The Effects of U.S. Community Colleges on Educational Attainment and Longevity

Kevin P. Connolly*

July 22, 2024

Abstract

More than six hundred two-year public colleges opened in the United States between 1945 and 1980, creating major differences in access to college by location and date of birth for otherwise similar people. Using the variation arising from these college openings, I estimate the causal effect of greater access to college on college attainment and longevity using linked Census and Social Security Administration data. The college openings led to about one-tenth of a year of additional college attainment on average for nearby college-age men and women. The openings resulted in "democratization" and not "diversion"—that is, to a higher probability of completing four years of college—and to a two percentage point increase in one's likelihood of living beyond age 68.

*University of Chicago (email: kevinconnolly@uchicago.edu). I thank Christopher Avery, David Autor, Greg Bruich, Raj Chetty, David Cutler, James Feigenbaum, Edward Glaeser, Claudia Goldin, Joshua Gottlieb, Sebastian Haneuse, Thomas Kane, Lawrence Katz, Adriana Lleras-Muney, Jack Mountjoy, Andrei Shleifer, participants in the Harvard Labor/Public Economics seminar, the Harvard Applied Microeconomics seminar, NBER Fellows in Aging and Health and Harvard PIER Fellows for helpful comments and suggestions. I am grateful to J. Clint Carter for help in accessing restricted-use Census data. This paper is based on work supported by Harvard Lab for Economic Applications and Policy, the Lynde and Harry Bradley Foundation, the National Institute on Aging through grants T32AG000186 and 5T32AG000243, and the Institute of Education Sciences, U.S. Department of Education, through grant R305B150010 to Harvard University. The views expressed are those of the author and not those of the U.S. Census Bureau, the National Institute on Aging, the Institute of Education Sciences, or the U.S. Department of Education. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1743. (CBDRB-FY21-P1743-R9271)

1 Introduction

Without a college degree, life in America is staggeringly shorter.
–Anne Case and Angus Deaton[1]

Relative to Americans with a four-year college degree, the adult lifespans of Americans without a college degree are shorter by an average of eight and a half years (Case and Deaton, 2023). There are good reasons to think that college attainment causes greater longevity, such as through greater income, better health behaviors, and/or stronger social networks. At the same time, the strong college-longevity correlation may instead reflect family background or other important differences between those with and without a college degree. I estimate the causal effect of increased college attainment on longevity, using as a natural experiment the postwar openings of U.S. community colleges. I also specifically address the longstanding and contested question of whether four-year degree attainment is increased ("democratization") or decreased ("diversion") by community colleges.[2]

In order to estimate these effects, I have created a college openings dataset containing both the date of opening and exact latitude and longitude of colleges that opened between 1940 and 1980, as well as the locations of colleges that were already in existence before the openings. I have also developed algorithms to geocode about 70,000 towns/townships in the 1940 Census and over 100,000 locations of birth in the Social Security Numident ("Numident") dataset—which contains a record of every individual who has ever received a Social Security number and contains information on place of birth, date of birth and date of death. By taking advantage of recently completed work linking Numident to the 1940 Census, 2000 Census and the American Community Survey (Massey et al., 2018), I am able to advance in important ways prior work on several research questions—most notably, the causal effect of educational attainment on longevity.

In my analysis I consider towns which, prior to the opening of a two-year public

[1]Case, A. & Deaton, A. (2023, October 3). Without a College Degree, Life in America is Staggeringly Shorter. *New York Times*.

[2]One related school of thought concerns the acquisition of human capital (Becker, 1964/2009; Goldin & Katz, 2010). Another school concerns "signaling" (Spence, 1973; Caplan, 2019). To address the question of whether college attainment correlations reflect causality, researchers have used research designs involving variation in college attainment among twins (Ashenfelter & Krueger, 1994; Bound & Solon, 1999) and arising from variation across states in one's risk of being drafted for the Vietnam War (Buckles et al., 2016). Some have theorized that what is learned in college leads to different thinking and decision-making patterns that account for greater longevity (Cutler & Lleras-Muney, 2006; Cutler & Lleras-Muney, 2010; Glied & Lleras-Muney, 2008; De Walque, 2010); others that certain pre-existing characteristics (such as one's discount factor) lead to both greater attainment and greater health (Fuchs, 1982).

college, were at least twenty-five miles from the nearest existing college. Among those who lived in one of these towns as a child, the treatment group consists of those who were college-age or younger at the time of the opening, and the control group those (sharing the same childhood town) who were older than college-age. I estimate a treatment effect of the community college openings of about one-tenth of a year of additional college attainment, with little or no detectable effect for towns beyond five miles from the college opening. The community college openings increased college attainment for those in their mid-20s or younger at the time of the opening, and led on net to "democratization" and not "diversion"—a one percentage point increase in the likelihood of completing four years of college.[3]

The community college openings increased the share of individuals surviving to age 68 by two percentage points. The effect on longevity is confined, like the college attainment effect, to men and women whose childhood town was within five miles of a college opening and is larger for men than women. Taken at face value, the point estimates imply large effects of the community college openings on lifespan. These results support the hypothesis that college attainment causes greater longevity.

There is a substantial literature estimating the causal effects of education in the United States on older-age mortality.[4] The present paper is the first in this literature that is not rendered inconclusive by missing mortality data. All other causal estimates in this literature are based on non-random samples of the study population, where inclusion in the sample is determined by the outcome, longevity. This problem of endogenous sample selection, econometrically equivalent to the "bad control" problem, means that these estimates "...d[o] not have a causal interpretation."[5] Unlike the papers in this literature, I am able to observe individuals in their youth (at the time of the education treatment) and account for all mortality from that point through age 68. In this way, I am able to avoid estimating mortality effects on a longevity-selected sample of the study population. Indeed, I show (Figure 4) that in my context the use of a longevity-selected sample leads to a substantially

[3]In this paper I study *net* democratization/diversion. Other researchers (with data on college enrollment) have also obtained evidence on gross democratization/diversion, which is beyond the scope of this paper.

[4]Lleras-Muney (2005), Glied & Lleras-Muney (2008), Mazumder (2008), Lillard & Molloy (2010), Sanchez (2010), Sansani (2011), Frisvold & Golberstein (2011, 2013), Black, Hsu et al. (2015), Fletcher (2015), Buckles et al. (2016), Halpern-Manners et al. (2020), Aaronson et al. (2021), Moorthy et al. (2022), Savelyev et al. (2022), Kocks (2023), Fletcher & Noghanibehambari (2024a), Cowan & Tefft (2024).

[5]Angrist & Pischke (2009), p. 65. The "bad control" variable is implicit. When one estimates mortality effects on a longevity-selected sample of the study population, the estimation is econometrically equivalent to estimating mortality effects on the full study population with an indicator variable (a "bad control") in the regression equation that is equal to 1 if the individual is observed in the longevity-selected sample.

different conclusion about the causal effect of the college openings on longevity.[6]

There are five papers that estimate causal effects of college attainment on older-age mortality, three in the U.S. context (Buckles et al., 2016; Cowan & Tefft, 2024; Fletcher & Noghanibehambari, 2024a) and two in non-U.S. contexts (Lacroix et al., 2019; González et al., 2023). The present paper is the first—in the U.S. or any other context—to estimate a causal effect of college attainment on older-age mortality while avoiding the problem of endogenous sample selection. There are good reasons to think that college differs in major ways from lower levels of education in how it affects health and longevity, including in the development of personality traits (such as self-directedness and an orientation to hard work), specific skills and knowledge, general cognitive skills, and real and perceived control of one's life (Ross & Mirowsky, 1999). Given that eighty-seven percent of U.S. birth cohorts are now graduating from high school (Digest of Education Statistics, 2023, Table 219.10), understanding the effects of college attainment on longevity is of great importance.

A large and growing literature estimates the causal effects of lower levels of educational attainment (high school and junior high school) on older-age mortality.[7] Six pub-

[6]In order to defend the causal interpretation of the results of the papers in this literature, one would have to show that the outcome of mortality constitutes an exception to the Angrist-Pischke argument. In Figure 4, and in much greater detail in a separate paper (Connolly, 2024), I show that it is not reasonable to assume that the outcome of mortality constitutes such an exception.

If one applies extreme value bounds (Manski, 1989), as Gerber and Green (2012) recommend (Chapter 7), all of the mortality estimates on longevity-selected samples in this literature are rendered inconclusive.

In certain cases one can make an assumption of mortality-rate monotonicity (e.g., Aaronson et al., 2021, p. 538), which guarantees against a Type I error. This assumption however provides no guarantee against either a magnitude error or a Type II error. Moreover, there is no good reason for believing that the assumption of mortality-rate monotonicity will hold, as opposed to either making no assumption at all about the distribution of the treatment effect or making a weaker monotonicity assumption (such as first-order stochastic dominance, under which there are no guarantees against Type I errors, Type II errors or magnitude errors). See Connolly (2024).

As Montgomery et al. (2018) note about post-treatment bias (of which endogenous sample selection bias is a special case), "When we include a post-treatment variable [i.e., a variable observed after the time of the treatment and potentially affected by the treatment, such as the "bad control" indicator variable mentioned in footnote no. 5] in the set of conditioning variables either directly or indirectly... standard estimates... [will] be biased regardless of sample size, measurement precision, or estimation method... [T]he bias of standard estimates... can be in any direction and of any magnitude depending on the value of unknown (and unknowable) parameters....Once we have conditioned on a post-treatment variable, we have eliminated the assurance of unconfoundedness provided by randomization" (pp. 764-765). The authors stress the inadequacy of empirically testing for post-treatment bias, and that the best solution "is not to test for post-treatment bias, but rather to carefully design [experiments] that prevent it in the first place" (pp. 772-773). "Conditioning on post-treatment variables eliminates the advantages of randomization because we are now comparing *dissimilar* groups" (p. 762). "In short, conditioning on post-treatment variables can ruin experiments; [researchers] should not do it" (p. 761).

[7]Lleras-Muney (2005), Glied & Lleras-Muney (2008), Mazumder (2008), Albouy & Lequin (2009), Lillard & Molloy (2010), Sanchez (2010), van Kippersluis et al. (2011), Lager & Torssander (2012), Clark & Royer

lished papers, all in European contexts, are able to avoid the problem of longevity-related endogenous sample selection in estimating causal effects of increased compulsory schooling on older-age mortality (Albouy & Lequin, 2009; Clark & Royer, 2013; Gathmann et al. 2015; Fischer et al., 2013; Lager & Torssander, 2012; Meghir et al., 2018). The present paper is the first—in the U.S. or any other context—to estimate a (valid) causal effect of *voluntary* educational attainment at any level on older-age mortality. There are good reasons to think that voluntary education, in its effects on health and longevity, is fundamentally different from compulsory education—which by definition involves only those who would drop out if they could, and who may be doing the bare minimum academically to get by. Given that observed variation in educational attainment is mostly variation in voluntary attainment, the relevance of estimates of the effect of compulsory education may be very limited when it comes to understanding the well-documented positive correlation between educational attainment and longevity.

It may be that the causal effect of additional educational attainment on longevity varies by academic ability, interest in education, family background, or other factors that determine the educational attainment distribution. Changes in compulsory schooling laws by definition affect only individuals at the left tail of the overall distribution of educational attainment. The present paper is the first to estimate a causal effect of educational attainment on older-age mortality for individuals who are not at the left tail of the educational attainment distribution. Likewise, it may be that what matters (to some substantial extent) for longevity is *relative* level of education (Albouy & Lequin, 2009, p. 167), which is affected by the voluntary decisions of individuals to attain additional education but not by increases in compulsory schooling.[8] The present paper is the first to estimate an effect of a change in the distribution of educational attainment.

Of the six papers that avoid the problem of longevity-related endogenous sample selection, three (Albouy & Lequin, 2009; Clark & Royer, 2013; Gathmann et al. 2015) have two additional important limitations that I am able to overcome.[9] First, all three

(2013), Fischer et al. (2013), Gathmann et al. (2015), Black, Hsu, et al. (2015), Fletcher (2015), Meghir et al. (2018), Davies et al. (2018), Moorthy et al. (2022), Davies et al. (2023), Malamud et al. (2023), Masuda & Shigeoka (2023).

[8]Increases in compulsory schooling do not change the level of education of the treatment group relative to the rest of the population—the treatment group remains at the left tail of the distribution of educational attainment.

[9]Albouy & Lequin (2009) study two reforms in France, avoiding longevity-related endogenous sample selection for the 1968 reform but not for the 1938 reform, as do Clark and Royer (2013) in the U.K. for the

papers are unable to observe mortality for individuals who migrate internationally, and the estimates are therefore subject to bias from endogenous sample selection arising from international migration.[10] There are good reasons to think that one's propensity to migrate internationally is in part a function of one's educational attainment (Borjas, 1987); there is also empirical evidence that increases in compulsory schooling have a causal effect on one's propensity to migrate internationally (by nine percent in European countries; Fenoll & Kuehn, 2017). I am able to account for international migration by observing mortality with the Social Security Numident dataset, which records deaths for individuals with Social Security numbers regardless of whether they are living in the U.S., in a U.S. territory, or abroad (Finlay & Genadek, 2021, p. 143).

Second, the compulsory schooling reforms studied in these three papers consist of all-at-once, nationwide reforms. By contrast, I use a research design that involves more plausible control groups—specifically, control groups that were born in the same year as treatment groups (rather than relying on cross-birth-cohort comparisons), and control groups that shared the same childhood town as treatment groups, and hence shared the same local labor market.[11]

Longevity, it has been suggested, is the most highly-valued component of social welfare.[12] There is great interest among researchers in the causes of longevity in the U.S. context, with at least forty-nine papers in recent years estimating effects of early-life natural experiments on older-age mortality.[13] Only one paper (Aizer et al., 2016) has been able to

1972 reform but not for the 1947 reform.

[10]Clark and Royer (2013) test empirically for a treatment effect on the overall rate of international migration (p. 2099), an approach that cannot fully address the bad-control problem (Montgomery et al. 2018, p. 772-773), and acknowledge that a possible treatment effect on the composition of emigrants could result in bias in the education-mortality treatment effect estimator.

[11]This advantage is shared by Meghir et al. (2018): "One of the advantages of this research...is that the Swedish reform allows us to study two groups of people, born in the same years and active in the same labor markets, but having been educated by two different education systems at the same time...rather than comparing across different birth cohorts or across groups of people brought up in different states that may differ in numerous ways." (p. 236).

As van Kippersluis et al. (2011) note, "Given that year of birth is a relatively coarse running variable in a [regression discontinuity research design], one might wonder that other events occurred around the threshold that may have had long-run consequences for the health and mortality of cohorts born at that time." (p. 715). The running variable in Clark and Royer (2013) is month of birth rather than year of birth, and to that extent is less affected by this limitation.

[12]Preston (1977), p. 163. See also Murphy & Topel (2006).

[13]In addition to the eighteen papers cited in footnote no. 3, all of which concern an education treatment, this literature includes Bedard & Deschênes (2006), Black, Sanders, et al. (2015), Aizer et al. (2016), Hollingsworth et al. (2024), Jacks et al. (2024), Salm (2011), Johnson & Taylor (2019), Atherwood (2022), Seltzer (2021), Helgertz & Warren (2023), Breen (2024), Fletcher & Noghanibehambari (2023, 2024b),

do so while avoiding the problem of endogenous sample selection, and the authors are able to do so only because their dataset has special features that are generally unavailable to researchers studying the mortality effects of natural experiments in the U.S.[14] In this paper, I show how researchers can estimate older-age mortality effects in the U.S. while avoiding the problem of endogenous sample selection, without requiring the special features of the Aizer et al. (2016) dataset. As researchers are often interested in natural experiments that involve treatment variation at the sub-county level, the town/township-level geocoded data that I have developed can likewise be of great value.

In my use of college openings as a natural experiment, my research design is most closely related to Currie and Moretti (2003), who estimate effects of postwar openings of U.S. four-year colleges on outcomes recorded in birth records, including birthweight, gestational age, and the incidence of smoking during pregnancy. The authors do not address the question of democratization/diversion and do not study the outcome of mortality. I advance the Currie and Moretti (2003) research design[15] in four important ways. First, I conduct my analysis at the level not of county but town/township, enabling me to rule out county trends in the outcome variables as potential confounders and to observe the dose-response relationship between the treatment and outcomes of interest at different distances from the college opening (0-5 miles, 5-10 miles, etc.). This town-level research design enables me to obtain important geographic evidence that rules out certain otherwise-plausible violations of the exclusion restriction, evidence that could not be obtained with a county- or state-level research design. Second, my analysis is based on one's childhood town, not one's location as an adult, allowing me to rule out the possibility of "endogenous mobility bias."[16] Third,

Noghanibehambari (2022), Noghanibehambari & Engelman (2022), Noghanibehambari & Fletcher (2023a, 2023b, 2023c, 2023d, 2024a, 2024b), Noghanibehambari, Fletcher, Schmitz et al. (2024), Noghanibehambari & Noghani (2023), Vu et al. (2023), Karger (2024), Taylor (2020), Saavedra (2013).

[14]Aizer et al. (2016) are in effect able to use censored (rather than truncated) mortality data for a subset of their results. With censored mortality data and the right outcome variable, one can estimate mortality effects that are not subject to endogenous sample selection. See Connolly (2024).

[15]Card (1993) and Kling (2001) estimate returns to college using the 1966 National Longitudinal Survey of Youth (NLSY) and variation in proximity to college (specifically, whether there was a four-year college in one's county of residence in 1966), a research design that is valid if distance to college is as good as randomly assigned (either unconditionally or conditional on some set of observed pre-treatment covariates). Carneiro and Heckman (2002), using the 1979 NLSY, which includes Armed Forces Qualification Test (AFQT) score (a measure of academic ability), find that AFQT is correlated with distance to college, suggesting that distance to college in itself is not as good as randomly assigned. One way around this problem is to use variation in proximity to the nearest college arising from college openings, which (with the right dataset and control variables) could be as good as randomly assigned.

[16]Endogenous mobility bias is a form of endogenous sample selection, in which the study population is

I am able to observe family background information—most notably, parents' educational attainment—and use it in covariate balance regressions, providing important evidence on the plausibility of the key identification assumption for my analysis. Fourth, whereas Currie and Moretti focus on the openings of four-year colleges and study their effects on outcomes for white mothers observed in birth records, I estimate the effects of two-year public college openings for men and women of all races and for the outcomes of college attainment, democratization/diversion, and longevity.

There has been debate since at least 1960 about whether two-year public colleges increase college attainment ("democratization"; Medsker, 1960) or decrease attainment by diverting students from four-year colleges ("diversion"; Clark, 1960). While a community college could increase four-year degree attainment by substantially reducing the cost of the first two years of college, there are also reasons to think that a community college could decrease four-year degree attainment.[17] A number of researchers in the 1960s through the 1980s found evidence of diversion, providing the empirical basis for the landmark 1989 book by Brint and Karabel, *The Diverted Dream*.[18] In the years since the publication of *The Diverted Dream*, researchers have addressed the question of democratization/diversion for community colleges in the modern period (1980-present) with increasingly sophisticated methodology and found mixed results. One set of papers finds evidence of net democratization (Rouse, 1995; Leigh & Gill, 2003; Mountjoy, 2022), while another set (Doyle, 2009; Reynolds, 2012) finds evidence of net diversion. I find that the U.S. community college openings led on net to democratization.[19]

Educational attainment is "[t]he one social factor that researchers agree is consistently linked to longer lives in every country where it has been studied... It is more impor-

selected based on what is potentially an outcome affected by the treatment: a person's location in adulthood. College openings might affect geographic mobility—e.g., individuals moving to places with college openings, either for the college education opportunities or, say, for job openings arising near college openings—potentially producing biased estimates of the causal effects of the college openings.

[17]Possible reasons include tracking into vocational degree programs, peer effects, major differences in per-student resources, a less-academic environment, and the additional obstacle of transferring for the final two years of college attainment.

[18]Brint & Karabel (1989). *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900-1985.* See p. 130-131 for a summary of the evidence.

[19]The results in this paper are for two-year public college openings taking place in higher education "deserts" (locations at least 25 miles from the nearest college). The democratization/diversion effect of a two-year public college may be different when another college is nearby. The two-year public college evolved substantially over the course of the twentieth century (Cohen et al., 2014), such that conclusions about democratization or diversion in the modern era may not apply to earlier eras.

tant than race; it obliterates any effects of income."[20] The causal link between education and health/longevity—whether one exists, its direction, and its magnitude—is of great importance for understanding the total rate of return to educational attainment (Cutler & Lleras-Muney, 2006; Oreopoulos & Salvanes, 2011), for economic models of the demand for human capital (Ben-Porath, 1967; Becker, 2009; Hazan, 2009; Cervellati & Sunde, 2013) and health capital specifically (Grossman, 1972, 2000, 2017), as well as for understanding the demographic transition and the fundamental causes of long-run economic growth (Acemoglu & Johnson, 2007; Cervellati & Sunde, 2005, 2015; Soares, 2005; Galor & Weil, 2000; Galor, 2011). Clark and Royer (2013) interpret their results—no reduction in mortality from increases in compulsory schooling in the U.K.—as "suggest[ing] that economic models that assume a strong causal effect of education on health ought to be carefully reconsidered" (p. 2089). Subsequent causal evidence from compulsory schooling natural experiments (notably Meghir et al., 2018) has reaffirmed the Clark-Royer null result for the causal effect of compulsory schooling on mortality. The results of the present paper—a strong positive effect of college attainment in the U.S. on longevity—suggest to the contrary that economic models involving this key parameter may be justified in assuming a strong causal effect of education on health/longevity, depending on the nature of the educational attainment.[21]

Increasing the longevity of the U.S. population has been an important goal of U.S. policymakers for more than four decades (U.S. Surgeon General, 1979). More recently, U.S. policymakers have expressly adopted the promotion of educational attainment—increasing high school completion rates (by 4.2 percentage points by 2030) and college-going rates (by 4.6 percentage points)—as a means of promoting health and longevity (Pronk et al., 2021). Clark and Royer (2013) interpret their null result as "cast[ing] doubt on the health returns to education policies, such as the U.S. Healthy People 2010 goal of increasing the high school completion rate to 90%" (p. 2089). The results of the present paper suggest to the contrary that health/longevity returns are strongly positive for at least some forms of educational attainment.

Starting in 2026, a "demographic cliff" will dramatically reduce overall enrollments

[20]Kolata (2007). As Grossman (2006) puts it, "years of formal schooling completed is the most important correlate of good health [and longevity]." See also Grossman (2015).

[21]That is, the causal effect of educational attainment may depend in important ways on whether the attainment is college level or high school, the historical and institutional context, the part of the educational attainment distribution affected, and whether the attainment was voluntary or compulsory, among other potentially important factors.

at U.S. community colleges, with enrollments in certain U.S. regions expected to fall by an average of 20-30% between 2026 and 2029 (Grawe, 2018, Chapter 5). State policymakers are considering whether and when to close existing community colleges (Harnisch et al., 2024, Figure 1). The results in this paper provide important evidence on the potential effects of closing a local community college on educational attainment and longevity.

The rest of this paper is organized as follows. Section 2 contains historical background information on the community colleges. In Section 3, I explain the research design, data sources and how I geocode and link the Census and Social Security datasets in order to carry out the research design. Section 4 contains the results, and Section 5 is the conclusion.

2 Historical Context

In its landmark 1947 report, the President's Commission on Higher Education ("Truman Commission") wrote:

> The time has come to make education through the fourteenth grade available in the same way high school education is now available...To achieve this, it will be necessary to develop much more extensively than at present such opportunities as are now provided in local communities by the 2-year junior college...such community colleges probably will have to carry a large part of the responsibility for expanding opportunities in higher education (Zook, 1947, p. 37).

In the decades following the Second World War, there were both large expansions in the facilities of existing public colleges and a vast increase in the number of public colleges. Figure 1 is a map of the postwar college openings as well as colleges that were already in existence by 1945 (including public and private, two- and four-year colleges). Figure 2 is a line graph showing the number of colleges over time for each type of college. The majority of openings in the postwar period, over 700, were two-year public colleges, with the fastest period of growth occurring between 1960 and the early 1970s.[22] By 1958, one out of three freshmen in public colleges were enrolled in two-year public colleges (Medsker, 1960, p. 13), and by the late 1960s new two-year public colleges were opening at a nationwide rate of five per month (U.S. Bureau of the Census, 1975, p. 382). By 1968 a two-year public college had opened in every state except Nevada (Medsker & Tillery, 1971, pp. 24-25).

[22]For the purposes of this research project I exclude for-profit colleges. For more information on how I created this dataset see Section 3.1.

Whereas the openings of pre-1940 two-year public colleges were mostly a matter of local funding and initiative, state governments played a major role in funding and controlling the new two-year public colleges in the postwar period. States' plans were "simultaneously efforts to respond to demand, to expand access, and to coordinate public higher education along the lines of a hierarchically segmented system..."[23] While the pre-1940 two-year public colleges were mostly upward extensions of public high schools with a curriculum dominated by the liberal arts, in the postwar period most community colleges were independent of high schools and 30-40% of the curriculum was vocational. Moreover, the majority of students enrolled in the postwar community colleges were of working-class background (Brint & Karabel, 1989, p. 74). Very few non-traditional students enrolled at two-year public colleges until the late 1960s and 1970s, which saw a major increase in the willingness of students to enroll at ages later than their early 20s, particularly among men.[24]

3 Data and Empirical Strategy

3.1 Data Sources

My empirical analysis requires data of three kinds: Census and Social Security Administration datasets, college openings data and geographic information. Of the first kind are the decennial U.S. Censuses (1940 and 2000), the Social Security Numident dataset ("Numident"), and the 2001-2015 American Community Survey (ACS). I use the 2000 Census and the 2001-2015 ACS to observe educational attainment for those affected by the community college openings. Numident—which contains a record of every person who has ever received a Social Security number—provides information on date of birth and city/town of birth, with which I calculate the distance from one's town of birth to the nearest college. The variable "date of death" in Numident enables me to calculate the lifespan of those affected by the college openings.

The second kind of data concerns college openings. For my analysis I have compiled a dataset containing the locations of colleges in the United States for each year between 1940 and 1985. Figure 1 illustrates the colleges in the U.S. that opened during the period 1945-

[23] Specifically, state plans in New York, Pennsylvania, Washington, and Georgia (Brint & Karabel, 1989, p. 72).

[24] For more information on how the two-year public college in the U.S. has evolved through the twentieth century, see Cohen et al. (2014). *The American Community College*. Ch. 1.

1985, as well as those that were already in existence in 1945. The primary data source is the 1996 Integrated Postsecondary Education Data System (IPEDS) dataset, supplemented by the 2009 IPEDS dataset (for the latitude and longitude of many of the colleges) and the 1984 Higher Education General Information Survey (HEGIS) dataset (for information on the year coursework was first offered). As far as possible, I have geocoded the exact location of the college; when not possible, I have geocoded the town in which the college is located.

The third kind of data is geographic. In order to calculate the main treatment variable for my analysis (the distance to the nearest college from one's childhood town at the time one was college-age), I have geocoded as many towns and townships as possible in the 1940 Census and in Numident. To this end, by far the most important geographic data source is the Geographic Names Information System (GNIS), a national dataset of geographic entities compiled from existing U.S. government maps and supplemented by state and local maps and a variety of historical documents.

3.2 Geocoding and Record Linkage Procedures

For my analysis I have developed an algorithm to geocode as many towns and townships as possible in the 1940 Census, using the Geographic Names Information System (GNIS). When it comes to geocoding "place of birth" in Numident, I adapt an algorithm by Taylor et al. (2016) that geocodes individuals from Numident to their county of birth. My adapted algorithm assigns individuals to the latitude and longitude of their "place of birth." Approximately 140,000 places of birth can be geocoded using this adapted algorithm.

I also link Numident to the 1940 Census to observe location of residence in 1940 (for those born before 1940). This linkage was carried out by Massey et al. (2018) and the resulting crosswalk from Numident to the 1940 Census has been made available to researchers at the Census Bureau. Finally, I link Numident to the 2000 Census and the 2001-2015 American Community Survey, in both cases using crosswalks provided to researchers by the Census Bureau.

3.3 Imputation of Missing Dates of Death

Numident contains date of death information for only about half of the postwar community college openings study population (born between 1921 and 1951). Many of

those for whom date of death is missing are still alive in 2021; the rest are deceased and the dates of death were never recorded by the Social Security Administration. Numident date of death information is highly comprehensive for deaths occurring in 2000 or later, regardless of age at death (Finlay & Genadek, 2021, Table 1). There is some question about the comprehensiveness of Numident date of death information for deaths occurring between 1973 and 1999 to individuals who died before the age of 65.[25] Accordingly, one can impute missing dates of death using the full-count 2000 Census in the following way. If date of death is missing and the person is alive in the year 2000, it can reasonably be inferred that the person is still alive in 2021. On the other hand, if a person's date of death is missing and he is not observed in the 2000 Census, then it can be reasonably inferred that the person died before the year 2000, and therefore likely before 1973.[26] The full-count 2000 Census has been linked to Numident at very high linkage rates, particularly for the relevant demographic (men and women born in the U.S. between 1921 and 1951).[27] This date of death imputation procedure is summarized in Appendix Figure 1.

3.4 Research Design

The basic empirical specification is a two-way fixed effects model, given by the following regression equation:

$$Y_{iyc} = \beta_1 AgeDist_{yc} + \Sigma_c \beta_{2,c} Town_c + \Sigma_y \beta_{3,y} CountyBirthCohort_y + \epsilon_{iyc} \quad (1)$$

where i indexes individuals, y childhood-county-by-birth years, and c childhood towns. The treatment variable, $AgeDist_{yc}$, is an indicator for both the distance from one's childhood town to the nearby college opening and one's age group at the time of the opening (e.g., 0-5 miles from the college opening and age 10-25 at the time of the opening). One set of fixed effects, $\Sigma_c \beta_{2,c} Town_c$, is for childhood town, which ensures that the estimated treatment effect will be orthogonal to any fixed characteristics of one's hometown (e.g., a town's long history

[25] See Hill and Rosenwaike (2001). Table 1. Since 2002 the Social Security Administration has collected additional date of death information for deaths occurring prior to 2000 (Finlay & Genadek, 2021), so that the results in Hill and Rosenwaike (2001) can be taken as a lower bound on the age-at-death-specific comprehensiveness of the pre-2000 Numident mortality data.

[26] I report results in two ways: first, assuming that Numident coverage of deaths occurring between 1973 and 2000 are highly comprehensive regardless of age at death (see Table 4, column 6), and second, without making any assumption about such comprehensiveness (Table 4, column 7).

[27] See Appendix Table 1.

of supporting public education). The second set of fixed effects, $\Sigma_y \beta_{3,y} CountyBirthCohort_y$, is for each childhood-county-by-birth-year, which ensure that the estimated treatment effect will be orthogonal to any trends over time in the outcome at the county, region, state or national level (such as a county-wide increase in funding for elementary school).

Using this two-way fixed effects model, the results in this paper have a causal interpretation under the assumption of conditional mean independence:

$$E(\epsilon|AgeDist, CountyBirthCohort, Town) = E(\epsilon|CountyBirthCohort, Town) \quad (2)$$

That is, conditional on one's birth cohort (or childhood-county-by-birth-year) and childhood town, one's age at the time of the local college opening was as good as randomly assigned. This assumption is also known as "exchangeability" and captures the idea that in a natural experiment the treatment and control groups should be otherwise similar (or "exchangeable"). While this assumption of exchangeability cannot be directly tested, one can obtain evidence on its plausibility. This assumption would be implausible given evidence that the treatment and control groups were substantially different in some way relevant to the outcome: for example, if treated men and women had substantially greater aptitude for college or had parents who were substantially wealthier or better-educated on average.[28]

When it comes to estimating the causal effect of the college openings on longevity, for most specifications in this paper at least one more assumption—about the distribution of the treatment effect and/or the pattern of missing data—is required for unbiased and consistent causal inference. When the outcome is "Lifespan in Years," there are two additional assumptions. First, one must assume that the method of projecting lifespans (see section 3.3) for those who are still alive in 2021 accurately captures the pattern of mortality for such individuals. Second, 11% of the study population are neither observed in the full-count 2000 Census nor have a date of death recorded by the Social Security Administration. The required assumption is that whether an individual is missing from the mortality dataset must be uncorrelated with the outcome "Lifespan in Years" (Stock & Watson, 2019, pp. 297-298). Under these two assumptions, together with exchangeability, the estimator of the causal effect of the college openings on the outcome "Lifespan in Years" will be unbiased

[28] I use covariate balance regressions (Table 2) and event study analyses (Figure 3) to obtain evidence on whether the exchangeability assumption is reasonable.

and consistent.[29]

In contrast, when instead of "Lifespan in Years" the outcome variable is "Survival to (at least) Age 68," the number of required assumptions falls from three to one. The mortality dataset in this case is double-censored,[30] with survival to at least age 68 being observable for every individual without requiring any lifespan projections. It follows that, under the single assumption of exchangeability, the causal effect estimator for "Survival to (at least) Age 68" will be unbiased and consistent without any assumption at all about either the distribution of the treatment effect or about missing mortality data.[31]

Finally, in order to interpret the causal mortality estimate as the effect of additional college attainment in itself and not the general effect of the college opening, an exclusion restriction must hold. That is, it must be the case that the college openings affected longevity only by means of increasing college attainment. It is conceivable that the college openings affected the lifespans of the treatment group (relative to the control group) in some other way (e.g., by creating employment opportunities that disproportionately benefited local college-age individuals). If the exclusion restriction does not hold, the estimated effect of the college openings on lifespan would not be entirely attributable to changes in college attainment arising from the college openings, and the instrumental-variables estimator—of the effect of an additional year of college on lifespan—will be biased and inconsistent. While no longer fully attributable to college attainment, the estimated causal effect on longevity would remain attributable to the general effect of college openings, and hence has importance for public policy.

[29]The second assumption—that an indicator variable for an individual's date of death being missing from the mortality dataset is uncorrelated with the outcome "Lifespan in Years"—almost certainly does not hold, given that those with missing death data died before the year 2000 and to a large extent before 1973. However, there is good reason to think that the degree of bias and inconsistency falls with the overall comprehensiveness of the mortality dataset (Connolly, 2024). The mortality dataset used in this paper—which has high coverage of deaths occurring between 1973 and 2020, plus projections for deaths occurring after 2020—is substantially more comprehensive than the publicly available mortality datasets that have been used to estimate mortality effects such as CenSoc-Numident, CenSoc-DMF and NARA Numident. That is to say, the degree of bias and inconsistency—when the outcome is 'Lifespan in Years"—will be nonzero but substantially smaller than if one had instead used a more limited mortality dataset. See Connolly (2024).

[30]A mortality dataset is left-truncated at age 50 when individuals who died before 50 are completely unobserved, and it is left-censored at age 50 when such individuals are partially observed, showing up in the dataset but with a date of death known only to be before age 50 (i.e., the exact date of death is not observed).

[31]That is, whether an individual survived to age 68 is observed for every individual in the study population; There is therefore no missing mortality data about which an assumption is required.

4 Results

4.1 Summary Statistics

Table 1 contains summary statistics. The study population is split evenly by gender. The typical person was born in 1937, enrolled in Social Security at age 18, and attained about 1.2 years of college.[32] Three-quarters of the study population survived to at least age 65. Eleven percent have dates of death that are left-censored at 1972, and 38 percent have dates of death that are right-censored at the year 2020. The overall sample size is about 81 million people. The subset most affected by the openings—those whose childhood hometown was brought within 5 miles of the nearest college by a two-year public college opening—contains about one million people, 24 percent of whom have an educational attainment observable in the 2000 Census or the 2001-2015 American Community Survey.

4.2 Covariate Balance Regressions

Table 2 contains covariate balance regressions, which provide evidence on whether the key assumption of exchangeability is reasonable. The sample is restricted to men and women whose educational attainment is observed in the 2000 Census or the 2001-2015 American Community Survey and for whom both father's and mother's educational attainment is observed in the 1940 Census. In columns 1 and 2, I find that the treatment variable does not predict either mother or father's years of college. In columns 3 and 4, I estimate the treatment effect of the community college openings on years of college and years of lifespan, respectively. In columns 5 and 6, I control for father's and mother's college attainment, and I find that the estimated treatment effect of the college openings changes by about 5% or less. In this way, Table 2 provides support for the assumption of exchangeability.

4.3 Event Study Analysis

Figure 3 contains four event study graphs. Each graph displays an outcome variable as a function of one's age at the time of the community college opening for the event

[32]The fact that men and women in these birth cohorts are enrolling in Social Security in early life—typically as a teenager—is what enables the use of censored, rather than truncated, mortality data. That is, such individuals are still observed even if they died outside of the period (1973-2021) in which the mortality data is highly comprehensive, enabling the imputation of their date of death with information from the full-count 2000 Census, as explained in Section 3.3.

study population: men and women whose childhood town was brought within 5 miles of the nearest public college by a community college opening. The event study graph plots linear regression coefficients in a two-way fixed effects model, with childhood town/township fixed effects and childhood-county-by-birth-year fixed effects. The coefficients are estimated relative to a reference group, those age 32-34 at the time of the college opening. In Figure 3.1, the outcome variable is "Years of College" (as observed in the 2000 Census/2001-2015 American Community Survey). Relative to men and women age 32-34, there is no statistical difference for men and women age age 35-40 in the mean number of years of college attained. Beginning with those age 26-28 or younger at the time of the opening, there is an increase in completed years of college of between five-hundredths and one-tenth of a year for each age category, statistically significant at the five-percent level.[33]

In Figure 3.2, the outcome is now an indicator for completing at least a bachelor's degree, addressing the question of whether the community college openings led on net to democratization or diversion. Those 26-28 or younger were 1-2 percentage points more likely to attain a bachelor's degree relative to those age 32-34, which is evidence for net democratization. The outcome in Figure 3.3 is "Lifespan in Years," with right-censored lifespans projected on the basis of age and gender using the 2019 Social Security life tables. The point estimates in Figure 3.3 show an increase of about three-tenths of a year on average for those in their mid-20s or younger at the time of the opening, although for most age groups the effect is not statistically different from the age 31-33 reference group. In Figure 3.4, the outcome variable is the same ("Lifespan in Years") but the sample is no longer restricted to those for whom we observe educational attainment and is consequently about four times larger than the sample in Figure 3.3. In Figure 3.4, there is an increase in lifespan of about 0.5 years on average for men and women in their mid-20s or younger at the time of the opening, relative to men and women in the same towns who were age 31-32 when the local community college opened.

4.4 Regression Estimates

In the following regression tables, I keep the same two-way fixed effects model that I use in the event study analysis while consolidating the age-specific indicator variables.

[33]In these event study graphs, and indeed throughout all of the results in this paper, the standard errors are clustered at the level of childhood town/township.

In Table 3, the first explanatory variable combines the youngest five age categories from the Figure 3 event study graphs (i.e., ages 11-25), and the omitted/reference category combines the oldest five age categories (ages 26-40). The second explanatory variable in Table 3 is identical to the first explanatory variable except that it refers to those whose childhood towns were brought within 5 to 20 miles of the nearest college by a community college opening.

The outcome variables in Table 3 are "Years of College" and indicators for completing at least an associate's degree and at least a bachelor's degree. All specifications in Table 3 include childhood town and childhood-county-by-birth-year fixed effects. In column 1, the first coefficient can be interpreted as follows: Among those whose childhood town/township was brought within 5 miles of the nearest college by a community college opening, being age 10-25 at the time of the opening is associated with about 7.4 hundredths of a year of additional college attainment on average relative to those (with the same childhood town) who were age 26-40 at the time of the opening—a 6% increase in college attainment.[34] For those whose childhood town was brought within 5 to 20 miles, the magnitude of the effect is smaller by an order of magnitude and not statistically distinguishable from zero.

In column 3 of Table 3, the outcome variable is an indicator for completing at least an associate's degree. The effect of the college openings (for those brought within 5 miles of a college by a college opening) is a 1.7 percentage-point increase in the likelihood of completing at least an associate's degree. In column 5, the effect of the postwar openings is a 1.4 percentage-point increase in the likelihood of completing at least a bachelor's degree—that is, evidence for the democratization and against the diversion hypothesis. Across all three columns in Table 3, there is no detectable effect of the college openings for towns beyond five miles from a community college opening.

There is a long-running debate, dating back to at least 1960, about whether the net effect of two-year public colleges on college attainment is positive ("democratization"; Medsker, 1960) or negative ("diversion"; Clark, 1960).[35] A number of researchers in the

[34]The mean of the dependent variable is 1.16, yielding a percent change of $\frac{0.0739}{1.16} = 6\%$.

[35]In this paper I address the question of net democratization versus net diversion—that is, how access to two-year college affects overall college attainment and four-year degree attainment. Other researchers (with data on college enrollment) have also addressed the question of the number of people who are "diverted" from starting college at a four-year institution by increased access to a two-year institution, as well as the number of people for whom the two-year college democratizes higher education (i.e., who are induced to attend college by increased access to a two-year college but would otherwise not attend college at all).

1960s through the 1980s found evidence of diversion, as Brint and Karabel note in *The Diverted Dream* (1991). In the years since *The Diverted Dream*, researchers have addressed the question with increasingly sophisticated methodology and found mixed results. One set of papers, which uses distance from one's high school to the nearest two- and four-year colleges for identification, finds evidence of net democratization (Rouse, 1995; Leigh & Gill, 2003; Mountjoy, 2022). Another set of papers uses propensity score matching and sensitivity analyses (Doyle, 2009; Reynolds, 2012) and finds evidence of net diversion. Both sets of papers study two-year public colleges in the modern period (1980 onward), and in each case, the key identification assumption is either selection (only) on observable variables, or that the instrument is valid conditional on observable variables. By contrast, the community college openings in the present paper are part of an earlier era of the two-year public college in the U.S., the era of the the comprehensive community college (1940-1980). The two-year public college evolved considerably since its first appearance in Illinois around 1900, and the two-year college in one era may well have had a different effect on overall college attainment than in the other eras.[36]

In the primary specification of Table 3, I control for trends across county birth cohorts in college attainment by using childhood-county-by-birth-year fixed effects. In this way one can rule out any trend (at the county-level or higher) as an explanation for the observed increase in college attainment. Using childhood town/township fixed effects one can also rule out as an explanation any fixed characteristic of individual towns/townships.[37] One cannot, however, rule out in this way an upward trend in college attainment within towns/townships, a trend that would violate the key assumption of exchangeability. Elsewhere—in the event study analysis and the covariate balance regressions—I find evidence against this possible explanation.[38]

[36]One important difference, for example, is the share of students enrolled in vocational programs, which was very low in the junior college era (1900-1940) and roughly 30-40% in the era of the comprehensive community college (1940-1980), but over 50% in the modern era (1980-present). Another difference is that in the modern era but not the earlier eras, one could enroll in a vocational/occupational program and then transfer with credit to a degree program at a four-year college. See Cohen et al. (2014).

[37]The inclusion of childhood town/township fixed effects in all specifications means that the observed effect is estimated relative to individuals with the same childhood town/township; in this way, one can rule out the possibility that the observed increase in college attainment can be explained by any fixed characteristics of the towns/townships near the college openings (or to be more precise, any characteristic of the townships that is fixed with respect to all of a given town's birth cohorts in the study population, the 1921-1951 birth cohorts).

[38]In particular, I find no pre-trend in college attainment across the birth cohorts (those age 26-39 at the time of the local community college opening), and I also find in covariate balance regressions that there are

In Table 4, I estimate the effect of the postwar community college openings on longevity outcomes. In column 1, the outcome is lifespan in years, in which right-censored lifespans are projected on the basis of gender and age using the 2019 Social Security Administration period life tables.[39] The treatment group—those living in towns brought within 5 miles of the nearest college by a community college opening and age 10-25 at the time of the opening—lived an additional 0.48 years on average. An effect on longevity is not detectable for towns brought within 5-20 miles of the college opening (consistent with the geographic pattern for college attainment observed in Table 3, column 1). As noted in Section 3.4, this treatment-effect estimator in column 1 will be unbiased and consistent under three assumptions: exchangeability, the assumption that the lifespan projections for individuals still alive in 2021 accurately capture the pattern of mortality for those individuals, and the assumption that having a missing date of death is uncorrelated with the outcome "Lifespan in Years."

In column 2, the outcome variable is now an indicator for surviving to at least age 55, and the estimated treatment effect of the community college opening is 0.5 percentage points. For ages 60, 65 and 68, the estimates are 0.77, 1.0, and 1.3 percentage points, respectively. In each case, there is again no detectable effect beyond five miles from the college opening.[40] The treatment effect estimators in columns 2-5 will be unbiased and consistent under three assumptions: exchangeability, the assumption that having a missing date of death is uncorrelated with the outcome "Lifespan in Years," and the assumption that deaths occurring between 1973 and 1999 are highly comprehensive regardless of age-at-death.

In column 6, the mortality dataset is treated as left-censored instead of left-truncated. That is, individuals whose date of death was never recorded by SSA, and who are not enumerated in the 2000 Census, are assumed to have died in 1972 or earlier. In this case, the estimated treatment effect of the college openings on survival to at least age 68 is +2.75 percentage points, an effect that is about twice as large as the effect in column 5, where the dataset is left-truncated and right-censored. The estimator in column 6 is unbiased and con-

no differences in family background—as measured by mother's and father's college attainment—across the treatment and control groups (Table 2).

[39]Social Security Administration Period Life Table (2019). <https://www.ssa.gov/oact/STATS/table4c6.html>

[40]The ages in columns 2-5 (55, 60, 65, and 68) are chosen because they can be estimated without any additional projections of lifespans. One cannot estimate age 50, for example, without making some assumption about left-censored deaths, because men and women born in 1921 with left-censored deaths are age 52 in 1973. Likewise, one cannot estimate an effect for age 70, because men and women born in 1951 were 69 years of age in 2020.

sistent under two assumptions: exchangeability, and the assumption that deaths occurring between 1973 and 1999 are highly comprehensive regardless of one's age at death.

In column 7, I estimate exactly the same regression as in column 6, except that I restrict the birth cohorts to include only individuals who were born between 1932 and 1951. As individuals in these birth cohorts turned 68 between 2000 and 2019, and as 2000-2019 mortality data is highly comprehensive regardless of age at death, the second assumption for the estimator in column 6 is no longer necessary. The estimated treatment effect in column 7 of the college openings on lifespan is +2.0 percentage points, and the only assumption required for unbiased and consistent inference is exchangeability.

In Figure 4, I conduct my analysis as if I had a truncated (or longevity-selected) mortality dataset and compare the estimated treatment effect under censoring with the estimated effect under truncation. In the first column of Figure 4, I plot the point estimate and 95-percent confidence interval from column 1 of Table 4. In column 2 of Figure 4, I estimate exactly the same regression specification, except with a truncated dataset. Specifically, I left-truncate the mortality dataset at 1988 and right-truncate the mortality dataset at 2005, a truncation pattern corresponding to that of the CenSoc-Numident and NARA Numident mortality datasets (Goldstein et al., 2023; Hollingsworth et al., 2024). The point estimate in the second column is 0.082, with a 95% confidence interval of [-0.017, 0.181].[41] Comparing columns 1 and 2, the 95-percent confidence intervals are non-overlapping and the point estimates differ in magnitude by a factor of six. Moreover, the estimate under truncation in the second column is not statistically different from zero. In this context at least, the use of a truncated mortality dataset leads to a substantially different conclusion about the treatment effect of the college openings on longevity.

In Figure 5, I plot the estimated treatment effects (for towns within 5 miles of a community college opening) on college attainment and longevity by race and gender. The estimate on college attainment in the first row, for white women, is about one-tenth of a year and statistically significant at the five percent level. The point estimates for the three other demographics—white men, black men and black women—are all positive and between 0 and one-tenth of a year, but not statistically different from zero or from the effect for white women. In rows 5-8, the outcome is lifespan in years, and the sample is restricted to those

[41] Going from column 1 to column 2 in Figure 4 the sample size falls by a factor of five but the standard error actually *decreases*. In a separate paper (Connolly, 2024), I discuss the effect of truncation on mortality estimates, including how standard errors are affected.

for whom we can observe educational attainment. The point estimates for white men and white women are positive, with that of white men equal to about three tenths of a year and statistically different from zero. The point estimates for black women and men are negative but not statistically distinguishable from zero or from the point estimates for white men and women. In the last four rows, the outcome is again "Lifespan in Years," but this time I estimate effects on the full sample (i.e., without restriction to those for whom educational attainment is observed). The point estimates for black men, white men and white women are positive, with those of black and white men equal to slightly more than 0.5 years and statistically different from zero. Moreover, the effect for white men is statistically larger than the effect for white women. The point estimate for black women is negative but not statistically different from zero or from the estimate for white women.

In Table 5, I estimate the effect of the community college openings on college attainment and on longevity using the education subsample, and I also estimate the ordinary least squares correlation between college attainment and lifespan. The first column of Table 5 is identical to column 1 from Table 3. In the third column, I estimate a treatment effect of 0.3 additional years of lifespan on this education subsample which is 40% smaller than the estimate (0.5) on the full sample (column 1 of Table 4). In column 2, I use an alternative age grouping for the treatment and control groups—ten instead of fifteen years—so that the treatment group are those whose childhood town was within 5 miles of a two-year public college opening and age 15-24 (inclusive) at the time of the opening, while the control group is analogous, but age 25-35. The estimated treatment effect on education (column 2) is 5.9 hundredths of a year on average, while the estimated treatment effect on lifespan is 0.11 years. The implied instrumental variables (IV) effect of the community colleges on lifespan is therefore 1.9 years (with ten-year treatment and control groups) and 4.1 years (with fifteen-year treatment and control groups). In column five, I estimate the ordinary least squares relationship between college attainment and lifespan, and the estimated association is equal to 0.5 years of lifespan for each additional year of college attainment.

There are at least five possible explanations for this result, that the estimated instrumental variables effect of years of college on years of lifespan arising from the community college openings is considerably larger than the estimated nationwide ordinary least squares association. First, the estimate from the community college openings is for "compliers": those whose childhood towns previously had no college within 25 miles but which

were brought within five miles of the nearest college by a two-year public college opening, and who were induced to attain additional years of college by the college opening. It follows that many of the compliers are men and women who could not afford to go to college but for the opening of the local two-year public college. The compliers stand in contrast to the average college-going individual in the national sample, who has a family background that is considerably better off in terms of parents' socioeconomic status and less likely to live in a college "desert" as a child. One important difference between the typical complier and the typical college-going individual is the counterfactual occupation outcome if the person did not attend college. For the compliers, particularly male compliers, it is conceivable that the counterfactual occupation could be much more likely to lead to a relatively short lifespan than for the typical college-educated American male. For example, a complier in the control group may be much more likely to serve in the military, or to have a working-class occupation with an elevated risk of a relatively early death. I am not able to observe occupation or cause of death directly in the available data for the community college openings study population, so I cannot directly test these possibilities.

Second, it could be that the college openings created positive spillovers in terms of lifespan. The ordinary least squares estimate of college attainment on lifespan does not capture positive spillovers because it is estimated at the individual level. By contrast, the community college treatment varies at the level of clusters—those age 25 or younger at the time of the opening whose childhood towns were within five miles of a two-year public college opening—which means that the IV estimate of college attainment on longevity from the community college openings will capture positive spillovers within these treated clusters. If one's spouse, siblings and/or peers attain additional college, it is conceivable that they could exert an influence on one's own health behaviors that could lead to greater longevity. Lleras-Muney et al. (2020) find evidence that the association between educational attainment and longevity for men is particularly strong in U.S. states (and within birth cohorts in U.S. states) that have greater levels of educational attainment (Figures A-9 and A-10).

Third, some considerable part of the difference between the OLS estimate and the IV estimate could be a result of sampling variability. The lifespan estimates are somewhat imprecise, particularly on the education subsample (the point estimate in Table 5, Column 4 is 0.3, with a 95% confidence interval of [0.0664, 0.5638]), and one cannot rule out IV estimates of considerably less than one year of lifespan for each additional year of college.

The imprecise estimates are a result of the fact that the "Lifespan in Years" outcome variable is inherently noisy with respect to educational attainment. Precise estimates of the effect of college attainment on lifespan would require a much larger first-stage effect on educational attainment than is generally observed in related papers—that is, papers evaluating natural experiments to estimate the causal effects of educational attainment using college openings (e.g., Currie & Moretti, 2003) or compulsory schooling laws (Lleras-Muney, 2005, and others).

Fourth, it is possible that at least one of the assumptions concerning the distribution of the treatment effect on mortality does not hold. As noted in Section 3.3, in order for the causal estimator to be unbiased and consistent when the outcome is "Lifespan in Years," the projected lifespans must accurately capture the pattern of mortality for individuals who are still alive in 2021, and whether an individual is missing from the mortality dataset must be uncorrelated with the outcome "Lifespan in Years." If either or both of these assumptions do not hold, then the reduced-form treatment effect estimator, and hence the estimated IV effect—will be biased and inconsistent.

Fifth, it may be the case that the exclusion restriction does not hold. That is, the college openings may have increased lifespan for local college-age men and women in ways other than through increasing college attainment. For example, a local college opening may have led to local economic growth that disproportionately benefited those who were college-age or younger at the time of the opening, relative to those sharing the same childhood town who were older than college-age. If the exclusion restriction does not hold, then the estimated IV effect will be biased and inconsistent. Regardless of whether the exclusion restriction holds, however, the reduced-form effect (interpreted as the general effect of the college openings on lifespan), which has importance for public policy, would still be valid under the assumption of exchangeability.

I am able to rule out some violations of the exclusion restriction based on the geographic evidence that I obtain with the childhood-town research design. Consider the possibility that the typical community college opening stimulated local economic growth, and that this economic growth—not the observed increase in educational attainment—is in large part what caused longer lifespans for individuals in the treatment group (i.e., individuals who had a childhood town that was within five miles of the opening and who were college-age or younger at the time of the college opening). This scenario is possible only if the local economic growth had large effects on longevity for those whose childhood towns

were less than five miles from the college opening, but zero effect for those whose childhood towns were 5-20 miles from the college opening. If the typical college opening had a substantial positive effect on local economic growth, it is reasonable to think either that the area of economic growth typically extended beyond five miles, or that individuals whose childhood towns were 5-20 miles from the college typically benefited to some extent from the economic growth, by commuting or moving to the area near the college which was experiencing economic growth. Instead, I find that the college openings had zero effect on either educational attainment or lifespan for individuals whose childhood towns were 5-20 miles from the college opening. More generally, if there is a violation of the exclusion restriction, the mechanism must operate in a way that fits the observed geographic pattern of estimated effects on educational attainment and lifespan.

Alternatively, the estimated effect of college attainment on longevity arising from the community college openings may in part be the result of some other unobserved factor that is varying between the treatment and control groups—that is, some violation of the key identification assumption of (conditional mean) exchangeability. In this case, as in the case of any natural experiment, there is some irreducible uncertainty about this assumption. As the locations and dates of the college openings were the result not of random assignment but of political processes at the state and local level, there is no a priori guarantee of exchangeability between the treatment and control groups arising from these openings. The research design enables me to rule out exchangeability violations arising from trends at the county level or higher, as well as from any fixed characteristics of one's childhood town. I have also obtained evidence that the exchangeability assumption seems reasonable using covariate balance tables and event study regressions. However, if there is some other trend orthogonal to observed family background variables that is varying at the level of towns along with the openings of the community colleges, it may explain at least part of the estimated effect of the openings on longevity.

A number of recent papers have used natural experiments to estimate causal effects of educational attainment on longevity. One set of papers concerns the effect of additional high school attainment, using changes in compulsory schooling laws as a source of quasi-experimental variation.[42] The magnitudes of the effects estimated in these papers range from 0 to a 6.9 percentage-point reduction in mortality over a ten-year period. One published

[42]See Galama et al. (2018) for a review of these papers.

paper estimates the causal effect of increased college attainment on mortality, using state-level variation in one's risk of being drafted for the Vietnam War (Buckles et al., 2016). The estimated effect of an additional year of college attainment in this case is a 2.6 percentage point reduction in mortality over a 26-year period (starting in one's 30s). Put in comparable terms, my estimate of the effect of one year of additional college attainment is a 6.6 percentage point reduction in mortality between ages 55 and 65. This estimated effect is larger than the estimate from Vietnam draft risk and at the higher end of the range of estimates for compulsory schooling laws.[43]

5 Conclusion

In this paper I examine the effects of the great expansion of two-year public college in the US between 1945 and 1980. These community college openings led to about one-tenth of a year of additional college attainment for local men and women who were college-age or younger at the time of the opening, to "democratization" and not "diversion" (i.e., a higher probability of completing a four-year college degree), and to between 0.1 and 0.5 years of additional lifespan.

The results of this paper rest on the assumption of exchangeability (i.e., the assumption that the treatment and control groups are substantially the same in every respect relevant to the outcomes). As with any natural experiment, there is some irreducible uncertainty about this assumption. I am able to rule out exchangeability violations arising from trends at the county level or higher, as well as from any fixed characteristics of one's childhood town. I have also obtained evidence that the exchangeability assumption seems reasonable, using covariate balance tables and event study regressions.

The college openings may have affected a number of important life outcomes that

[43]These three types of natural experiments—compulsory schooling law changes, variation in the risk of being drafted for the Vietnam War, and community college openings—involve both different types of education (high school, general college, and community college) and different types of compliers. The complier for compulsory-schooling-law natural experiments is the high school student who would drop out at the earliest point permitted by law and only attains additional schooling because it is legally compulsory. In the case of Vietnam draft risk, the complier is a person who is attaining additional years of college solely because he wants to reduce his risk of being drafted for the Vietnam War. By contrast, in the case of local college openings, the complier is a person who is attaining additional years of college because it has been made more affordable. As a result of these differences——the different types of education and the different compliers—it is reasonable to think that the magnitudes of the estimands (the effect of increased educational attainment on mortality that the researcher is seeking to estimate) and hence of the estimates will differ.

are beyond the scope of this paper. Some relate to education, such as the institution attended and one's area of study. There are also longer-term outcomes, including income, political participation (e.g., propensity to vote), outcomes relating to culture (e.g., religiosity), criminal justice, geographic mobility, as well as inter-generational college attainment and longevity. All of these outcomes are important topics for future research.

References

Aaronson, D., Mazumder, B., Sanders, S. G., & Taylor, E. J. (2021). Estimating the Effect of School Quality on Mortality in the Presence of Migration: Evidence from the Jim Crow South. *Journal of Labor Economics*, *39*(2), 527–558.

Acemoglu, D., & Johnson, S. (2007). Disease and Development: the Effect of Life Expectancy on Economic Growth. *Journal of Political Economy*, *115*(6), 925–985.

Aizer, A., Eli, S., Ferrie, J., & Lleras-Muney, A. (2016). The Long-Run Impact of Cash Transfers to Poor Families. *American Economic Review*, *106*(4), 935–971.

Albouy, V., & Lequien, L. (2009). Does Compulsory Education Lower Mortality? *Journal of Health Economics*, *28*(1), 155–168.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Ashenfelter, O., & Krueger, A. (1994). Estimates of the Economic Return to Schooling from a New Sample of Twins. *The American Economic Review*, *84*(5), 1157–1173.

Atherwood, S. (2022). Does a Prolonged Hardship Reduce Life Span? Examining the Longevity of Young Men who Lived Through the 1930s Great Plains Drought. *Population and Environment*, *43*(4), 530–552.

Becker, G. S. (2009). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Bedard, K., & Deschênes, O. (2006). The Long-Term Impact of Military Service on Health: Evidence from World War II and Korean War Veterans. *American Economic Review*, *96*(1), 176–194.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, *75*(4), 352–365.

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.

Black, D. A., Sanders, S. G., Taylor, E. J., & Taylor, L. J. (2015). The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South. *American Economic Review*, *105*(2), 477–503.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review*, 531–553.

Bound, J., & Solon, G. (1999). Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling. *Economics of Education Review*, *18*(2), 169–182.

Breen, C. F. (2024). The Longevity Benefits of Homeownership: Evidence from Early 20th-Century US Male Birth Cohorts.

Brint, S., & Karabel, J. (1989). *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900-1985*. Oxford University Press.

Buckles, K., Hagemann, A., Malamud, O., Morrill, M., Wozniak, A., et al. (2016). The Effect of College Education on Mortality. *Journal of Health Economics*, *50*(100), 99–114.

Caplan, B. (2019). *The Case Against Education*. Princeton University Press.

Card, D. (1993). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. Working Paper 4483. *National Bureau of Economic Research*.

Carneiro, P., & Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, *112*(482), 705–734.

Case, A., & Deaton, A. (2023). Without a College Degree, Life in America Is Staggeringly Shorter [Available at: https://www.nytimes.com/2023/10/03/opinion/life-expectancy-college-degree.html (Accessed: June 27th, 2024)]. *New York Times*.

Cervellati, M., & Sunde, U. (2005). Human Capital Formation, Life Expectancy, and the Process of Development. *American Economic Review*, *95*(5), 1653–1672.

Cervellati, M., & Sunde, U. (2013). Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited. *Econometrica*, *81*(5), 2055–2086.

Cervellati, M., & Sunde, U. (2015). The Economic and Demographic Transition, Mortality, and Comparative Development. *American Economic Journal: Macroeconomics*, *7*(3), 189–225.

Clark, B. R. (1960). The "Cooling-Out" Function in Higher Education. *American Journal of Sociology*, *65*(6), 569–576.

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

Cohen, A. M., Brawer, F. B., & Kisker, C. B. (2014). *The American Community College*. Jossey-Bass.

Connolly, K. (2024). Estimating Mortality Effects with Truncation and Censoring [*Unpublished Manuscript*].

Cowan, B. W., & Tefft, N. (2024). College Access and Adult Health. *American Journal of Health Economics*.

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*, 1495–1532.

Cutler, D. M., & Lleras-Muney, A. (2006). Education and Health: Evaluating Theories and Evidence. *NBER Working Paper*, (w12352).

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding Differences in Health Behaviors by Education. *Journal of Health Economics*, *29*(1), 1–28.

Davies, N. M., Dickson, M., Davey Smith, G., Van Den Berg, G. J., & Windmeijer, F. (2018). The Causal Effects of Education on Health Outcomes in the UK Biobank. *Nature Human Behaviour*, *2*(2), 117–125.

Davies, N. M., Dickson, M., Davey Smith, G., Windmeijer, F., & Van Den Berg, G. J. (2023). The Causal Effects of Education on Adult Health, Mortality and Income: Evidence from Mendelian Randomization and the Raising of the School Leaving Age. *International Journal of Epidemiology*, *52*(6), 1878–1886.

De Walque, D. (2010). Education, Information, and Smoking Decisions: Evidence from Smoking Histories in the United States, 1940–2000. *Journal of Human Resources*, *45*(3), 682–717.

Doyle, W. R. (2009). The Effect of Community College Enrollment on Bachelor's Degree Completion. *Economics of Education Review*, *28*(2), 199–206.

Fenoll, A. A., & Kuehn, Z. (2017). Compulsory Schooling Laws and Migration Across European Countries. *Demography*, *54*(6), 2181–2200.

Finlay, K., & Genadek, K. R. (2021). Measuring All-Cause Mortality with the Census Numident File. *American Journal of Public Health*, *111*(S2), S141–S148.

Fischer, M., Karlsson, M., & Nilsson, T. (2013). Effects of Compulsory Schooling on Mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health*, *10*(8), 3596–3618.

Fletcher, J., & Noghanibehambari, H. (2023). Toxified to the Bone: Early-Life and Childhood Exposure to Lead and Men's Old-Age Mortality. Working Paper 31957. *National Bureau of Economic Research*.

Fletcher, J., & Noghanibehambari, H. (2024a). The Effects of Education on Mortality: Evidence Using College Expansions. *Health Economics*, *33*(3), 541–575.

Fletcher, J., & Noghanibehambari, H. (2024b). The Siren Song of Cicadas: Early-Life Pesticide Exposure and Later-Life Male Mortality. *Journal of Environmental Economics and Management*, *123*, 102903.

Frisvold, D., & Golberstein, E. (2011). School Quality and the Education–Health Relationship: Evidence from Blacks in Segregated Schools. *Journal of Health Economics*, *30*(6), 1232–1245.

Frisvold, D., & Golberstein, E. (2013). The Effect of School Quality on Black-White Health Differences: Evidence from Segregated Southern Schools. *Demography*, *50*(6), 1989–2012.

Fuchs, V. R. (1982). Time Preference and Health: An Exploratory Study. In *Economic Aspects of Health* (pp. 93–120). University of Chicago Press.

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. *Oxford Research Encyclopedia of Economics and Finance*.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

Galor, O., & Weil, D. N. (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, *90*(4), 806–828.

Gathmann, C., Jürges, H., & Reinhold, S. (2015). Compulsory Schooling Reforms, Education and Mortality in Twentieth Century Europe. *Social Science & Medicine*, *127*, 74–82.

Gerber, A. S., & Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton & Company.

Glied, S., & Lleras-Muney, A. (2008). Technological Innovation and Inequality in Health. *Demography*, *45*(3), 741–761.

Goldin, C., & Katz, L. F. (2010). *The Race Between Education and Technology*. Harvard University Press.

Goldstein, J. R., Alexander, M., Breen, C., González, A. M., Menares, F., Osborne, M., Snyder, M., & Yildirim, U. (2023). CenSoc Mortality File: Version 3.0. https://censoc.berkeley.edu/data/

González, F., Martínez, L. R., Muñoz, P., & Prem, M. (2023). Higher Education and Mortality: Legacies of an Authoritarian College Contraction. *Journal of the European Economic Association*.

Grawe, N. D. (2018). *Demographics and the Demand for Higher Education*. JHU Press.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, *80*(2), 223–255.

Grossman, M. (2000). The Human Capital Model. In *Handbook of Health Economics* (pp. 347–408, Vol. 1). Elsevier.

Grossman, M. (2006). Education and Nonmarket Outcomes. *Handbook of the Economics of Education*, *1*, 577–633.

Grossman, M. (2015). The Relationship between Health and Schooling: What's New? *Nordic Journal of Health Economics*, *3*(1), 7–17.

Grossman, M. (2017). *The Demand for Health: a Theoretical and Empirical Investigation*. Columbia University Press.

Halpern-Manners, A., Helgertz, J., Warren, J. R., & Roberts, E. (2020). The Effects of Education on Mortality: Evidence from Linked US Census and Administrative Mortality Data. *Demography*, *57*(4), 1513–1541.

Harnisch, T., Burns, R., Heckert, K., Kunkel, K., & Weeden, D. (2024). State Priorities for Higher Education in 2023: Survey of State Higher Education Officers. *State Higher Education Executive Officers Association*.

Hazan, M. (2009). Longevity and Lifetime Labor Supply: Evidence and Implications. *Econometrica*, *77*(6), 1829–1863.

Helgertz, J., & Warren, J. R. (2023). Early Life Exposure to Cigarette Smoking and Adult and Old-Age Male Mortality. *Demographic Research*, *49*, 651–692.

Hill, M. E., & Rosenwaike, I. (2001). The Social Security Administration's Death Master File: the Completeness of Death Reporting at Older Ages. *Soc. Sec. Bull.*, *64*, 45.

Hollingsworth, A., Karbownik, K., Thomasson, M. A., & Wray, A. (2024). The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality. *American Economic Review*, *114*(7), 2201–2238.

Jacks, D. S., Pendakur, K., Shigeoka, H., & Wray, A. (2024). Later-life Mortality and the Repeal of Federal Prohibition [*Forthcoming*]. *Journal of Public Economics*.

Johnson, J. E., & Taylor, E. J. [Evan J]. (2019). The Long Run Health Consequences of Rural-Urban Migration. *Quantitative Economics*, *10*(2), 565–606.

Karger, E. (2021). The Long-Run Effect of Public Libraries on Children: Evidence from the Early 1900s. *SocArXiv. Doi*, *10*.

Kling, J. R. (2001). Interpreting Instrumental Variables Estimates of the Returns to Schooling. *Journal of Business & Economic Statistics*, *19*(3), 358–364.

Kocks, G. (2023). School Desegregation and Long-Run Health.

Kolata, G. (2007). A Surprising Secret to a Long Life: Stay in School [Available at: https://www.nytimes.com/2007/01/03/health/03aging.html (Accessed: June 27th, 2024)]. *New York Times*.

Lacroix, G., Laliberté-Auger, F., Michaud, P.-C., & Parent, D. (2021). The Effect of College Education on Health and Mortality: Evidence from Canada. *Health Economics*, *30*, 105–118.

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.

Leigh, D. E., & Gill, A. M. (2003). Do Community Colleges Really Divert Students from Earning Bachelor's Degrees? *Economics of Education Review*, *22*(1), 23–30.

Lillard, D. R., & Molloy, E. (2010). Live and Learn or Learn and Live: Does Education Lead to Longer Lives. *Available at SSRN 1880767*.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, *72*(250), 189.

Lleras-Muney, A., Price, J., & Yue, D. (2020). The Association between Educational Attainment and Longevity Using Individual Level Data from the 1940 Census. Working Paper 27514. *National Bureau of Economic Research*.

Malamud, O., Mitrut, A., & Pop-Eleches, C. (2023). The Effect of Education on Mortality and Health: Evidence from a Schooling Expansion in Romania. *Journal of Human Resources*, *58*(2), 561–592.

Manski, C. F. (1989). Anatomy of the Selection Problem. *Journal of Human Resources*, 343–360.

Massey, C. G., Genadek, K. R., Alexander, J. T., Gardner, T. K., & O'Hara, A. (2018). Linking the 1940 US Census with Modern Data. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *51*(4), 246–257.

Masuda, K., & Shigeoka, H. (2023). Education and Later-Life Mortality: Evidence from a School Reform in Japan. *NBER Working Paper*, (w31472).

Mazumder, B. (2008). Does Education Improve Health? A Reexamination of the Evidence from Compulsory Schooling Laws. *Economic Perspectives*, *32*(2).

Medsker, L. L. (1960). The Junior College: Prospect and Progress. *New York: McGraw-Hill Book Co.*

Medsker, L. L., & Tillery, D. (1971). Breaking the Access Barriers: A Profile of Two-Year Colleges.

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

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How Conditioning on Posttreatment Variables Can Ruin your Experiment and What To Do About It. *American Journal of Political Science*, *62*(3), 760–775.

Moorthy, A., Figinski, T. F., & Lloro, A. (2022). Revisiting the Effect of Education on Later Life Health. *FEDS Working Paper*.

Mountjoy, J. (2022). Community Colleges and Upward Mobility. *American Economic Review*, *112*(8), 2580–2630.

Murphy, K. M., & Topel, R. H. (2006). The Value of Health and Longevity. *Journal of Political Economy*, *114*(5), 871–904.

National Center for Health Statistics (US) and National Center for Health Services Research. (2011). *Health, United States*. US Department of Health, Education, Welfare, Public Health Service.

Noghanibehambari, H. (2022). In Utero Exposure to Natural Disasters and Later-Life Mortality: Evidence from Earthquakes in the Early Twentieth Century. *Social Science & Medicine*, *307*, 115189.

Noghanibehambari, H., & Engelman, M. (2022). Social Insurance Programs and Later-Life Mortality: Evidence from New Deal Relief Spending. *Journal of Health Economics*, *86*, 102690.

Noghanibehambari, H., & Fletcher, J. (2023a). Childhood Exposure to Birth Registration Laws and Old-Age Mortality. *Health Economics*, *32*(3), 735–743.

Noghanibehambari, H., & Fletcher, J. (2023b). Early-Life Income Shocks and Old-Age Mortality: Evidence from World War I Veterans' Bonus. *Available at SSRN 4499916*.

Noghanibehambari, H., & Fletcher, J. (2023c). In Utero and Childhood Exposure to Alcohol and Old Age Mortality: Evidence from the Temperance Movement in the US. *Economics & Human Biology*, *50*, 101276.

Noghanibehambari, H., & Fletcher, J. (2023d). Long-Term Health Benefits of Occupational Licensing: Evidence from Midwifery Laws. *Journal of Health Economics*, *92*, 102807.

Noghanibehambari, H., & Fletcher, J. (2023e). The Early Bird Catches the Worm: The Effect of Birth Order on Old-Age Mortality. *Population and Development Review*, *49*(3), 531–560.

Noghanibehambari, H., & Fletcher, J. (2024a). Dust to Feed, Dust to Gray: The Effect of in Utero Exposure to the Dust Bowl on Old-Age Longevity. *Demography*, *61*(1), 87–113.

Noghanibehambari, H., & Fletcher, J. (2024b). Unequal Before Death: The Effect of Paternal Education on Children's Old-Age Mortality in the United States. *Population Studies*, 1–27.

Noghanibehambari, H., Fletcher, J., Schmitz, L., Duque, V., & Gawai, V. (2024). Early-Life Economic Conditions and Old-Age Male Mortality: Evidence from Historical County-Level Bank Deposit Data. *Journal of Population Economics*, *37*(1), 1–33.

Noghanibehambari, H., & Noghani, F. (2023). Long-Run Intergenerational Health Benefits of Women Empowerment: Evidence from Suffrage Movements in the US. *Health Economics*, *32*(11), 2583–2631.

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, *25*(1), 159–184.

Preston, S. H. (1977). Mortality Trends. *Annual Review of Sociology*, *3*, 163–178.

Pronk, N., Kleinman, D. V., Goekler, S. F., Ochiai, E., Blakey, C., & Brewer, K. H. (2021). Promoting Health and Well-Being in Healthy People 2030. *Journal of Public Health Management and Practice*, *27*(Supplement 6), S242–S248.

Reynolds, C. L. (2012). Where to Attend? Estimating the Effects of Beginning College at a Two-Year Institution. *Economics of Education Review*, *31*(4), 345–362.

Ross, C. E., & Mirowsky, J. (1999). Refining the Association Between Education and Health: The Effects of Quantity, Credential, and Selectivity. *Demography*, *36*(4), 445–460.

Rostron, B., Boies, J., & Arias, E. (2010). Education Reporting and Classification on Death Certificates in the United States. *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research*, (151), 1–21.

Rouse, C. E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment. *Journal of Business & Economic Statistics*, *13*(2), 217–224.

Saavedra, M. H. (2013). Early Childhood Conditions and Mortality: Evidence from Japanese American Internment. *SSRN Working Paper 1932111*.

Salm, M. (2011). The Effect of Pensions on Longevity: Evidence from Union Army Veterans. *The Economic Journal*, *121*(552), 595–619.

Sanchez, Y. (2010). The Longevity Gains of Education. *Manuscript, Univ. Chicago, Dept. Econ.*

Sansani, S. (2011). The Effects of School Quality on Long-Term Health. *Economics of Education Review*, *30*(6), 1320–1333.

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.

Seltzer, N. (2021). The Intergenerational Consequences of Climate Change on Later Life Mortality: Evidence from Individual-Level US Death Records.

Soares, R. R. (2005). Mortality Reductions, Educational Attainment, and Fertility Choice. *American Economic Review*, *95*(3), 580–601.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Stock, J., & Watson, M. (2019). *Introduction to Econometrics, 4th Edition*. London: Pearson Education.

Taylor, E. J. [Evan J.], Stuart, B. A., & Bailey, M. J. (2016). Summary of Procedure to Match NUMIDENT Place of Birth County to GNIS Places. *Center for Economic Studies, US Census Bureau*.

Taylor, S. (2020). Growing Up in a Steel Town: Early-Life Pollution Exposure and Later-Life Mortality. *Social Science History*, *44*(1), 117–142.

U. S. Office of the Surgeon General. (1979). *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention* (Vol. 79). US Department of Health, Education, Welfare, Public Health Service.

United States Bureau of the Census. (1975). *Historical Statistics of the United States, Times to 1970*. US Department of Commerce, Bureau of the Census.

U.S. Department of Education, National Center for Education Statistics. (1996, 2009). Integrated Postsecondary Education Data System (IPEDS).

Van Kippersluis, H., O'Donnell, O., & Van Doorslaer, E. (2011). Long-Run Returns to Education: Does Schooling Lead to an Extended Old Age? *Journal of Human Resources*, *46*(4), 695–721.

Vu, H., Noghanibehambari, H., Fletcher, J., & Green, T. (2023). Prenatal Exposure to Racial Violence and Later Life Mortality among Males: Evidence from Lynching. *Available at SSRN 4613372*.

Zook, G. F. (1947). *Higher Education for American Democracy: A Report of the President's Commission on Higher Education* (Vol. 1). US Government Printing Office.

Figure 1

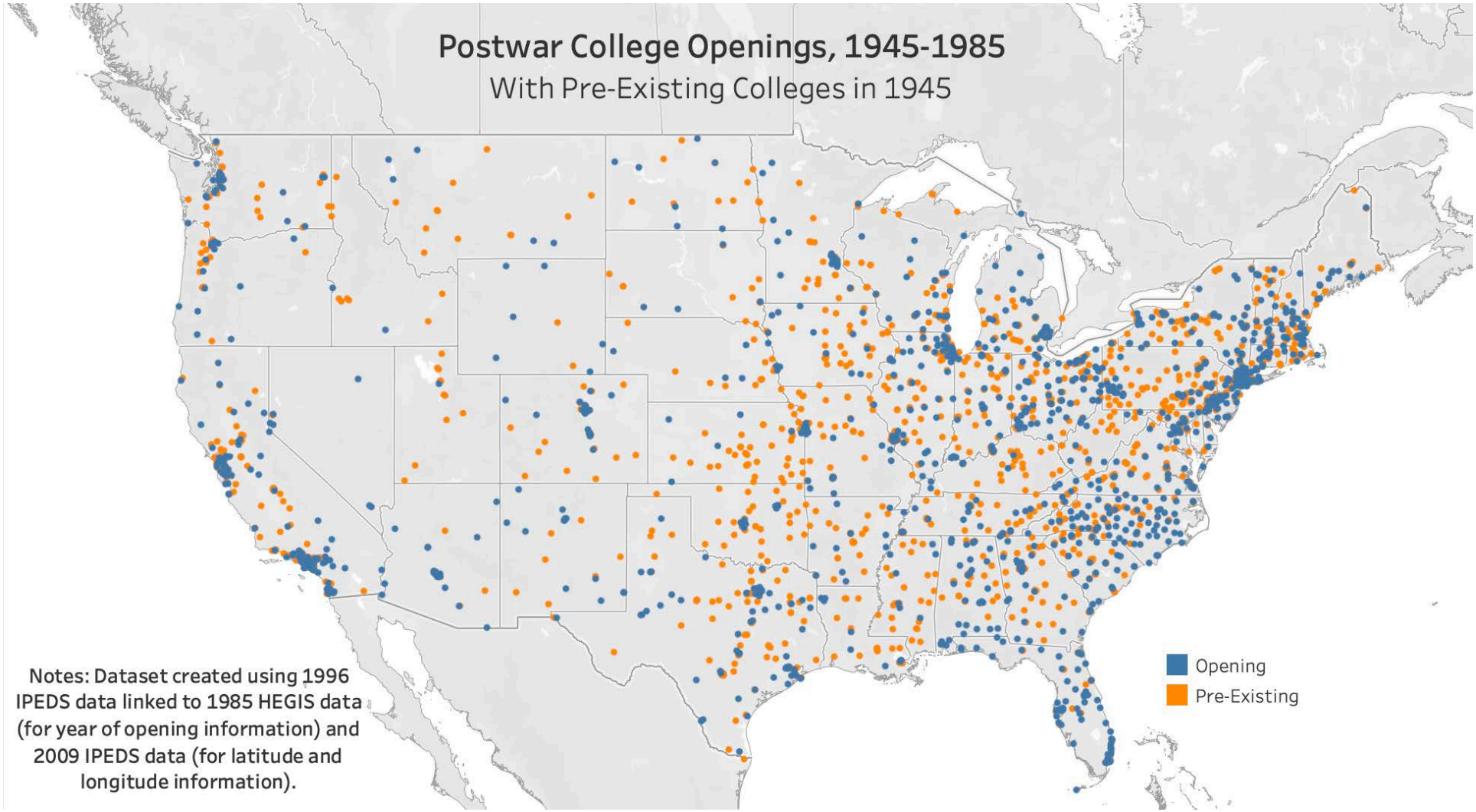
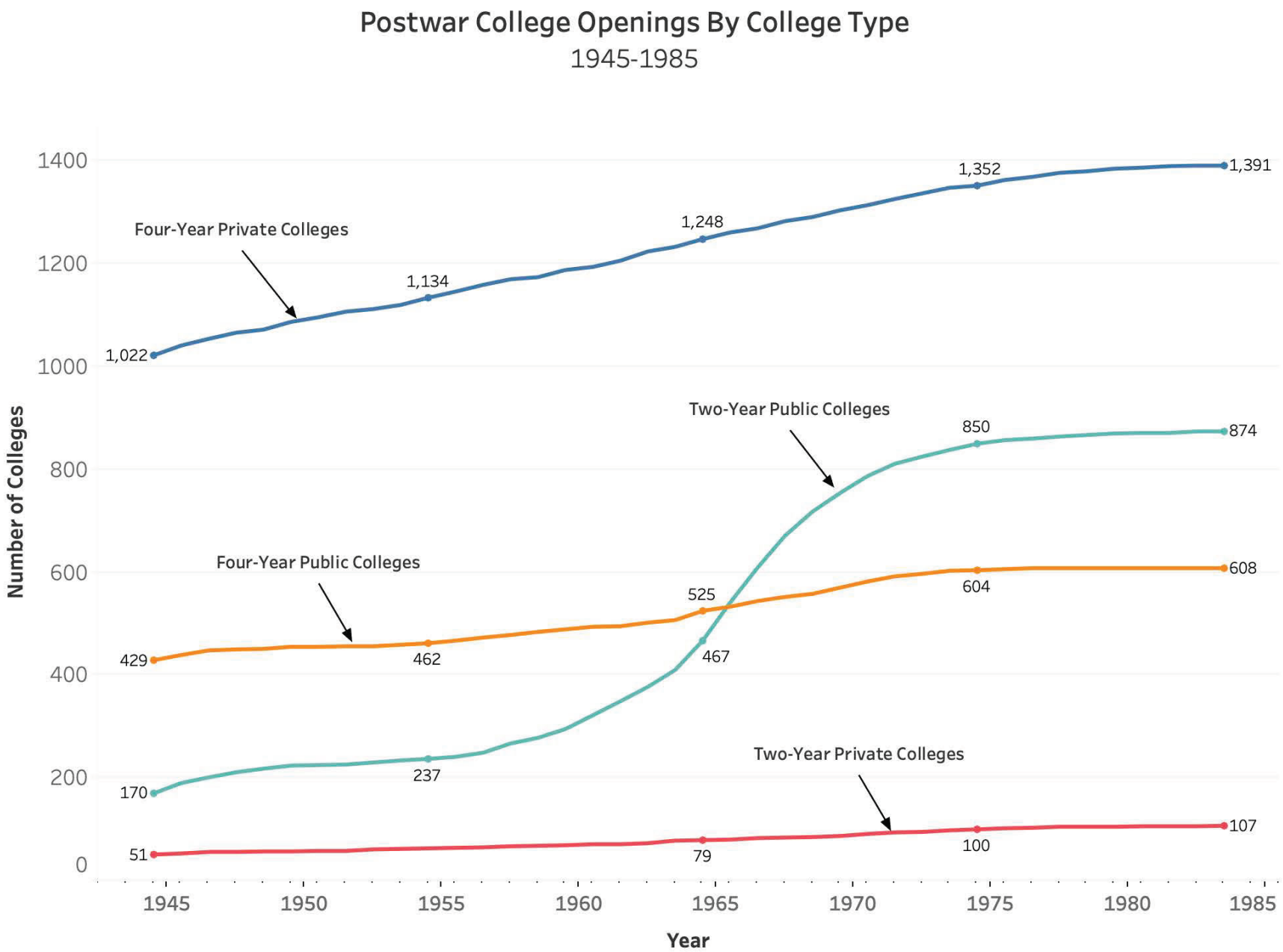
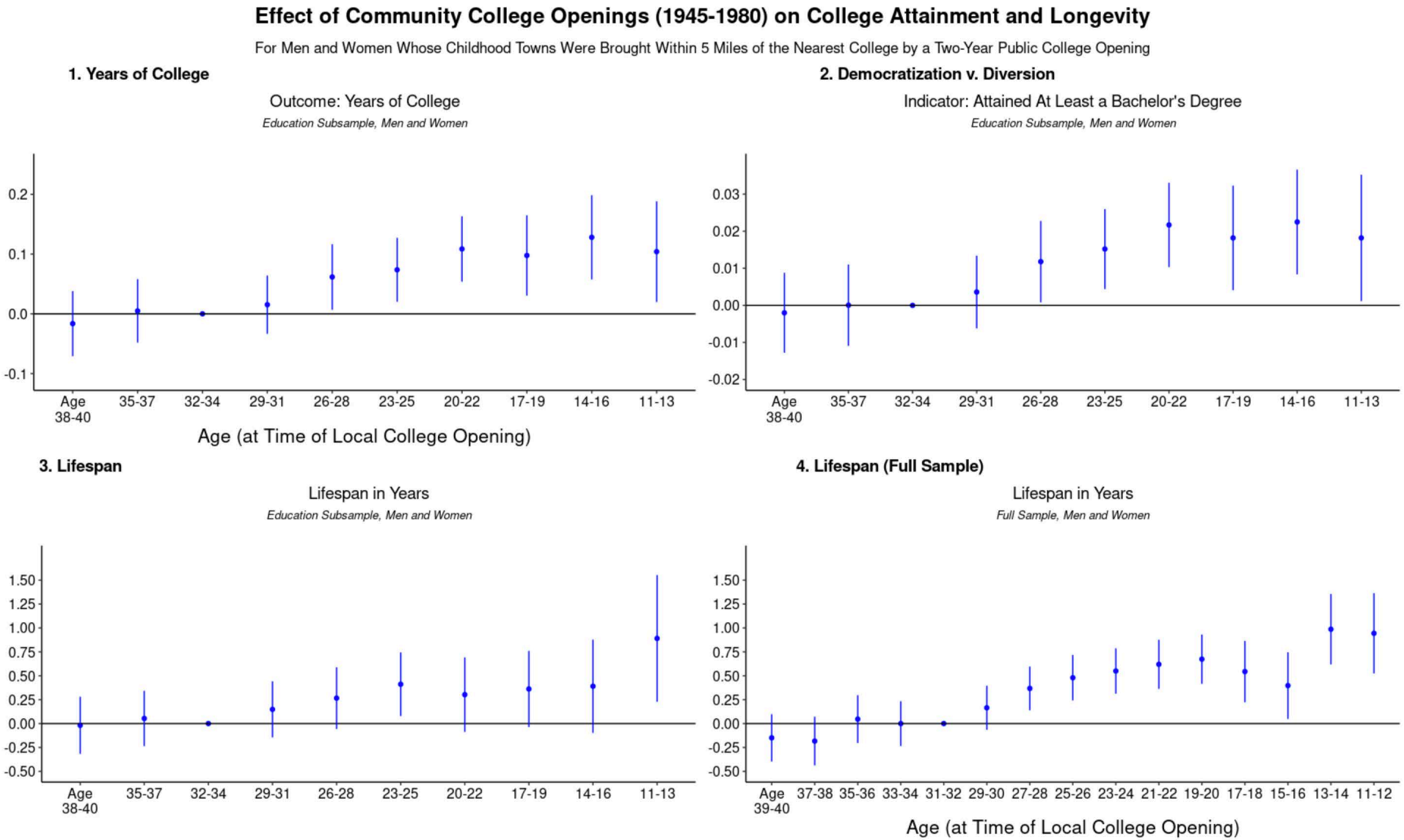

Postwar College Openings, 1945-1985
With Pre-Existing Colleges in 1945
Notes: Dataset created using 1996 IPEDS data linked to 1985 HEGIS data (for year of opening information) and 2009 IPEDS data (for latitude and longitude information).
Opening
Pre-Existing

Figure 2



Notes: Dataset created using 1996 IPEDS data linked to 1985 HEGIS data (for year of opening information) and 2009 IPEDS data (for latitude and longitude information).

Figure 3

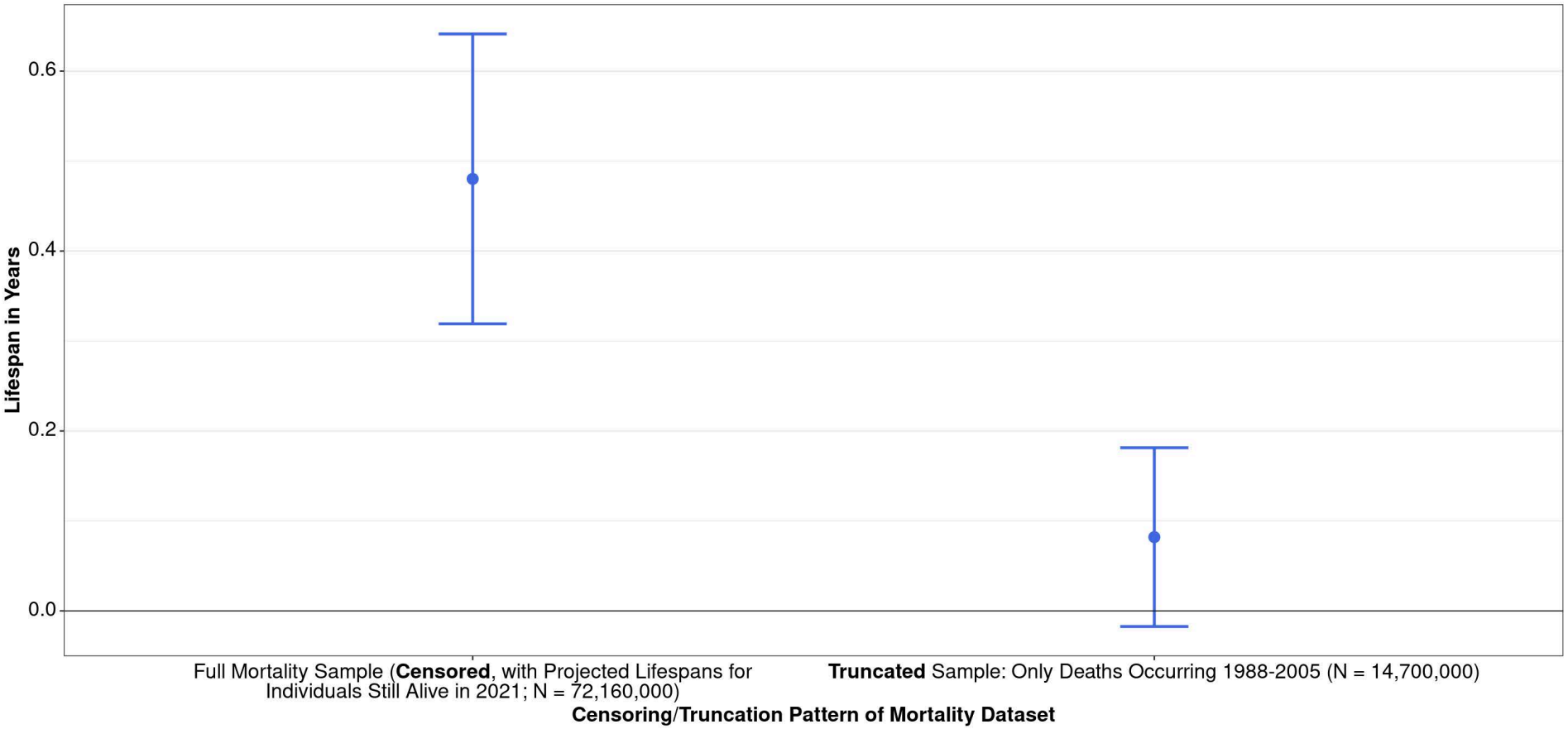


Notes: All specifications include childhood (i.e., 1940 or Numident) town fixed effects and childhood-county-by-birth-year fixed effects, with standard errors clustered at the level of childhood town. Coefficients are relative to the coefficient for age 32-34 (31-32 for full sample). Error bars are 95% confidence intervals.

Figure 4

The Estimated Effect of College Openings on Lifespan under Different Patterns of Truncation/Censoring

The observed effect of the college openings changes when one changes the mortality sample from the censored sample (Table 4, Column 1) to a truncated sample in which one observes only individuals who died between 1988 and 2005, suggesting that the latter estimate is misleading with respect to the estimand of interest.



Notes: This figure plots the point estimates and 95% confidence intervals. The first truncation/censoring pattern is the one used in Table 4, column 1, and the second corresponds to mortality datasets based on a public version of the Social Security administration death data known as Numident, including CenSoc Numident (which has been used in Fletcher & Noghanharambari (2024) to estimate the effect of U.S. college openings on lifespan) as well as NARA Numident (Hollingsworth et al., (2024)).

Figure 5

Effect of Community College Openings (1945-1980) on College Attainment and Longevity

By Gender and Race

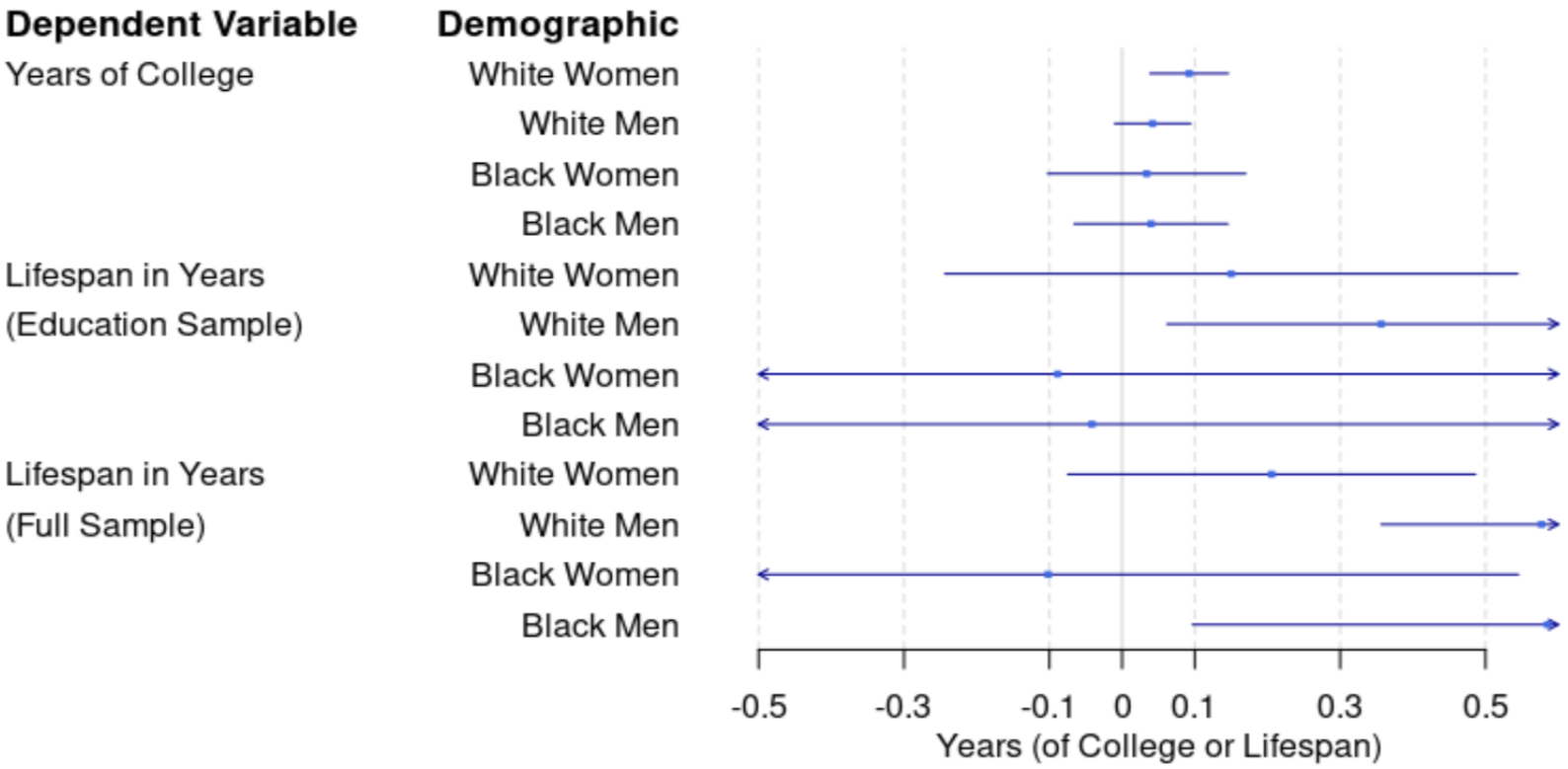


Table 1

Summary Statistics

Variable	Full Sample Mean	Full Sample Std. Dev.	Full Sample Share Nonmissing	Event Study Sample Mean	Event Study Sample Std. Dev.	Event Study Sample Share Nonmissing
Year of Birth	1937	9.4	1	1939	9.2	1
Age at Social Security Enrollment	18.32	6.8	0.98			
Race and Gender						
Male	0.51	0.5	1	0.50	0.5	1
White	0.84	0.37	0.98	0.81	0.39	0.99
Black	0.12	0.33	0.98	0.13	0.34	0.99
Other	0.03	0.16	0.98	0.05	0.21	0.99
Unknown	0.01	0.11	0.98	0.01	0.11	0.99
Location Information						
Location Observable at Birth but not in 1940	0.68	0.47	1	0.79	0.41	1
Location Observable at Birth and in 1940	0.30	0.46	1	0.20	0.40	1
Location Observable in 1940 but not at Birth	0.02	0.15	1	0.02	0.12	1
Age At Which Location is Observed	3.22	5.65	1	2.12	4.81	1
Educational Attainment						
Mother's Years of College	0.18	0.73	0.33	0.18	0.71	0.24
Father's Years of College	0.27	0.96	0.31	0.23	0.88	0.22
Years of College	1.16	1.89	0.23	1.16	1.84	0.24
Attained At Least An Associate's Degree	0.23	0.42	0.23	0.24	0.43	0.24
Attained At Least A Bachelor's Degree	0.19	0.39	0.23	0.19	0.39	0.24
Longevity						
Lived To At Least Age 55	0.84	0.37	1	0.83	0.37	1
Lived To At Least Age 65	0.75	0.43	1	0.74	0.44	1
Lifespan (Projected)	74.0	18.8	1	73.6	19.1	1
Lifespan Left-Censored	0.11	0.31	1	0.11	0.31	1
Lifespan Right-Censored	0.38	0.48	1	0.41	0.49	1
N	80,950,000			1,003,000		

Table 2

Community College Openings: Covariate Balance Table

	Parents' College Attainment				*Controlling for Parents' College Attainment*	
	Mother's Years of College	Father's Years of College	Years of College	Years of Lifespan	Years of College	Years of Lifespan
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a	-0.0034	0.0134	0.0702	0.044	0.0671	0.0422
Two-Year Public College Opening and Age 10-25 at Time of Opening	(0.0187)	(0.0197)	(0.0372)	(0.2156)	(0.0359)	(0.2147)
Indicator: ...5 to 20 Miles…	-0.0009	0.0245	0.0244	0.0136	0.0173	0.0092
	(0.0127)	(0.0151)	(0.0267)	(0.1605)	(0.0259)	(0.1605)
Mother's Years of College					0.2782***	0.1951***
					(0.0026)	(0.0068)
Father's Years of College					0.3011***	0.1845***
					(0.0025)	(0.0052)
Childhood Town of Residence FEs	Yes	Yes	Yes	Yes	Yes	Yes
Childhood County by Birth Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	67,000	67,000	67,000	67,000	67,000	67,000
Dep. Var. Mean	0.201	0.29	0.977	85.6	0.977	85.6
Observations	5,287,000	5,287,000	5,287,000	5,287,000	5,287,000	5,287,000

Notes: The underlying dataset consists of the subset of Numident born 1921-1940 that can be linked to the 1940 Census, for which father's and mother's educational attainment is observed in the 1940 census, for which own educational attainment can be observed in the 2000 census or the 2001-2015 American Community Survey, and whose childhood town is geocodable from the 1940 Census or Numident. Standard errors are clustered at the level of childhood town. The omitted explanatory variable for each distance category (0 to 5 miles, and 5 to 20 miles) consists of those who were age 26 to 40 at the time of the college opening. The college openings in this analysis are those that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles.

Table 3

Effect of Two-Year Public College Openings on College Attainment

	Indicator: Attained at Least a(n)…		
	Years of College	Associate's Degree	Bachelor's Degree
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 10-25 at Time of Opening	0.0739***	0.0173***	0.0141***
	(0.0206)	(0.0046)	(0.0042)
Indicator: ... 5-20 Miles…Age 10-25…	-0.0091	-0.0008	-0.002
	(0.0114)	(0.0026)	(0.0023)
Childhood Town FEs	Yes	Yes	Yes
Childhood County By Birth Year FEs	Yes	Yes	Yes
Dep. Var. Mean	1.16	0.232	0.19
Observations	18,300,000	18,300,000	18,300,000

Notes: Standard errors are clustered at the level of childhood town. The community colleges study population consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence) and which can be linked to either the 2000 Census long form or the 2001-2015 American Community Survey. The omitted explanatory variable for each distance category (0 to 5 miles, and 5 to 20 miles) consists of those who were age 26 to 40 at the time of the college opening. The college openings in this analysis are two-year public college openings that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles.

Table 4

Effect of Community College Openings on Longevity

Birth Cohorts	**1921-1951**						**1932-1951**
Truncation/Censoring Pattern	*Left-Truncated, Right Censored with Projections*	*Left-Truncated, Right-Censored*				*Left-Censored, Right-Censored*	*Left-Censored, Right-Censored*
		Indicator: Lived to at Least Age…					
	Lifespan in Years	...55	…60	…65	…68	…68	…68
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 10-25 at Time of Opening	0.4802***	0.0052***	0.0077***	0.0100***	0.0132***	0.0275***	0.0201***
	(0.0822)	(0.0016)	(0.0018)	(0.0020)	(0.0022)	(0.004)	. (0.0036)
Indicator: ... 5-20 Miles…	-0.0389	-0.0009	-0.0011	-0.0014	-0.0016	-	-
	(0.0556)	(0.0010)	(0.0012)	(0.0015)	(0.0016)		
Childhood Town FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Childhood County By Birth Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	78.800	0.93600	0.89600	0.83800	0.79400	*Supp.*	*Supp.*
Observations	72,160,000	72,160,000	72,160,000	72,160,000	72,160,000	80,700,000	80,700,000
Number of Clusters	135,000	135,000	135,000	135,000	135,000	*Supp.*	*Supp.*

Notes: The underlying dataset consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence) and which has a date of death in Numident or the date of death is right censored in 2020. Standard errors are clustered at the level of childhood town. The omitted explanatory variable for each distance category (0 to 5 miles, and 5 to 20 miles) consists of those who were age 26 to 40 at the time of the college opening. SEAs are the state economic areas, which are groups of counties within states. The college openings in this analysis are those that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles. "-" indicates that the sign of the coefficient is negative and the magnitude is not statistically significant; "*Supp* ." indicates the figure has not yet been released from the Census Bureau.

Table 5

Effect of Community College Openings on College Attainment and Longevity

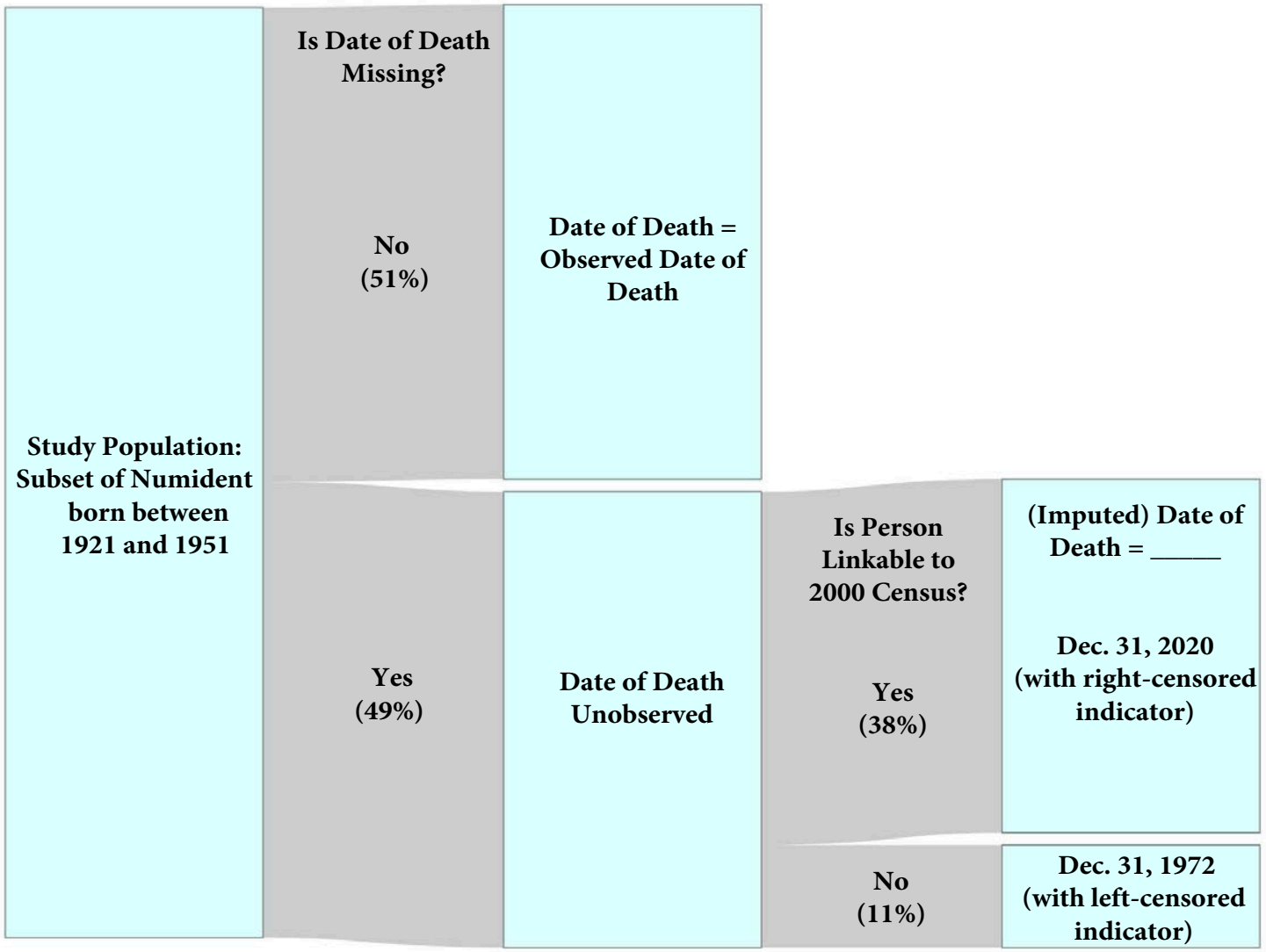
First Stage, Reduced Form, OLS

	Years of College		Lifespan in Years		
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 10-25 at Time of Opening	0.0739***		0.3016*		
	(0.0206)		(0.1200)		
Indicator: ... 5-20 Miles…	-0.0091		0.0542		
	(0.0114)		(0.0818)		
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 15-24 at Time of Opening		0.0586**		0.114	
		(0.0199)		(0.1171)	
Indicator: ... 5-20 Miles…		-0.0069		0.0574	
		(0.0113)		(0.0838)	
Years of College					0.4951***
					(0.0046)
Childhood Town FEs	Yes	Yes	Yes	Yes	Yes
Childhood County By Birth Year FEs	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.16	1.16	83.0	83.0	83.0
Observations	18,300,000	18,300,000	18,300,000	18,300,000	18,300,000
Number of Clusters	114,000	114,000	114,000	114,000	114,000

Notes: The underlying dataset consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence) and which can be linked to either the 2000 Census long form or the 2001-2015 American Community Survey. Standard errors are clustered at the level of childhood town. The omitted explanatory variable for the 10-25 age category is the 26-40 age category; the omitted explanatory variable for the 15-24 age category is the 25-35 age category. The college openings in this analysis are those that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles.

Appendix Figure 1

Imputation of Missing Death Dates



Appendix Table 1

Linkage (PIK) Rates from Numident to the 2000 and 2010 Census

	2000 Census	2010 Census
Overall PIK (Linkage) Rate	87.8%	88.4%
PIK Rates by Age (at time of Census enumeration)		
Age 45-64	*[Not Yet Available]*	93.0%
Age 65-74		94.4%
Age 75+		93.5%

Factors Affecting PIK Rate

More Likely to be PIKed:	Less Likely to be PIKed:
Taxpayers	Mobile/Transient Populations
Workers	US Residents without SSN
Persons on Medicare	Children not claimed on tax forms
Persons Receiving housing/medical/food assistance	
Men ever registered for Selective Service	

Notes: For PIK rates in upper panel, see Rastogi and O'Hara (2012), Table 12 and Masey et al. (2018), Table 4. The lower panel largely reflects a given person's presence in federal administrative data (see Appendix A of Masey et al. (2018)).

Appendix Table 2

Correlation Between Treatment Variable and Linkage to 2000 Census and American Community Survey

	Indicator: Linkable to…		
	2000 Census	2000 Census Long Form	2000 Census Long Form or 2001-2015 ACS
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 10-25 at Time of Opening	0.0080***	-0.0004	-0.0012
	(0.0022)	(0.0015)	(0.0019)
Indicator: ... 5-20 Miles…	-0.0002	-0.0010	-0.0015
	(0.0015)	(0.0012)	(0.0015)
Childhood Town FEs	Yes	Yes	Yes
Childhood County By Birth Year FEs	Yes	Yes	Yes
Dep. Var. Mean	0.78000	0.13400	0.24700
Observations	72,160,000	72,160,000	72,160,000
Number of Clusters	135,000	135,000	135,000

Notes: The underlying dataset consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence) and which has a date of death in Numident or the date of death is right censored in 2020. Standard errors are clustered at the level of childhood town. The omitted explanatory variable for each distance category (0 to 5 miles, and 5 to 20 miles) consists of those who were age 26 to 40 at the time of the college opening. SEAs are the state economic areas, which are groups of counties within states. The college openings in this analysis are those that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles.

Appendix Table 3

Effect of Community College Openings on College Attainment and Longevity, By Gender and Race

	Education Subset:			*Full Sample:*
	Years of College	Indicator: Competed At Least a Bachelor's Degree	Lifespan in Years	Lifespan in Years
Indicator: Childhood Town Brought Within 5 Miles of Nearest College by a Two-Year Public College Opening and Age 10-25 at Time of Opening	0.042	0.0102	0.3565*	0.5770***
	(0.0266)	(0.0057)	(0.1503)	(0.1124)
Indicator: Female X Indicator: Childhood Town…	0.0502	0.0051	-0.2063	-0.3713**
	(0.0274)	(0.0060)	(0.2010)	(0.1430)
Indicator: Black X Indicator: Childhood Town…	-0.0022	0.01	-0.398	0.0084
	(0.0539)	(0.0119)	(0.5046)	(0.2493)
Indicator: Black X Indicator: Female X Indicator: Childhood Town…	-0.0562	-0.0204	0.1591	-0.3159
	(0.0694)	(0.0149)	(0.6716)	(0.3300)
Childhood Town by Race by Gender FEs	Yes	Yes	Yes	Yes
Childhood County By Birth Year By Race by Gender FEs	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.16	0.19	83.0	78.800
Observations	18,300,000	18,300,000	18,300,000	72,160,000
Number of Clusters	114,000	114,000	114,000	135,000

Notes: The underlying dataset for columns 1-3 consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence) and which can be linked to either the 2000 Census long form or the 2001-2015 American Community Survey. The underlying dataset for column 4 consists of the subset of Numident born 1921-1951 with a geocodable town of birth (or geocodable 1940 town of residence). Standard errors are clustered at the level of childhood town. The omitted explanatory variable for each distance category (0 to 5 miles) consists of those who were age 26 to 40 at the time of the college opening. The college openings in this analysis are those that opened in places where the distance to the nearest college (of any kind) prior to the opening was at least 25 miles.