysis based on two-way fixed effect estimation introduced by de Chaisemartin & D'Haultfœuille (2020). They show that the overall DD estimation produced by OLS could be biased downward due to negative weights assigned to some pairs in group-time DD coefficients. They propose an alternative method that is free of negative weight contamination. We replicate the event-study analysis using their proposed method.[21] The results are reported in Figure 6. We observe a similar pattern as the event-study results produced by OLS. Interestingly, the post-treatment coefficients of 4-year college opening (top panel) reveal larger effects than those reported in Figure 4. This suggest that some negative weights assigned by OLS (and observed in the bacon-decomposition of Figure 5) bias the overall effects downward.

6. Conclusion

The benefits of education, as shown by a large body of literature, can go beyond its labor market returns. This study investigated the potential long term effects of college expansion on mortality and longevity. Our results shed new light on the long-run health benefits of education and add to the literature on social returns of education and, specifically, college education. We found that a new 4-year college construction in the local area increases any college education by about 5.2 percent from the mean and adds to the average schooling by 0.09 additional years. It also contributes to age-at-death by about 0.13 months. A back-of-an-envelope calculation suggests that

[21] In footnote 10, we show the regressions that we estimate for the event study analysis.

the effects could be as large as 12 months for those who do attend a 4-year college after a college opening and would have not attended if the college had not been constructed.

An event-study analysis showed that for unaffected cohorts college expansions had no effects, reducing concerns over pre-trends in the outcome, while the effects start to rise for the affected cohorts, those younger than 17 at the time of college expansion. We also explored whether the county-expansion-induced migration of possibly college-educated cohorts to college-expanding counties has driven our results by regressing the observable characteristics of individuals on college counts. The evidence is not consistent and strong enough to point to this selective migration. Moreover, a series of placebo tests, in which we assign the measure of colleges to individuals in ages later than 17, supports our empirical method.

To illustrate the contribution of college expansion to longevity of the US population, we extrapolate our findings to the second half of twentieth century college openings and mortality trends. Between the years 1940 to 1990, there has been 489 new 4-year college construction across US counties. The share of local area (own and neighboring county) that were affected by the expansion accounts for 0.4 percent of the US population. Between the years 1970-2018, the average age at death, conditional on survival up to age 47, increased from 72.19 to 77.09 years.[22] Considering the fact that college educated individuals accounted for 18.5 percent of deaths and implementing the results of Table 6 and first stage results of Table 8, one can calculate that the 1940-1990 college expansions increased average age at death for the whole US population by 4.35

[22] We look at this time period rather than 1940-1990 (period of college expansion calculation) since we have assigned college expansion at age 17 and we restrict our attention to those aged at least 47, per Numident death coverage. This means that the earlier cohort that we observe must have been 17 years old in 1940 and have reached age 47, which points to a mortality window starting at 1970.

months. This is equivalent to 7.5 percent of the observed increase in the longevity trend between the years 1970-2018.[23]

[23] Since the effects of Table 6 are local average treatment effects, which shows the changes in outcome for those in local areas rather than the whole population, we need to multiply the effects by the share of locally affected population to the whole US population. The final number is calculated as the product of total new college construction (489), the inverse value of first stage effect (1/0.0108), the main effect of 4-year college on longevity (0.1318), the share of college educated people in death records (0.185), and the average share of local population to the whole US population over the same period (0.004).

References

Acemoglu, D., & Angrist, J. (2000). How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual*, *15*, 9–59. https://doi.org/10.1086/654403

Albouy, V., & Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics*, *28*(1), 155–168. https://doi.org/10.1016/J.JHEALECO.2008.09.003

Angrist, J. D., & Keueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, *106*(4), 979–1014. https://doi.org/10.2307/2937954

Backlund, E., Sorlie, P. D., & Johnson, N. J. (1999). A comparison of the relationships of education and income with mortality: the national longitudinal mortality study. *Social Science & Medicine*, *49*(10), 1373–1384. https://doi.org/10.1016/S0277-9536(99)00209-9

Bailey, M., Clay, K., Fishback, P., Haines, M., Kantor, S., Severnini, E., & Wentz, A. (2016). U.S. County-Level Natality and Mortality Data, 1915-2007. *Inter-University Consortium for Political and Social Research*. https://doi.org/https://doi.org/10.3886/E100229V4

Bautista, M. A., González, F., Martinez, L. R., Munoz, P., & Prem, M. (2020). Does Higher Education Reduce Mortality? Evidence from a Natural Experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3717849

Behrman, J. R., Kohler, H.-P., Jensen, V. M., Pedersen, D., Petersen, I., Bingley, P., & Christensen, K. (2011). Does More Schooling Reduce Hospitalization and Delay Mortality? New Evidence Based on Danish Twins. *Demography*, *48*(4), 1347–1375. https://doi.org/10.1007/S13524-011-0052-1

Black, D. A., Hsu, Y. C., & Taylor, L. J. (2015). The effect of early-life education on later-life mortality. *Journal of Health Economics*, *44*, 1–9.

Buckles, K., Hagemann, A., Malamud, O., Morrill, M., & Wozniak, A. (2016). The effect of college education on mortality. *Journal of Health Economics*, *50*, 99–114. https://doi.org/10.1016/J.JHEALECO.2016.08.002

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early Childhood Investments Substantially Boost Adult Health. *Science*, *343*(6178), 1478–1485. https://doi.org/10.1126/SCIENCE.1248429

Card, D. (1999). The Causal Effect of Education on Earnings. *Handbook of Labor Economics*, *3*(1), 1801–1863. https://doi.org/10.1016/S1573-4463(99)03011-4

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, *69*(5), 1127–1160. https://doi.org/10.1111/1468-0262.00237

Cellini, S. R., & Chaudhary, L. (2014). The labor market returns to a for-profit college education. *Economics of Education Review*, *43*, 125–140. https://doi.org/10.1016/J.ECONEDUREV.2014.10.001

Chevalier, A., Harmon, C., Walker, I., & Zhu, Y. (2004). Does Education Raise Productivity, or Just Reflect it? *The Economic Journal*, *114*(499), F499–F517. https://doi.org/10.1111/j.1468-0297.2004.00256.x

Clark, D., & Royer, H. (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*, *103*(6), 2087–2120. https://doi.org/10.1257/AER.103.6.2087

Conti, G., Heckman, J. J., & Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *The Economic Journal*, *126*(596), F28–F65. https://doi.org/10.1111/ECOJ.12420

Conti, G., Heckman, J., & Urzua, S. (2010). The Education-Health Gradient. *American Economic Review*, *100*(2), 234–238. https://doi.org/10.1257/AER.100.2.234

Cornaglia, F., Crivellaro, E., & McNally, S. (2015). Mental health and education decisions. *Labour Economics*, *33*, 1–12. https://doi.org/10.1016/J.LABECO.2015.01.005

Cowan, B. W., & Tefft, N. (2020). *College Access and Adult Health*. https://doi.org/10.3386/W26685

Cuijpers, P., & Smit, F. (2002). Excess mortality in depression: a meta-analysis of community studies. *Journal of Affective Disorders*, *72*(3), 227–236. https://doi.org/10.1016/S0165-0327(01)00413-X

Currie, J., & Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*, *118*(4), 1495–1532. https://doi.org/10.1162/003355303322552856

Cutler, D. M., & Lleras-Muney, A. (2006). Education and Health: Evaluating Theories and Evidence. *National Bureau of Economic Research*, 37. https://doi.org/10.3386/W12352

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/AER.20181169

de Walque, D. (2007). Does education affect smoking behaviors?. Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics*, *26*(5), 877–895. https://doi.org/10.1016/j.jhealeco.2006.12.005

Dickson, M., & Harmon, C. (2011). Economic returns to education: What We Know, What We Don't Know, and Where We Are Going—Some brief pointers. *Economics of Education Review*, *30*(6), 1118–1122. https://doi.org/10.1016/J.ECONEDUREV.2011.08.003

Doll, R., & Hill, A. B. (1956). Lung Cancer and Other Causes of Death in Relation to Smoking. *British Medical Journal*, *2*(5001), 1071. https://doi.org/10.1136/BMJ.2.5001.1071

Everett, B. G., Rehkopf, D. H., & Rogers, R. G. (2013). The Nonlinear Relationship Between Education and Mortality: An Examination of Cohort, Race/Ethnic, and Gender Differences. *Population Research and Policy Review 2013 32:6*, *32*(6), 893–917. https://doi.org/10.1007/S11113-013-9299-0

Fenelon, A., & Preston, S. H. (2012). Estimating Smoking-Attributable Mortality in the United States. *Demography*, *49*(3), 797–818. https://doi.org/10.1007/S13524-012-0108-X

Fischer, M., Karlsson, M., & Nilsson, T. (2013). Effects of Compulsory Schooling on Mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health 2013, Vol. 10, Pages 3596-3618*, *10*(8), 3596–3618. https://doi.org/10.3390/IJERPH10083596

Fletcher, J. M. (2015). New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. *Social Science & Medicine*, *127*, 101–107. https://doi.org/10.1016/J.SOCSCIMED.2014.09.052

Fletcher, J. M., & Frisvold, D. E. (2011). College selectivity and young adult health behaviors. *Economics of Education Review*, *30*(5), 826–837. https://doi.org/10.1016/J.ECONEDUREV.2011.04.005

Fletcher, J. M., & Frisvold, D. E. (2014). The long run health returns to college quality. *Review of*

Economics of the Household, *12*(2), 295–325. https://doi.org/10.1007/S11150-012-9150-0

Fruehwirth, J. C. (2014). Can Achievement Peer Effect Estimates Inform Policy? A View from Inside the Black Box. *The Review of Economics and Statistics*, *96*(3), 514–523. https://doi.org/10.1162/REST_A_00385

Galama, T. J., Lleras-Muney, A., & van Kippersluis, H. (2018). *The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence.* https://doi.org/10.3386/W24225

Gathmann, C., Jürges, H., & Reinhold, S. (2015). Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science & Medicine*, *127*, 74–82. https://doi.org/10.1016/J.SOCSCIMED.2014.01.037

Goldstein, J. R., Alexander, M., Breen, C., Miranda González, A., Menares, F., Osborne, M., Snyder, M., & Yildirim, U. (2021). Censoc Project. In *CenSoc Mortality File: Version 2.0. Berkeley: University of California*. https://censoc.berkeley.edu/data/

Gonzalez, F., & Quast, T. (2013). Non-Linearities in the Relationship between Aggregate Income and Mortality Rates. *Eastern Economic Journal 2013 41:1*, *41*(1), 51–69. https://doi.org/10.1057/EEJ.2013.36

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*. https://doi.org/10.1016/J.JECONOM.2021.03.014

Halpern-Manners, A., Helgertz, J., Warren, J. R., & Roberts, E. (2020). The Effects of Education on Mortality: Evidence From Linked U.S. Census and Administrative Mortality Data. *Demography*, *57*(4), 1513–1541. https://doi.org/10.1007/S13524-020-00892-6

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, *103*(6), 2052–2086. https://doi.org/10.1257/AER.103.6.2052

Jaffe, D. H., Eisenbach, Z., Neumark, Y. D., & Manor, O. (2006). Effects of husbands' and wives' education on each other's mortality. *Social Science & Medicine*, *62*(8), 2014–2023. https://doi.org/10.1016/J.SOCSCIMED.2005.08.030

Jamison, E. A., Jamison, D. T., & Hanushek, E. A. (2007). The effects of education quality on income growth and mortality decline. *Economics of Education Review*, *26*(6), 771–788. https://doi.org/10.1016/J.ECONEDUREV.2007.07.001

Jemal, A., Thun, M. J., Ward, E. E., Henley, S. J., Cokkinides, V. E., & Murray, T. E. (2008). Mortality from Leading Causes by Education and Race in the United States, 2001. *American Journal of Preventive Medicine*, *34*(1), 1–8.e7. https://doi.org/10.1016/J.AMEPRE.2007.09.017

Johnson, N. J., Sorlie, P. D., & Backlund, E. (1999). The impact of specific occupation on mortality in the U.S. National Longitudinal Mortality Study. *Demography 1999 36:3*, *36*(3), 355–367. https://doi.org/10.2307/2648058

Kalediene, R., & Petrauskiene, J. (2005). Inequalities in mortality by education and socio-economic transition in Lithuania: equal opportunities? *Public Health*, *119*(9), 808–815. https://doi.org/10.1016/J.PUHE.2004.11.004

Kamhöfer, D. A., Schmitz, H., & Westphal, M. (2019). Heterogeneity in Marginal Non-Monetary Returns to Higher Education. *Journal of the European Economic Association*, *17*(1), 205–244. https://doi.org/10.1093/JEEA/JVX058

Kravdal, Ø. (2008). A broader perspective on education and mortality: Are we influenced by other people's education? *Social Science & Medicine*, *66*(3), 620–636.

https://doi.org/10.1016/J.SOCSCIMED.2007.10.009
Lacroix, G., Laliberté-Auger, F., Michaud, P.-C., & Parent, D. (2019). *The effect of college education on health and mortality: Evidence from Canada*. https://doi.org/10.1002/hec.3975
Lager, A. C. J., & Torssander, J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proceedings of the National Academy of Sciences*, *109*(22), 8461–8466. https://doi.org/10.1073/PNAS.1105839109
Lefèbvre, M., Pestieau, P., & Ponthiere, G. (2018). Premature mortality and poverty measurement in an OLG economy. *Journal of Population Economics 2018 32:2*, *32*(2), 621–664. https://doi.org/10.1007/S00148-018-0688-X
Leigh, D. E., & Gill, A. M. (1997). Labor market returns to community colleges: Evidence for returning adults. *Journal of Human Resources*, *32*(2), 334–353. https://doi.org/10.2307/146218
Lindahl, M. (2005). Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human Resources*, *40*(1), 144–168. https://doi.org/10.3368/jhr.xl.1.144
Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, *72*(1), 189–221. https://doi.org/10.1111/0034-6527.00329
Lleras-Muney, A., Price, J., & Yue, D. (2020). *The Association Between Educational Attainment and Longevity using Individual Level Data from the 1940 Census*. https://doi.org/10.3386/W27514
Long, M. C. (2010). Changes in the returns to education and college quality. *Economics of Education Review*, *29*(3), 338–347. https://doi.org/10.1016/J.ECONEDUREV.2009.10.005
Lundborg, P., Lyttkens, C. H., & Nystedt, P. (2016). The Effect of Schooling on Mortality: New Evidence From 50,000 Swedish Twins. *Demography*, *53*(4), 1135–1168. https://doi.org/10.1007/S13524-016-0489-3
Luy, M., Giulio, P. Di, & Caselli, G. (2011). Differences in life expectancy by education and occupation in Italy, 1980–94: Indirect estimates from maternal and paternal orphanhood. *Http://Dx.Doi.Org/10.1080/00324728.2011.568192*, *65*(2), 137–155. https://doi.org/10.1080/00324728.2011.568192
Lynch, S. M. (2003). Cohort and life-course patterns in the relationship between education and health: A hierarchical approach. *Demography 2003 40:2*, *40*(2), 309–331. https://doi.org/10.1353/DEM.2003.0016
Madsen, M., Andersen, A.-M. N., Christensen, K., Andersen, P. K., & Osler, M. (2010). Does Educational Status Impact Adult Mortality in Denmark? A Twin Approach. *American Journal of Epidemiology*, *172*(2), 225–234. https://doi.org/10.1093/AJE/KWQ072
Malamud, O., & Wozniak, A. (2012). The Impact of College on Migration. *Journal of Human Resources*, *47*(4), 913–950. https://doi.org/10.3368/JHR.47.4.913
Martins, P. S., & Jin, J. Y. (2008). Firm-level social returns to education. *Journal of Population Economics 2008 23:2*, *23*(2), 539–558. https://doi.org/10.1007/S00148-008-0204-9
Mazumder, B. (2008). Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives*, *32*(Q II), 2–16.
Mazzonna, F. (2014). The long lasting effects of education on old age health: Evidence of gender differences. *Social Science & Medicine*, *101*, 129–138. https://doi.org/10.1016/J.SOCSCIMED.2013.10.042

Meara, E. R., Richards, S., & Cutler, D. M. (2017). The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981–2000. *Https://Doi.Org/10.1377/Hlthaff.27.2.350*, *27*(2), 350–360. https://doi.org/10.1377/HLTHAFF.27.2.350

Meghir, C., Palme, M., & Simeonova, E. (2018). Education and Mortality: Evidence from a Social Experiment. *American Economic Journal: Applied Economics*, *10*(2), 234–256. https://doi.org/10.1257/APP.20150365

Muller, A. (2002). Education, income inequality, and mortality: a multiple regression analysis. *BMJ*, *324*(7328), 23. https://doi.org/10.1136/BMJ.324.7328.23

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, *25*(1), 159–184. https://doi.org/10.1257/JEP.25.1.159

Preston, S. H., Glei, D. A., & Wilmoth, J. R. (2010). A new method for estimating smoking-attributable mortality in high-income countries. *International Journal of Epidemiology*, *39*(2), 430–438. https://doi.org/10.1093/IJE/DYP360

Psacharopoulos, G. (1985). Returns to Education: A Further International Update and Implications. *The Journal of Human Resources*, *20*(4), 583. https://doi.org/10.2307/145686

Ross, C. E., Masters, R. K., & Hummer, R. A. (2012). Education and the Gender Gaps in Health and Mortality. *Demography*, *49*(4), 1157–1183. https://doi.org/10.1007/S13524-012-0130-Z

Ruggles, S., Flood, S., Goeken, R., Grover, J., & Meyer, E. (2020). IPUMS USA: Version 10.0 [dataset]. *Minneapolis, MN: IPUMS*. https://doi.org/10.18128/D010.V10.0

Savelyev, P. A. (2020). Conscientiousness, Extraversion, College Education, and Longevity of High-Ability Individuals. *Journal of Human Resources*, *58*(1), 0918-9720R2. https://doi.org/10.3368/JHR.58.1.0918-9720R2

Savelyev, P. A., Ward, B. C., Krueger, R. F., & McGue, M. (2022). Health endowments, schooling allocation in the family, and longevity: Evidence from US twins. *Journal of Health Economics*, *81*, 102554. https://doi.org/10.1016/J.JHEALECO.2021.102554

Schulz, R., Drayer, R. A., & Rollman, B. L. (2002). Depression as a risk factor for non-suicide mortality in the elderly. *Biological Psychiatry*, *52*(3), 205–225. https://doi.org/10.1016/S0006-3223(02)01423-3

Snyder, S. E., & Evans, W. N. (2006). The effect of income on mortality: Evidence from the social security Notch. *Review of Economics and Statistics*, *88*(3), 482–495. https://doi.org/10.1162/rest.88.3.482

Spoerri, A., Schmidlin, K., Richter, M., Egger, M., & Clough-Gorr, K. M. (2014). Individual and spousal education, mortality and life expectancy in Switzerland: a national cohort study. *J Epidemiol Community Health*, *68*(9), 804–810. https://doi.org/10.1136/JECH-2013-203714

van Kippersluis, H., O'Donnell, O., & van Doorslaer, E. (2011a). Long-Run Returns to Education. *Journal of Human Resources*, *46*(4), 695–721. https://doi.org/10.3368/JHR.46.4.695

van Kippersluis, H., O'Donnell, O., & van Doorslaer, E. (2011b). Long-Run Returns to Education. *Journal of Human Resources*, *46*(4), 695–721. https://doi.org/10.3368/JHR.46.4.695

Wulsin, L. R., Vaillant, G. E., & Wells, V. E. (1999). A systematic review of the mortality of depression. *Psychosomatic Medicine*, *61*(1), 6–17.

Zajacova, A. (2006). Education, gender, and mortality: Does schooling have the same effect on mortality for men and women in the US? *Social Science & Medicine*, *63*(8), 2176–2190. https://doi.org/10.1016/J.SOCSCIMED.2006.04.031

Tables

Table 1 - Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Female	3,849,792	0.426	0.4945	0	1
White	3,849,792	0.9067	0.2908	0	1
Black	3,849,792	0.0893	0.2852	0	1
Hispanic	3,849,792	0.0182	0.1335	0	1
First Generation Immigrant	3,849,792	0.0034	0.0578	0	1
Second-generation Immigrant	3,849,792	0.1446	0.3517	0	1
Birth Year	3,849,792	1929.293	4.7755	1923	1940
Death Year	3,849,792	1998.5607	4.6756	1988	2005
Death Age (Years)	3,849,792	68.7814	6.2309	47	82
Death Age (Months)	3,849,792	831.3591	74.7004	565	995
4-Year Colleges (Own+Neighboring):					
Total	3,849,792	11.7987	14.8451	0	63
Not in IPEDS List	3,849,792	11.6559	14.6745	0	63
Not Granting at Least Associate Degree	3,849,792	11.5155	14.3269	0	60
Private	3,849,792	2.4138	2.5338	0	13
Public	3,849,792	9.0399	12.3974	0	53
2-Year Colleges (Own+Neighboring):					
Total	3,849,792	3.4337	5.9778	0	44
Not in IPEDS List	3,849,792	2.6455	5.0074	0	40
Not Granting at Least Associate Degree	3,849,792	2.6713	5.0598	0	40
Private	3,849,792	1.4873	4.3511	0	38
Public	3,849,792	1.1234	2.1722	0	14
4-Year College (Own County)	3,849,792	3.1103	5.0768	0	26
2-Year College (Own County)	3,849,792	.9759	2.3556	0	18

Table 2 – Endogenous Migration: College Expansion and Endogenous Changes in Demographic and Socioeconomic Characteristics of Individuals

	Outcomes:								
	Female	White	Black	Hispanic	First Generation Immigrants	Second Generation Immigrants	Mother's Years of Schooling	Father's Years of Schooling	Family SEI Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. County and Birth Cohort FE									
4-Year	-0.00001	-0.00049	0.00065	0.0013	-0.00195***	-0.0161***	0.09736***	0.11161***	-0.16778**
College	(0.00083)	(0.00101)	(0.00095)	(0.00091)	(0.00068)	(0.00564)	(0.03323)	(0.03547)	(0.07409)
2-Year	0.00113	0.00202**	-0.00126	0.00102*	-0.0009**	-0.00677	0.04205	0.04438	-0.11867
College	(0.00101)	(0.00094)	(0.00077)	(0.00056)	(0.00042)	(0.00959)	(0.05122)	(0.05455)	(0.07944)
Observations	3855600	3855600	3855600	3855600	3855600	3855600	3543964	3308720	3611645
R-squared	0.00344	0.23268	0.24211	0.25398	0.01063	0.18534	0.12994	0.13006	0.01839
Mean DV	0.43	0.91	0.09	0.02	0.00	0.14	8.18	7.78	3.16
Panel B. County and Birth Cohort FE+County by Birth Cohort Trend									
4-Year	0.0026**	-0.00068	0.00032	-0.00027	-0.00021	0.00107	0.00445	0.00627	0.01953
College	(0.00104)	(0.00062)	(0.0006)	(0.0014)	(0.00028)	(0.00168)	(0.01287)	(0.01103)	(0.02482)
2-Year	0.00033	-0.00089	0.00078	-0.00019	0.0001	0.00161	-0.01918	-0.0077	-0.03119
College	(0.00118)	(0.00061)	(0.00056)	(0.00081)	(0.00035)	(0.00137)	(0.01225)	(0.00873)	(0.02751)
Observations	3855600	3855600	3855600	3855600	3855600	3855600	3543964	3308720	3611645
R-squared	0.00442	0.23462	0.24401	0.25585	0.01284	0.19193	0.134	0.1337	0.0197
Mean DV	0.43	0.91	0.09	0.02	0.00	0.14	8.18	7.78	3.16

Notes. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3 - College Openings and Endogenous Changes in Contemporaneous Health Outcomes

	Infant Mortality per 1,000 Live Births	Total Mortality per 1,000 Population	Fertility Rate (Births per 1,000 Women)	Share of Births Born in Hospital
	(1)	(2)	(3)	(4)
4-Year College	0.15513	0.02608	-1.37559	0.00442
	(0.33724)	(0.06561)	(1.19055)	(0.0029)
2-Year College	0.26838	-0.00843	-1.99237	0.00558
	(0.42777)	(0.065)	(1.25396)	(0.00387)
Observations	54252	54230	54212	42169
R-squared	0.56034	0.8054	0.91067	0.97033
Mean DV	33.678	9.503	101.401	0.836

Notes. Each column represents a separate regression. Regressions include county fixed effects, year fixed effects, and county-specific linear time trend. The data covers the years 1940-1958. Standard errors, clustered at the county level, are in parentheses.

*** p<.01, ** p<.05, * p<0.1

Table 4 - Placebo Tests: Assigning Colleges at Ages Later than 17

	Outcome: Age-at-death (Months)								
	Colleges Assigned at Age 25			Colleges Assigned at Age 30			Colleges Assigned at Age 35		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
4-Year College	0.0414	0.0414	-0.0882	0.0549	0.0504	-0.0249	0.0538	0.0549	0.0306
	(0.0508)	(0.0562)	(0.0551)	(0.0485)	(0.0528)	(0.0692)	(0.0406)	(0.0445)	(0.0643)
2-Year College	0.0676*	0.0762**	-0.0646	0.0348	0.0446	-0.0764*	0.0386	0.0423	0.0321
	(0.035)	(0.0366)	(0.0627)	(0.0279)	(0.0294)	(0.0406)	(0.0283)	(0.0299)	(0.0359)
Observations	3849792	3849792	3849792	3849792	3849792	3849792	3849792	3849792	3849792
R-squared	0.4448	0.4469	0.4474	0.4448	0.4469	0.4474	0.4448	0.4469	0.4474
Mean DV	831.4	831.4	831.4	831.4	831.4	831.4	831.4	831.4	831.4
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
County by Birth Cohort Linear Trend	No	No	Yes	No	No	Yes	No	No	Yes

Notes. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.

*** p<.01, ** p<.05, * p<0.1

Table 5 - The Association between College Openings and Successful Merging of Numident with the 1940-Census Data

	Outcome: Successful merging between Numident and 1940 census		
	(1)	(2)	(3)
4-Year College	0.0002	-0.0002	0.0004
	(0.0004)	(0.0004)	(0.0003)
2-Year College	-0.0014***	-0.0015***	-0.0005
	(0.0004)	(0.0004)	(0.0004)
Observations	40,300,107	40,300,107	40,300,107
R-squared	0.0133	0.0158	0.0162
Mean DV	0.09523	0.09523	0.09523
%Change 4-Year College	0.2576	-.1621	0.3747
%Change 2-Year College	-1.423	-1.626	-0.546
County FE	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes
Individual Controls	No	Yes	Yes
County by Birth Cohort Linear Trend	No	No	Yes

Notes. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 6 - Main Results: College Expansion and Age-at-death

	Outcome: Death Age (Months)		
	(1)	(2)	(3)
4-Year College	0.1922***	0.2029***	0.1318**
	(0.0445)	(0.0462)	(0.0671)
2-Year College	-0.0233	-0.0278	0.0473
	(0.0605)	(0.066)	(0.0539)
Female		6.6777***	6.6864***
		(0.0892)	(0.0894)
Other		-2.6102***	-2.6399***
		(0.5485)	(0.5493)
Black		-3.8***	-3.8327***
		(0.1204)	(0.1209)
Hispanic		0.2661	0.2797
		(0.2815)	(0.2851)
First Generation Immigrant		1.4886**	1.5458***
		(0.5854)	(0.5818)
Second-generation Immigrant		0.7931***	0.8181***
		(0.1084)	(0.1086)
Observations	3849792	3849792	3849792
R-squared	0.4448	0.4469	0.4474
Mean DV	831.4	831.4	831.4
Elasticity of 4-Year College	0.00273	0.00288	0.00187
Elasticity of 2-Year College	-0.00009	-0.00011	0.00020
County FE	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes
Individual Controls	No	Yes	Yes
County by Birth Cohort Linear Trend	No	No	Yes

Note. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

Table 7 - Robustness Checks of College Expansion and Age-at-death

	Column 3 Table 6	Adding Mother-Father Schooling+Family SEI Score	Adding House Value+Father Wage	Adding Mother Fixed Effect	Huber-White Robust SE	SE clustered at Birth Cohort	SE clustered at state-level
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
4-Year College	0.1318**	0.1557**	0.1476	0.1933	0.1318**	0.1318**	0.1318**
	(0.0671)	(0.0718)	(0.1053)	(0.2181)	(0.0626)	(0.0601)	(0.0608)
2-Year College	0.0473	-0.0008	-0.048	-0.0765	0.0473	0.0473	0.0473
	(0.0539)	(0.0597)	(0.1031)	(0.2069)	(0.0671)	(0.0478)	(0.0624)
Observations	3849792	3195730	1298744	951210	3849792	3849792	3849792
R-squared	0.4474	0.4556	0.4293	0.6887	0.4474	0.4474	0.4474
Mean DV	831.4	828.8	838.1	831.0	831.4	831.4	831.4

	SE clustered at county by birth cohort level	Adding a polynomial function of county population	Adding Birth Month by Birth Year FE	Adding Month of Death FE	Birth State=State of Residence at 1940	Age in 1940<15	Drop colleges if not in IPEDS list
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
4-Year College	0.1318*	0.1301**	0.1332**	0.1281*	.1307*	.1545**	0.1253*
	(0.0744)	(0.0629)	(0.0672)	(0.0672)	(.069)	(.0682)	(0.0672)
2-Year College	0.0473	0.0602	0.0484	0.0475	.0595	.0646	0.0536
	(0.0622)	(0.0579)	(0.0538)	(0.0547)	(.0557)	(.0575)	(0.0638)
Observations	3849792	3849792	3849792	3849792	3472541	3124362	3849792
R-squared	0.4474	0.4474	0.4491	0.4484	.4525	.4069	0.4474
Mean DV	831.4	831.4	831.4	831.4	829.7	817.5	831.4

	Drop colleges if not granting at least an associate degree	Public Colleges	Private Colleges	Outcome in logarithm (Semi-Log specification)	Outcome: Age at Death>60	Outcome: Age at Death>70	Replace Outcome with Death Age in Years
	(15)	(16)	(17)	(18)	(19)	(20)	(21)
4-Year College	0.1175*	0.1584	0.1343	0.0002**	.0013***	.0009**	0.0108*
	(0.07)	(0.1289)	(0.0823)	(0.0001)	(.0005)	(.0004)	(0.0056)
2-Year College	0.0668	-0.0046	0.1423	0.0001	.0007*	-.0002	0.004
	(0.0651)	(0.07)	(0.1228)	(0.0001)	(.0004)	(.0006)	(0.0045)
Observations	3534637	3488475	3849792	3849792	3849792	3849792	3849792
R-squared	0.4605	0.4272	0.4474	0.4489	.2336	.2773	0.4465
Mean DV	828.4	824.6	831.4	6.7	0.9	0.4	68.8

Notes. Each column represents a separate regression. All regressions include county fixed effect, birth year fixed effect, a county trend in birth year, and individual controls. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, which are clustered at the county level except for columns 6-9, are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 8 - College Expansion and Education-Income at 1960 Census

	First Stage Outcomes:					Reduced Form Outcomes:	
	Education: At Least One Year of College	Education: At Least Two Years of College	Education: At Least Three Years of College	Education: At Least Four Years of College	Years of Schooling	Log of Wage Income	Log of Total Income
	(1)	(2)	(3)	(4)	(5)	(7)	(8)
4-Year College	0.0108***	0.009***	0.007***	0.0048***	0.0933***	0.0568**	0.0378**
	(0.0033)	(0.003)	(0.0024)	(0.0018)	(0.0277)	(0.0233)	(0.0181)
2-Year College	0.0005	-0.0018	-0.0039	-0.0012	0.0397	0.0211	0.0067
	0(.0035)	(0.0029)	(0.0025)	(0.0019)	(0.0314)	(0.0286)	(0.02)
Observations	341834	341834	341834	341834	341834	341834	341687
R-squared	0.0547	0.0473	0.0385	0.0337	0.127	0.316	0.3798
Mean DV	0.204	0.156	0.113	0.086	11.252	5.145	5.404
Elasticity of 4-Year College	0.2284	0.2492	0.2673	0.2409	0.03568	0.04745	0.03008
Elasticity of 2-Year College	0.00405	-0.01906	-0.05687	-0.02376	0.00583	0.00677	0.00206
PUMA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PUMA by Birth Cohort Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. *Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county-PUMA level, are in parentheses.*
*** p<0.01, ** p<0.05, * p<0.1

Table 9 - Heterogeneity of the Main Results by Gender, Race, and Ethnicity

	Outcome: Death Age (Months)			
	(1)	(2)	(3)	(4)
Male×4-Year College	0.027***			
	(0.01)			
Male×2-Year College	0.0619**			
	(0.0257)			
Male	-7.2191***			
	(0.0926)			
Other×4-Year College		0.0826		
		(0.0653)		
Other×2-Year College		0.035		
		(0.1134)		
Other		-3.5842***		
		(0.627)		
Black×4-Year College			-0.0117	
			(0.0141)	
Black×2-Year College			0.1178***	
			(0.0349)	
Black			-4.017***	
			(0.1455)	
Hispanic×4-Year College				0.0238
				(0.0213)
Hispanic×2-Year College				-0.0744**
				(0.0315)
Hispanic				0.5247
				(0.3231)
4-Year College	0.1173*	0.1311*	0.1325**	0.1312*
	(0.0696)	(0.0672)	(0.0668)	(0.0671)
2-Year College	0.0126	0.0467	0.0407	0.0498
	(0.0523)	(0.0538)	(0.0538)	(0.0537)
Observations	3849792	3849792	3849792	3849792
R-squared	0.4474	0.4474	0.4474	0.4474
Mean DV	831.4	831.4	831.4	831.4
Birth Cohort FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes

Notes. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.

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

Figures

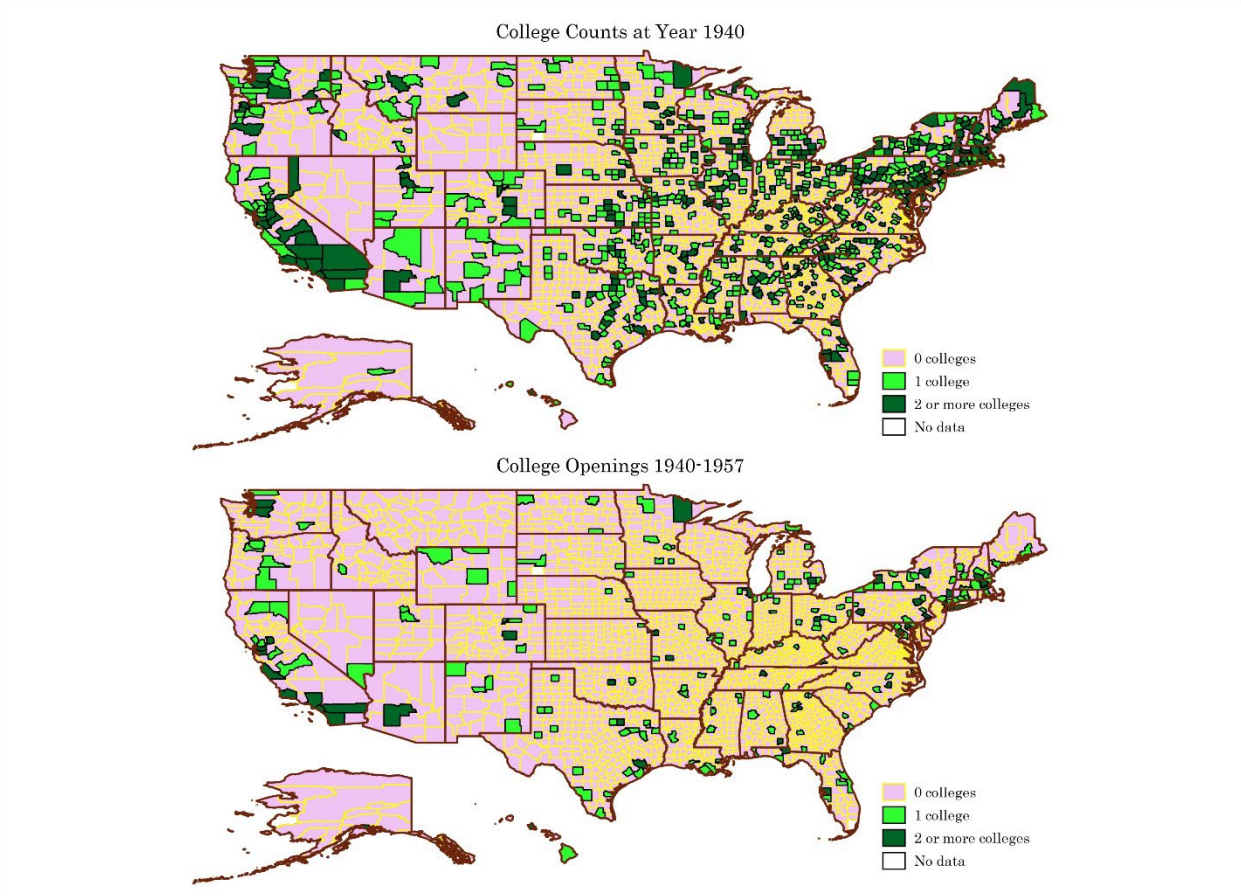


Figure 1 - Geographic Distribution of College Inventory at 1940 and College Expansion over the Years 1940-1957

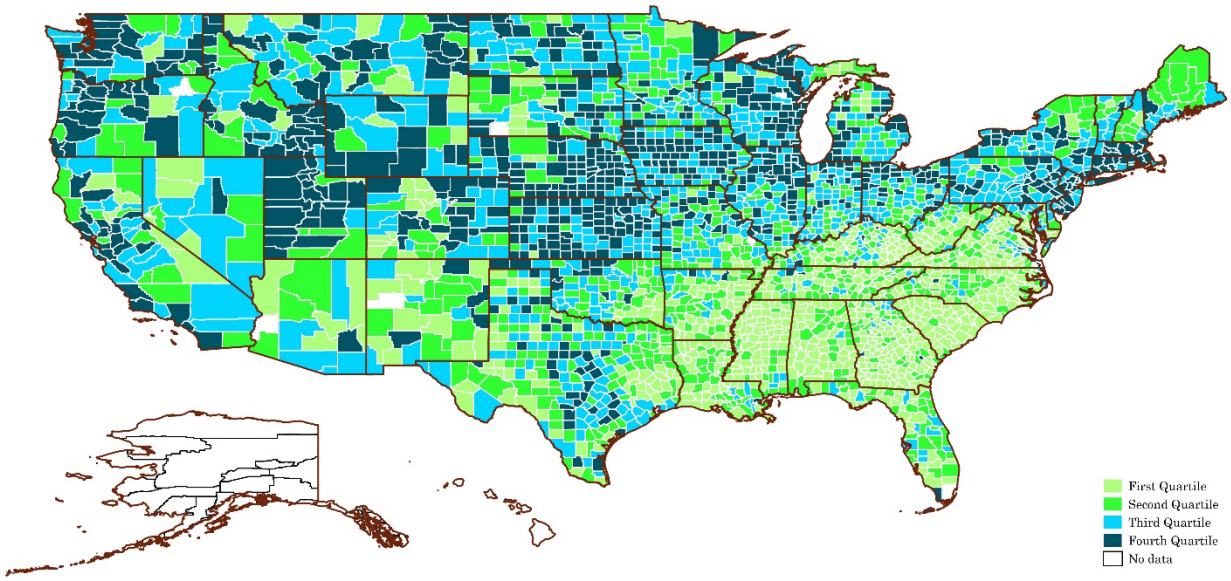


Figure 2 - Geographic Distribution of Age-at-death by Place of Residence during Childhood and Adolescence at 1940 for Cohorts Born between 1923-1940 and Died between 1988-2005

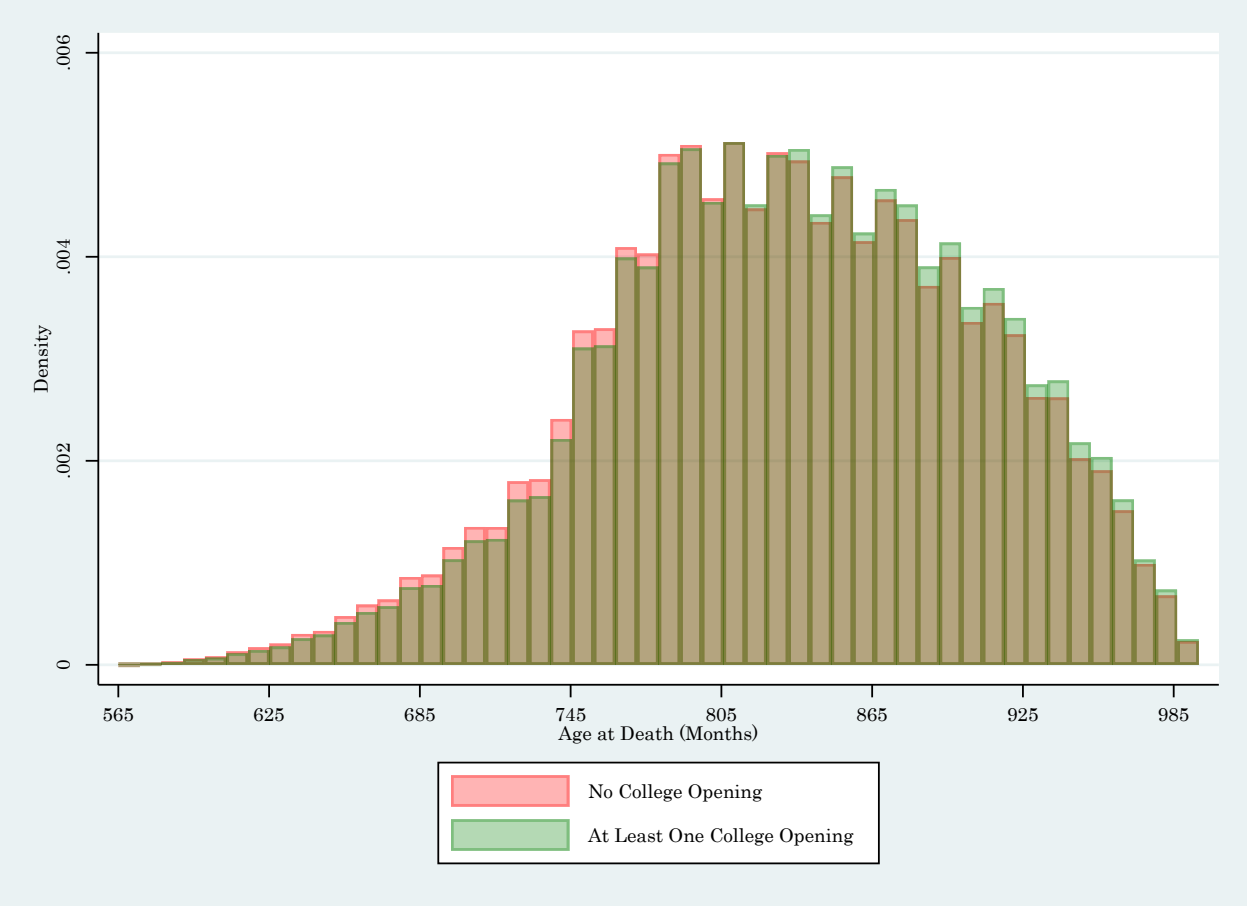


Figure 3 - Density Distribution of Age-at-death for Counties without/with College Expansion

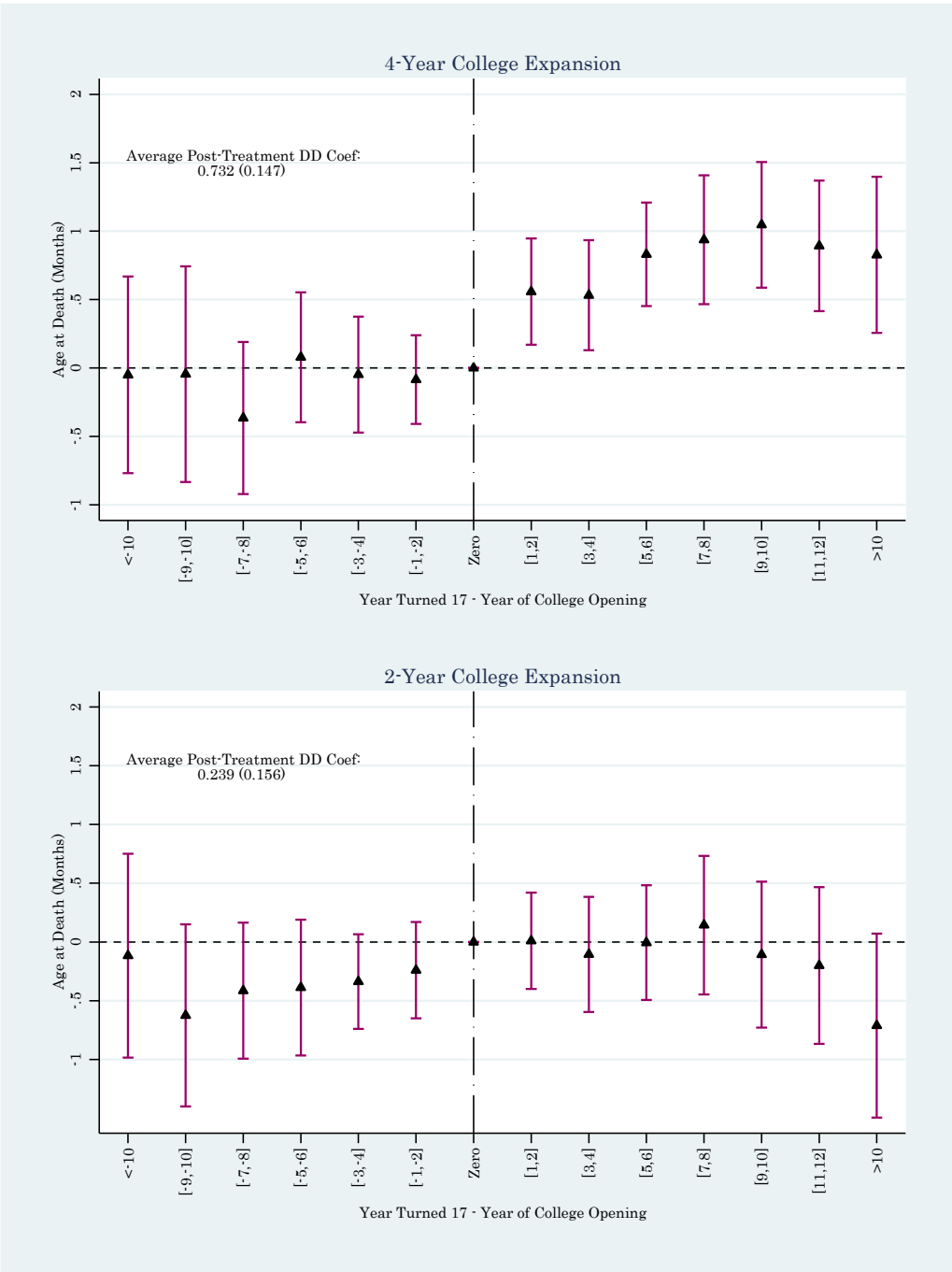


Figure 4 - Event-Study Results of College Expansion on Age-at-death

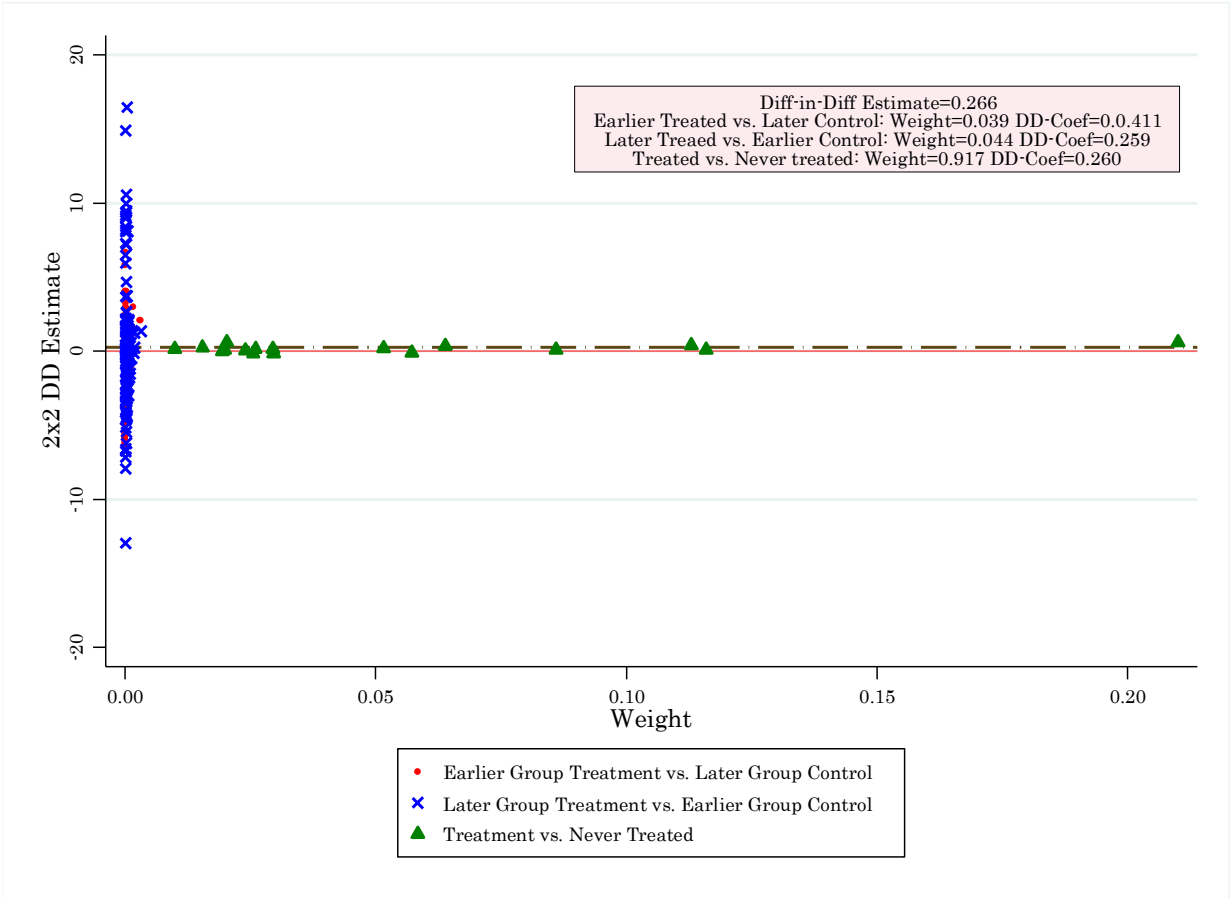


Figure 5 - Bacon Decomposition of Difference-in-Difference Estimates

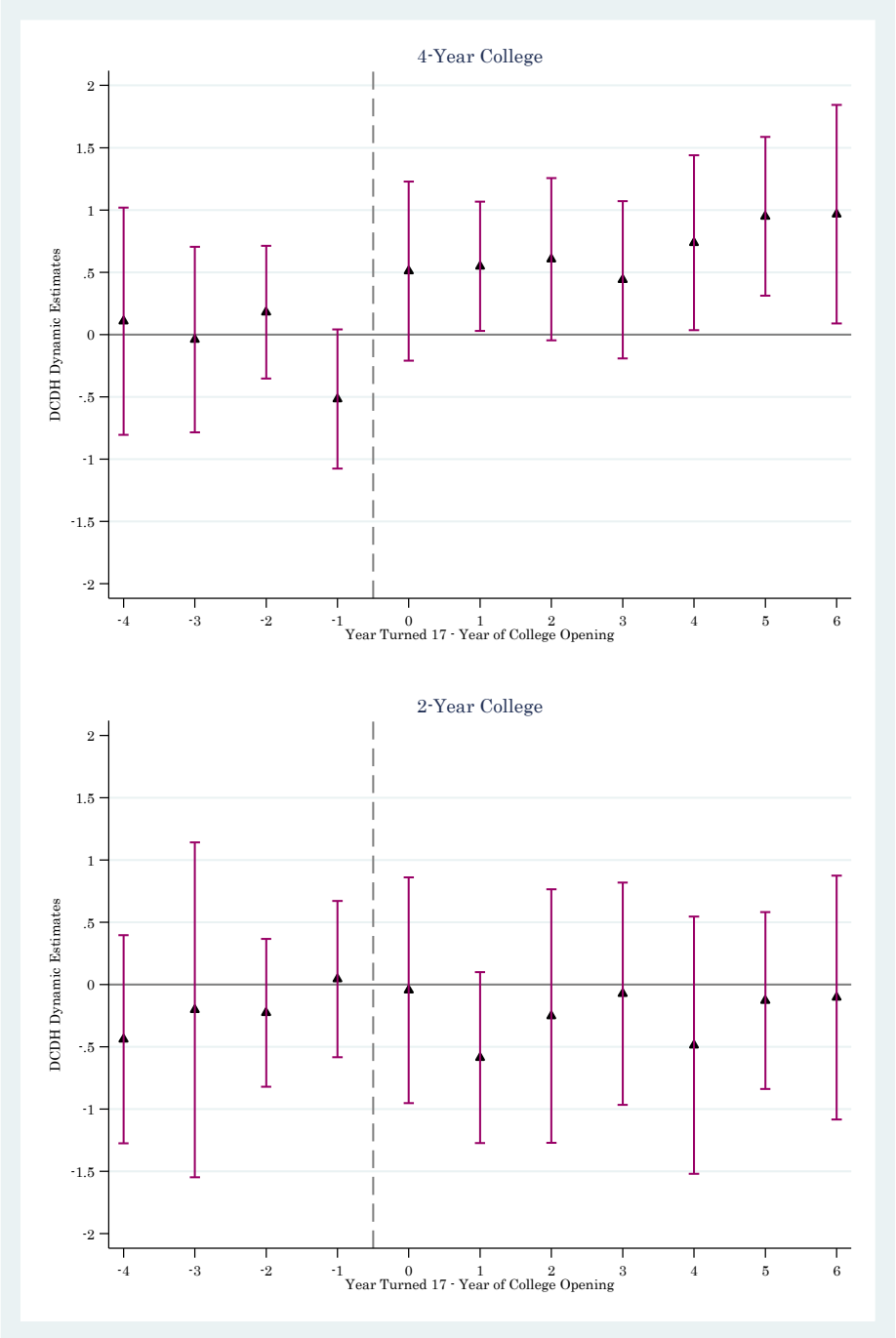


Figure 6 - Event Study of De Chaisemartin and D'Haultfoeuille Dynamic Treatment Difference-in-Difference Estimates

Appendix A

In the main analysis of the text, we aggregate the number of colleges to those of own county colleges and within-state neighboring counties. This appendix shows the results for two alternative measures of colleges: number of own county colleges and number of own county colleges in addition to all neighboring county colleges regardless of being within the same state or not. The effects, reported in Appendix Table A-1, are comparable to the main findings of Table 6.

Appendix Table A-1 - Alternative Measures of College Aggregation

	Outcome: Death Age (Months)/Explanatory variable:					
	Only Own County Colleges			Own County+Neighboring County Colleges (No within State Restriction)		
	(1)	(2)	(3)	(4)	(5)	(6)
4-Year College	0.2954***	0.3125***	0.2301*	0.213***	.2248***	0.1873***
	(0.107)	(0.1134)	(0.1346)	(0.0429)	(0.0443)	(0.0723)
2-Year College	-0.0211	-0.0292	0.0172	-0.0355	-0.04	0.0445
	(0.143)	(0.1567)	(0.1076)	(0.0615)	(0.0669)	(0.0549)
Female		6.68***	6.6889***		6.6778***	6.6864***
		(0.0892)	(0.0894)		(0.0892)	(0.0894)
Other		-2.6128***	-2.641***		-2.6106***	-2.6405***
		(0.5487)	(0.5494)		(0.5483)	(0.5493)
Black		-1.202**	-1.2048**		-1.1889**	-1.1922**
		(0.5494)	(0.5499)		(0.5491)	(0.5499)
Hispanic		0.3035	0.3084		0.2614	0.2794
		(0.2853)	(0.2883)		(0.2811)	(0.2847)
First Generation Immigrant		1.4406**	1.5335***		1.508**	1.5474***
		(0.5818)	(0.5793)		(0.5875)	(0.5825)
Second-generation Immigrant		0.7804***	0.8144***		0.7987***	0.8179***
		(0.1074)	(0.1085)		(0.1086)	(0.1086)
Observations	3855600	3855600	3855600	3849792	3849792	3849792
R-squared	0.4448	0.4469	0.4474	0.4448	0.4469	0.4474
Mean DV	831.4	831.4	831.4	831.4	831.4	831.4
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	No	Yes	Yes
County by Birth Cohort Linear Trend	No	No	Yes	No	No	Yes

Note. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.

Appendix B

This appendix replicates the main results of Table 6 for the DMF dataset which covers male's death between 1975-2005 that are linked to the 1940 census. While the marginal effects of the period that matches the death in Numident data are very close to the main results, the earlier death data reveals relatively smaller coefficients.

Appendix Table B-1- Replicating the Main Results using Censoc DMF Data 1975-2005

	DMF, Males 1975-2005	DMF, Males 1988-2005	Numident, Males 1988-2005	Numident, Females 1988-2005
	(1)	(2)	(3)	(4)
4-Year College	.0335 (.2286)	.1339 (.1081)	.1286 (.0962)	.129 (.09)
2-Year College	.1939 (.1208)	.0545 (.0836)	.0058 (.0707)	.1011 (.0949)
Observations	2000822	1530277	2209701	1640090
R-squared	.1905	.4337	.4399	.4536
Mean DV	779.9	825.2	826.7	837.6
Elasticity of 4-Year College	.00053	.00199	.00183	.00182
Elasticity of 2-Year College	.0009	.00024	2.4e-05	.00041
County FE	Yes	Yes	Yes	Yes
Birth Cohort FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
County by Birth Cohort Linear Trend	Yes	Yes	Yes	Yes

Note. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

Appendix C

Appendix Table C-1 reports the results of two-sample two-stage least square regressions in which the college education (having at least one year of college education) is instrumented with the presence of 4-year and 2-year colleges at the local area of residence. The local area in this part is either a county or a PUMA whichever is larger. In areas where a county constitutes several PUMAs, we consider county as the local area. In areas where a PUMA represents several counties, we consider PUMA as the local area. In all other areas, PUMA (which represents a county) is the aggregated measure of local area.

In the census 1960, we observe the education of individuals and in the census 1940 we observe individuals' county of residence which is used to capture the number of colleges when they turn 17. Since these two samples are not linked, we apply a two-sample 2SLS method in which the primary independent variable (educational attainment) comes from the 1960 census and the outcome (mortality) comes from the 1940 census that is linked to Numident. To eliminate cohorts who may have not yet completed education, we focus on individuals aged 22 and above. To mitigate the migration issues, we implement two additional restrictions. We keep individuals whose state of residence 5 years ago is different than their current state of residence. We restrict the sample to those who are less than 30 at 1960. The results are comparable without these exclusions.

Appendix Table C-1 - College Education and Age-at-death: Two-Sample 2SLS Results

	Outcome: Age-at-death (Months):		
	(1)	(2)	(3)
Education: Some College or More	25.430***	25.141***	19.562*
	(9.373)	(9.351)	(11.616)
Observations	1,144,557	1,144,557	1,144,557
Mean DV	786.6	786.6	786.6
Birth Year FE	Yes	Yes	Yes
PUMA-County FE	Yes	Yes	Yes
Individual Controls	No	Yes	Yes
PUMA-County by Birth Year Trend	No	No	Yes

Note. Each column represents a separate regression. The excluded instruments are 4-year and 2-year colleges at the PUMA level. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the PUMA level, are in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

Appendix D

This appendix shows the heterogeneity of the first stage results (Table 8) by race/gender/ethnicity. Each panel in Appendix Table D-1 shows the results of first stage where our measures of college expansion is interacted with dummies for gender, race, and ethnicity.

Appendix Table D-1 - Heterogeneity in the First Stage Effects by Demographic Characteristics

	Outcomes:			
	Education: At Least One Year of College	Education: At Least Two Years of College	Education: At Least Three Years of College	Education: At Least Four Years of College
	(1)	(2)	(3)	(4)
Panel A				
4-Year College×Male	0.0018	0.0016	0.0018*	0.0017**
	(0.0012)	(0.0011)	(0.001)	(0.0008)
2-Year College×Male	0.0006	0.0003	-0.0009	-0.001
	(0.0015)	(0.0014)	(0.0012)	(0.001)
4-Year College	0.010***	0.0083***	0.0062***	0.0041**
	(0.0032)	(0.0028)	(0.0023)	(0.0017)
2-Year College	0.0002	-0.0019	-0.0034	-0.0008
	(0.0035)	(0.0029)	(0.0025)	(0.0019)
Male	0.0533***	0.0515***	0.0428***	0.038***
	(0.0039)	(0.0034)	(0.0031)	(0.0024)
R-Squared	0.0549	0.0475	0.0387	0.0338
Panel B				
4-Year College×Other	0.0017	0.0027	0.0019	0.0039
	(0.005)	(0.0052)	(0.0049)	(0.0045)
2-Year College×Other	0.0106	0.0055	0.0039	-0.0015
	(0.0074)	(0.0077)	(0.0071)	(0.0065)
4-Year College	0.0108***	0.009***	0.007***	0.0048***
	(0.0033)	(0.003)	(0.0024)	(0.0018)
2-Year College	0.0004	-0.0018	-0.0039	-0.0012
	(0.0035)	(0.0029)	(0.0025)	(0.0019)
Other	-0.0771***	-0.0507***	-0.0462***	-0.0342***
	(0.0168)	(0.0155)	(0.0131)	(0.0124)
R-Squared	0.0548	0.0474	0.0386	0.0337
Panel C				
4-Year College×Black	-0.0036	-0.0043**	-0.0042**	-0.004**
	(0.0023)	(0.0021)	(0.0019)	(0.0016)
2-Year College×Black	0.0039	0.0039*	0.0035*	0.0036**
	(0.0026)	(0.0023)	(0.0019)	(0.0016)

4-Year College	0.011***	0.0093***	0.0073***	0.005***
	(0.0034)	(0.003)	(0.0025)	(0.0018)
2-Year College	0.0001	-0.0022	-0.0042*	-0.0016
	(0.0035)	(0.0029)	(0.0025)	(0.0019)
Black	-0.1093***	-0.084***	-0.0648***	-0.0521***
	(0.008)	(0.0067)	(0.0056)	(0.0047)
R-Squared	0.0548	0.0475	0.0388	0.034
Panel D				
4-Year	-0.0008	-0.0022	-0.003*	-0.0024
College×Hispanic	(0.002)	(0.0019)	(0.0017)	(0.0015)
2-Year	-0.0039	-0.0005	0.0029	0.0025
College×Hispanic	(0.0028)	(0.0026)	(0.0024)	(0.0021)
4-Year College	0.0108***	0.0091***	0.0071***	0.0049***
	(0.0034)	(0.003)	(0.0025)	(0.0018)
2-Year College	0.0006	-0.0018	-0.004	-0.0014
	(0.0035)	(0.0029)	(0.0025)	(0.0019)
Hispanic	-0.1725***	-0.1382***	-0.1069***	-0.0855***
	(0.0108)	(0.0088)	(0.007)	(0.0059)
R-Squared	0.0548	0.0474	0.0385	0.0337
Observations	341834	341834	341834	341834
Mean DV	0.204	0.156	0.113	0.086

Notes. *Each column within each panel represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county-PUMA level, are in parentheses.*
*** p<0.01, ** p<0.05, * p<0.1

Appendix E

This appendix shows the heterogeneity of the results across census regions, birth cohorts, and based on county characteristics. We investigate the heterogeneity of the results across subsamples based on other observable dimensions. In so doing, we divide the sample by individuals' census region of residence in 1940 (columns 1-4), birth cohort (columns 5-6), father's education (columns 7-8), and county population (columns 9-10). The results are reported in Appendix Table E-1

Compared to the main results (column 3, Table 6), the effect of 4-year college is larger by a factor of 2 in the Midwest region. It drops by roughly 60 percent for the South and West regions and is zero for the Northeast region. The effects of 2-year college expansions are considerably larger for the Northeast, Midwest, and South regions. In southern states, an additional 2-year college is associated with 0.35 months higher age at death, almost 7.5 times larger than its aggregate effect in the main results. Both 4-year and 2-year college expansions are more effective for later cohorts than earlier cohorts. The effect of 2-year college expansions among cohorts born after 1932 is not only larger than the main results (by a factor of 6.5) but also statistically significant at 10 percent level. Not surprisingly, the college expansion is more effective for individuals whose father is more educated. Finally, low population counties benefit more from 4-year college expansions although the marginal effects are not precise and significant at 10 percent level (column 9 vs column 10).

Appendix Table E-1 - Heterogeneity of the Main Results by Subsamples

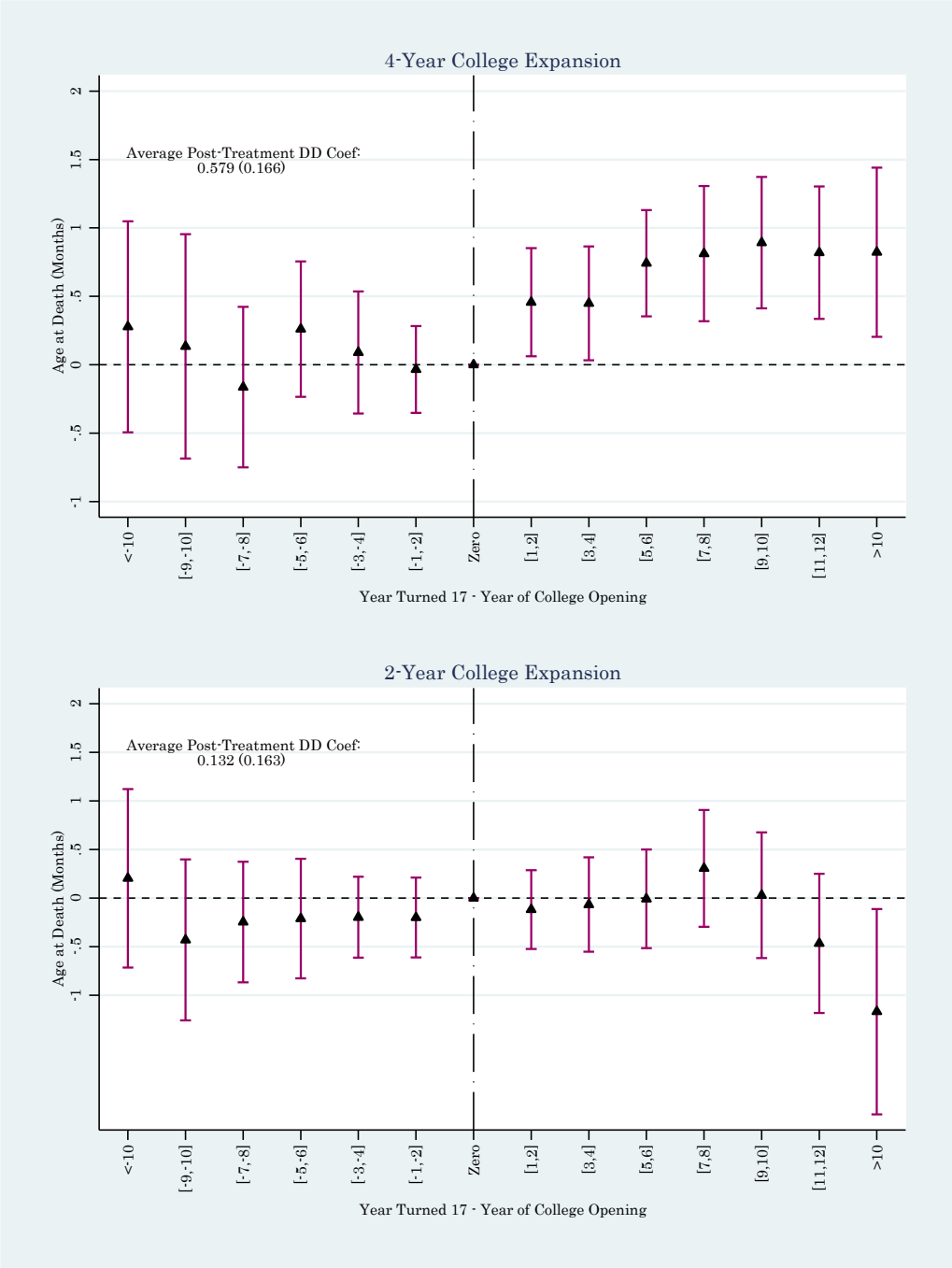
	Outcome: Death Age (Months)/Subsample:									
	Region: Northeast	Region: Midwest	Region: South	Region: West	Birth Year<1932	Birth Year≥1932	Below Median Father Education	Above Median Father Education	Below Median County Population	Above Median County Population
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
4-Year College	-0.008	0.2773**	0.0488	0.0592	0.0434	0.2851	0.101	0.2117**	0.318	0.1174*
	(0.1076)	(0.1118)	(0.2321)	(0.1403)	(0.0864)	(0.2123)	(0.0917)	(0.0948)	(0.2232)	(0.0692)
2-Year College	0.125	0.1644	0.3461	-0.1622	0.0566	0.3117*	0.026	0.0209	-0.2741	0.0706
	(0.1034)	(0.1459)	(0.2633)	(0.1206)	(0.0926)	(0.182)	(0.0858)	(0.0747)	(0.2327)	(0.0638)
Observations	952090	1223248	1297867	376587	2634682	1215110	2279673	1024386	1927602	1922190
R-squared	0.43	0.4484	0.4524	0.4609	0.1709	0.216	0.4379	0.4859	0.4532	0.4407
Mean DV	834.5	832.9	827.0	833.6	860.2	768.7	832.0	823.7	829.3	833.4
Birth Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Each column represents a separate regression. Individual controls include dummies for female, white, black, Hispanic, first-generation immigrant, and second-generation immigrant. Standard errors, clustered at the county level, are in parentheses.

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

Appendix F

The Bacon-decomposition (reported in Figure 5) suggest that the largest effects in producing the OLS estimations are driven by comparing treated and never treated counties (i.e., counties with an expansion to counties with no colleges at all). The event-study in the text (reported in Figure 4) shows no preexisting trend for the whole sample. In this appendix, we replicate the event-study for the sample that excludes counties that received the treatment before the start of the sample. Therefore, we focus on treated and never-treated counties. The results are reported in Appendix Figure F-1. The estimated coefficients and the observed pattern for both 4-year and 2-year college openings are quite similar to the event-study of the whole sample.



Appendix Figure F-1 - Event-Study Results of College Expansion on Age-at-death Using Treated and Never Treated Counties

Appendix G

In the main text, we provided suggestive evidence that education and income profiles improve as a result of new college openings and argue that these improvements are likely mechanism channels. Another possible change in the county characteristics that could partly explain the effects on longevity is compositional changes in occupations and industries. For instance, if a new college opening is accommodated by bringing new firms and establishments into the county and the new firms bring better-paid occupations, then the results could partly be driven by improvements in job quality. However, we do not have county-by-year level data on the share of different occupations over the sample period. One solution is to use decennial census data for the years 1940, 1950, and 1960. The drawback of the census is that for post-1950 years we do not have access to county identifiers. The variable PUMA which we exploited in section 5.3 is also only available for the 1960 census. Therefore, we are left with IPUMS de-identified counties that we do have access to for all the years 1940-1960. We should note that it limits our sample to about 500 counties, usually in metro areas with a large population. We assign the number of colleges to this data based on the year and de-identified county. We then restrict the sample to people aged 20-60 and those who report a specific occupation. We then build several dummy variables indicating the category of occupation. We regress these dummies on college measures, conditional on year fixed effects, birth-year fixed effects, county fixed effects interacted with birth year trend, and individual covariates. The results are reported in Appendix Table G-1. We do not observe a significant change in the composition of occupations across various categories. However, we find an increase in the share of people employed as managers and officials following a 4-year college opening. The impact suggests about a 7 percent increase from the mean of the outcome. Overall, these results do not offer a discernible pattern of changes in occupation composition as the pathway.

Appendix Table G-1 - Changes in Occupations Following College Openings

	Outcomes: Occupation in the following categories:								
	Professionals, Instructors, Doctors, Engineers, Other Technical	Farmers and Farm Managers	Managers, Officials, and Proprietors	Sales workers	Craftsmen and Repairmen	Apprentice Operatives, Mine Operatives, Other Operatives	Service Workers (private and public)	Farm Laborers	Other Laborers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
4-Year College	.0028	-.0055	.0064***	.0048	.0017	-.0052	.011	-.0161	.0013
	(.0047)	(.0036)	(.0018)	(.0064)	(.009)	(.0086)	(.011)	(.0137)	(.0031)
2-Year College	-.0013	.0004	.0003	.0038*	.0011	-.0015	.0012	-.0084	.001
	(.002)	(.0015)	(.001)	(.0023)	(.0035)	(.0033)	(.006)	(.0082)	(.0011)
Observations	1758358	1758358	1758358	1758358	1758358	1758358	1758358	1758358	1758358
R-squared	.0322	.0313	.045	.1602	.0644	.1168	.0542	.0309	.046
Mean DV	0.052	0.076	0.087	0.263	0.106	0.070	0.103	0.123	0.013
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year and Birth-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. *Each column represents a separate regression. Individual controls include dummies for female, white, and black. Standard errors, clustered at the county level, are in parentheses.*
*** p<0.01, ** p<0.05, * p<0.1