Geography, Education, and the Widening Divide in U.S. Midlife Mortality

Christopher Foote
Federal Reserve Bank of Boston

Ellen Meara
Harvard University and NBER

Jonathan Skinner
Dartmouth College and NBER

Luke Stewart
University of Pennsylvania

July 31, 2024

Preliminary – Please do not cite or quote

Abstract

There has been a sharp increase in the education-mortality gradient, with the gap in life expectancy between college and non-college graduates widening from 2.6 years in 1992 to 6.3 years in 2019 (Case and Deaton, 2023). During the same period, there has been a 32 percent increase in regional mortality inequality, accompanied by the emergence of a rural health penalty. Using county-level data since 1992, we demonstrate that these three trends have arisen because of a fundamental shift in the geographic patterns of mortality among college and non-college graduates. First, we find a sharp decline in both mortality rates and geographic inequality for college graduates; by 2019 mortality rates for college graduates became more similar across counties. Second, the reverse was true for non-college graduates; spatial inequality became amplified, especially in rural areas. Third, this increasing *concentration* of high mortality and poor health behaviors meant that 25 percent of counties (weighted by population) accounted for 62 percent of the increasing education gap (and 81 percent excluding deaths of despair). Finally, we find that differences in health behaviors, particularly smoking, are strongly associated with the rising educational gradient in mortality. Smoking rates among college graduates are uniformly low, while those of non-college populations vary by more across counties and are strongly associated with rising non-college mortality at the county level. In a counterfactual in which all non-college populations are assigned the average smoking rate of college graduates, the increase in the education-mortality gradient since 1992 vanishes. Other explanations for the change in the gradient—such as selection into education, changes in county income, the loss of factory jobs, or differences in state-level policies—received less support.

I. Introduction

During the past three decades, there has been a dramatic rise in the mortality-education gradient, with the gap in life expectancy between college and non-college graduates at midlife (age 25-64) widening from 2.6 years in 1990 to 6.3 years in 2019 (Case and Deaton, 2023). There has also been a rapid increase in regional mortality inequality, with the county-level coefficient of variation rising from 0.19 in 1992 to 0.31 in 2019. These trends coincide with a third shift. The health and mortality advantage historically enjoyed by residents of rural counties compared with large urban counties has reversed since the 1990s, with rural counties now experiencing a midlife mortality rate that is 140 deaths per 100,000 higher than that experienced in large urban counties. In this paper, we first consider, and reject, common hypotheses that these trends are driven primarily by rising deaths of despair, selection into college education, or immigration patterns.

Instead, we demonstrate that the three trends are symptoms of a fundamental shift in geographic patterns of health and mortality. First, the rise in spatial inequality occurs primarily for people who did not graduate from college; college graduates in all states have experienced a remarkable reduction in midlife mortality since 1992 as well as both beta and sigma convergence in mortality rates. By contrast, trends in non-college mortality—and the education-mortality gradient—have widened dramatically; in many urban counties, the education gap has barely budged (or narrowed) since 1992, while in rural counties, the education gap has doubled or tripled. We find that between 1992 and 2019, just 25 percent of counties (weighted by population) accounted for 62 percent of the increase in the education-mortality gap, and 81 percent of the gap excluding deaths of despair (poisonings, suicide, alcohol-related liver disease).

In regressions of county-level midlife mortality on the natural log of county income, we included variables to reflect hypotheses that 1) rising dispersion in the state policy environment,

2) long term declines in economic opportunity for workers in manufacturing and industries affected by global trade or automation, or 3) health behaviors such as smoking or obesity-inducing diet and activity might explain trends in mortality by education and place. In these models, broadly, the lack of mortality gains for non-college graduates are associated with behaviors in that county and income in the county. More specifically, non-college graduates in some counties experienced large declines in smoking rates over time, and large gains in income, and the mortality benefits of income and healthy behaviors rose in importance over time.

2. Model (outline)

To create a framework for understanding the changing patterns of mortality, income, health behaviors, and state policies, in the next section we adopt the Case and Deaton (2007) health depreciation model. Rather than focusing on health investment (as in Grossman, 1968), Case and Deaton (2005) focus on factors that lead to rising inequities in health across education or income groups, whether because of blue collar jobs that lead to worsening physical health, or because of consumption choices such as unhealthy food or smoking that depreciate health capital. The key insight of these models, however, is that changes in state policies, risky consumption choices, or income often affects midlife mortality with a long and variable lag; the smoking habit avoided at age 18 (because of high cigarette taxes) only reduces mortality rates 30 years later. This model also implies that if state-level policies from the 1980s and 1990s are only now bearing fruit, that the variance of mortality explained by state-level policies should be widening over time. And while other policies (such as restrictions on opioid prescribing) may have more immediate effects on deaths of despair, it is still the case that establishing causality from the closing of a manufacturing plant or raising state cigarette taxes is difficult in the usual

difference-in-difference framework. Still, the model implies that, where possible, we should allow for lagged effects of economic and state policy factors, that these could work indirectly through their influence on health behaviors such as smoking, obesity, and inactivity. For this reason, we consider the most flexible measure of state-level policy effects, which are state-year fixed effects; thus, we view these measures as the upper bounds on the potential average effects of state-level policies.

3. Data

We use the restricted individual-level mortality measures from the National Center for Health Statistics (NCHS). These measures come from death certificates and include the causes of death for each decedent and demographic information such as age, sex, race, ethnicity, education, and county of residence. We compute mortality rates in a given county or state by dividing the total number of deaths by the relevant population from the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER).

SEER has county-level population data for demographic subgroups, allowing us to compute mortality rates by race, ethnicity, sex, and age. We compute age-specific mortality rates within the age bins of 25-34 years old, 35-44 years old, 45-54 years old, and 55-64 years old, then we age-standardize these rates using the 2000 decennial Census age distribution for an overall midlife mortality rate. For mortality data before 1989, we use publicly available data from the CDC (Centers for Disease Control) WONDER database.

Beginning in 1989, death certificates include education level; while some states were slow to report education, by 1992 most states were reporting education on death certificates. We

limit our attention to the 44 states with the most accurate educational reporting.[1] Education-specific populations by county are not readily available from SEER. To compute these populations, we use data from the Census and the American Community Survey (ACS).

The smallest identifiable geographic area in this data is the Public Use Microdata Area (PUMA), which must have populations over 100,000. County boundaries fall cleanly within PUMA boundaries. Frequently, multiple counties are combined to make up a PUMA, but in denser areas, counties map directly to PUMAs or multiple PUMAs make up one county. Generally, we compute the proportion of the population of an age-race-sex-PUMA cell by educational attainment, then apply these proportions to PUMA-county crosswalks to compute the proportion of the population of an age-race-sex-county cell by educational attainment. We then multiply these shares by the total population within the cell taken from SEER.

The Census data are only available in 1990, 2000, and 2010, while the ACS data are available from 2006 onward (in the form of multiple-year ACS estimates). Due to constraints on the quality of education data in the mortality records in 1989-1991 and to avoid complications due to the COVID-19 pandemic, we conduct education-specific analyses in 1992, 2000, 2010, and 2019. To get county-level education-specific population estimates in 1992, we linearly interpolate between the 1990 and 2000 Census figures.

For data on rates of health behaviors and outcomes (like smoking, obesity, and inactivity), we compute county-level averages from the Behavioral Risk Factor Surveillance System. The County FIPS code is available in the BRFSS microdata from 1996 to 2010, so we pool the microdata from these years to estimate county-level rates of these outcomes. We also

[1] We omit Alaska, DC, and Hawaii in all analyses. For education-specific analyses, we omit Georgia, Oklahoma, Rhode Island, and South Dakota.

use the County Health Ranking data, which includes detailed county-level data on health and social determinants going back to roughly 2007.[2]

Although changes in the definitions and boundaries of counties are rare, they do occur in our sample. For counties that split or combine partway through our years of interest, we treat them as one combined unit in all years. Per capita income data for both states and counties come from the Bureau of Economic Analysis. For further discussion of data sources and a full description of the counties in our sample, see the data appendix.

4. Basic Empirical Patterns

Case and Deaton (2023) estimated an increase in the non-college graduate – college graduate life expectancy gap from 2.6 years in 1992 to 6.3 years in 2019. (The gap widened further during the Covid-19 pandemic, but we restrict our attention to the period 1992-2019.) This increase is unlikely to have resulted from causal effects of education since one would have to explain why the impact of education on mortality increased so rapidly over time. A simple estimate of the welfare cost arising from this rise in the mortality gap is just shy of $40 trillion.[3]

One concern with this estimated rise in the mortality gradient is that it simply reflects selection as a higher percentage of the midlife population has graduated from college (Hoxby, 2023); thus, adults who by 2019 were not college graduates could be systematically in poorer

[2] See https://www.countyhealthrankings.org/. Because there is a roughly 3-year lag between the County Health Data sources and when they are issued, we label (e.g.) the 2010 release as 2007.

[3] The counterfactual is one in which the gap would have remained constant at 2.6 years; thus the population of people in midlife has experienced a loss of 3.7 years (6.3 minus 2.6). Valuing a life year at $100,000 and multiplying times the number of people in 2019 between ages 25-64 without a college degree yields $39 trillion. This would be an underestimate if we ignore future generations not yet in their midlife who similarly experience lower life expectancy (and life-years past age 85, which is the Case and Deaton upper limit) but an overestimate because future deaths have not been discounted.

health than non-college graduates in 1992. We evaluate the selection argument in **Figure 1**, which plots the education-specific mortality rate (y-axis) on education percentile of college graduates (x-axis) by year for college and non-college graduates. **Figure 1** demonstrates that the education percentile of college and non-college populations, defined here as the midpoint of each education group's percentile range, changed little, falling from the 88th to 83rd percentile for college graduates, and from the 38th to the 33rd for non-college graduates. Yet, mortality rates decrease monotonically for college graduates from 1992 to 2019 (262 to 133 deaths per 100,000), while, for non-college graduates, the decline slows after 2000 and deaths per 100,000 rise after 2010, such that mortality rises slightly comparing 1992 and 2019 (458 vs. 463 deaths per 100,000). Indeed, as shown by **Figure 1**, the national increase in the educational gradient results from a *drop* in the college mortality rate that took place alongside a relatively *stable* non-college mortality rate. This is the opposite pattern that one would expect if selection were driving the change in the education gradient.

One may still be concerned that most of the increase in non-college mortality arises solely because of high school dropouts rather than high-school graduates, as suggested by Novosad et al. (2022). However, in Appendix., 2022, Figure A.1, we use the Novosad et al. approach to show that the progression across the education gradient is monotonic, with mortality rates for college graduates dropping more quickly than for high school graduates.

In **Figure 2**, we document the gradual increase in spatial inequality, as measured by the coefficient of variation (weighted by population) by county since 1992. The overall growth is from 0.19 in 1992 to 0.31 in 2019, with a more rapid increase during the 2000s and 2010s. **Figure 3** links the education-gap analysis to the trend of rising relative mortality rates in rural areas. The panel shows that, although the overall non-college mortality rate was roughly stable

from 1992 to 2019, in the most densely populated urban counties, this rate fell significantly (Group 1). Moving to increasingly rural counties in Group 6, however, this pattern is reversed. A qualitatively similar urban-rural gradient is also found in college mortality rates, but it is much less pronounced, and college mortality declined even in the most rural areas. Putting the non-college and college patterns together indicates that the largest differences between non-college and college mortality changes were concentrated in the most rural counties.

The national mortality gap can be modelled as a population-weighted average of the mortality gap in each county, defined as the county's non-college mortality rate minus the county's college mortality rate. In turn, the *change* in the national gap is a population-weighted average of each county's change in non-college mortality minus its change college mortality.

Figure 4 plots population-weighted kernel densities for 1992 and 2019 mortality rates among non-college populations (top panel) and college populations (bottom panel). For non-college persons, average mortality rates changed little between 1992 and 2019, although the county-level variance of non-college mortality increased. By contrast, the average college mortality rate fell by about 100 per 100,000 persons over this period. Combined with the near-zero change in mean non-college mortality, the college mortality decline essentially accounts for the rising educational gradient from 1992 to 2019.

The lower panel of **Figure 4** also shows that the county-level *variance* of college mortality rates declined from 1992 to 2019, in contrast to the pattern for non-college mortality. This pattern indicates that the trend of rising geographic variance of midlife mortality rates is driven by non-college populations.

The top panel of **Figure 5** depicts population-weighted densities of within-county changes in mortality rates for the two educational groups. College mortality changes (red line)

are closely distributed around the mean of approximately 100 fewer deaths per 100k persons, while non-college changes (blue line) are widely distributed around zero. The lower panel shows the population-weighted density of county-level gap changes (each county's non-college mortality change less its college mortality change). The distribution is skewed to the right, with counties experiencing large increases in the gap accounting for a disproportionate large share of the overall gap change. We can get a sense for the importance of these counties on the right side of the distribution by calculating the fraction of increase attributed to the counties with the greatest increase in the education-mortality gap. For example, based on this graph, we find that 25% of counties (weighted by population) accounts for 62% of the overall increase in the education gap. One may be concerned that these results are driven by rising deaths of despair (defined in Case and Deaton (2015) as deaths arising from suicides, alcohol, and drug overdoses) concentrated in a small number of counties. However, when excluding deaths of despair, the proportion of the increased educational gradient explained by 25 percent of (population-weighted) counties increased to 81 percent. This is because deaths of despair had such an adverse impact on non-college mortality even those counties with otherwise stable education gradients.

5. Factors Explaining the Rising Education Gradients in Mortality

5.1 Within-County Changes in College and Non-College Mortality Rates: 1992 to 2019.

What variables explain the disparate behavior of education-specific mortality changes at the county level? We answer this question with a series of regression models that place county-level changes in either college or non-college mortality rates on the left-hand-side. First, the declining variance of college mortality rates seen in the lower panel of **Figure 4** suggests convergence in college mortality rates. To borrow from the language of macroeconomics, *sigma*

convergence occurs when a cross-sectional variance declines over time, as seen with the college mortality rates in Figure 1. Sigma convergence is typically accompanied by *beta convergence*, which in this case would occur if counties with relatively high college mortality rates in 1992 experienced relatively large declines in college mortality from 1992 to 2019. The binned scatter plots in **Figure 6** show that evidence for beta convergence is exceptionally strong for college mortality rates. For non-college rates, there is some evidence of beta convergence, but it is weaker.

Second, the widening variance of non-college mortality rates (that is, the lack of non-college sigma convergence) evident in the top panel of **Figure 4** suggests that several factors are pushing apart the county-level distribution of non-college rates, thereby offsetting the tendency of (weak) beta convergence to push the non-college rates closer together. The binned scatter plot in the top left panel of **Figure 7** shows that one factor generating variance in non-college rates is the smoking rate, which has a strong positive correlation with changes in non-college mortality. As noted above, we use BRFSS data to estimate smoking and obesity rates by educational group and county between 1996 and 2010.[4] In counties with non-college smoking rates below 20%, non-college mortality rates fell by about 100 per 100,000, close to the mean decline for the college population overall. Conversely, counties with non-college smoking rates of 30% or more averaged over this period saw their non-college mortality rates *rise* by about 100 per 100,000.

The upper right panel of **Figure 7** shows that the college smoking rate is much less correlated with college mortality rates, primarily because there is very little dispersion in college smoking rates—across counties, college smoking rates are less than 10%. The lower panels of **Figure 7** show that obesity rates are positively correlated with changes in mortality for both

[4] The smoking and obesity rates are constant within each education group/county cell, because BRFSS sample sizes are not large enough to estimate yearly rates at the education-county level.

educational groups, but that obesity is less likely to drive differences in mortality changes across the two educational groups. The lower panels show that obesity rates are more similar across the groups, and that obesity has a smaller correlation with mortality than smoking does (note the difference in vertical scales across the top and bottom panels of the figure).

Finally, changes in college and non-college rates can also be correlated differently with common countywide variables, most importantly those that reflect local economic conditions. **Figure 8** shows binned scatter plots of college and non-college mortality changes against changes in countywide average per capita income (top panel) and changes in employment-to-population ratios (bottom panel). The top panel shows that changes in income are more strongly correlated with non-college mortality than with college mortality. This pattern suggests that non-college populations are more sensitive to local economic conditions than college graduates. However, the lower panel shows that measuring conditions with employment rather than income provides less evidence for this conjecture.

5.2 Accounting for County-Level Influences Simultaneously.

We next turn to regression models that account for potential determinants of education-specific changes in county-level mortality rates. **Tables 1 and 2** project changes in non-college and college mortality rates on a collection of covariates, including education-specific smoking and obesity rates from BRFSS; average countywide per capita income in 1992, and the change in income between 1992 and 2019; the 1992 employment-to-population ratio and its change; the county's 1992 manufacturing share of employment; an index of progressivity in state policy, such that higher values are consistent with more-liberal economic and social policies;[5] the

[5] Our policy index, created by CW, borrows from the political science literature on… **MORE HERE**

county's 1992 college population share (Bor et al, 2024); and indicator variables for the six NCHS urban-rural designations.[6] The first nine regressions in each table account for a potential influence on mortality-rate changes individually, while the tenth regression combines them. All regressions in Tables 1 and 2 are weighted by the county's education-specific population in 1992, and standard errors are clustered by state.

Turning first to the non-college analysis in Table 1, each of the first nine regressions indicates a non-trivial correlation of a particular variable (or set of variables) with county-level changes in non-college mortality. Column 1 shows that the NCHS indicator variables explain about 26 percent of the variance of non-college mortality change, while column 2 shows that the smoking and obesity rates account for about 32 percent. Columns 3 through 9 show that individual covariates reflecting economic conditions or policy activity are also correlated strongly with non-college mortality changes, as each of the individual covariates is significant at the 1 percent level or higher.

Placing all covariates in the same regression, as in column 10, causes many of them to decline in absolute value and lose statistical significance. The smoking rate, the NCHS indicator variables, the 1992 level of county income, and the change in county income remain strongly significant. The manufacturing share remains significant at 10%.[7] All told, the combined regression in column 10 explains about 47 percent over the overall variance in non-college mortality changes.

Table 2 depicts the analysis of college mortality changes, indicating less evidence of correlation with the potential influences than in the non-college analysis. Consistent with the

[6] The indicator for the most-urban counties (NCHS Code = 1) is omitted.
[7] The college share remains significant but switches sign; in column 10, higher college shares are associated with increases in non-college mortality.

lower panel of **Figure 3**, there is less of an urban-rural gradient in college mortality changes (column 1). In general, the economic and policy covariates generate much lower R^2 than they did for the non-college mortality changes in **Table 1**.

In **Table 3** we extend the analysis in two ways. First, we include state fixed effects in some regressions, and second, we allow for the 1992-2019 change in a county-education mortality rate to depend on the initial (1992) level of this rate. Column 1 of this table replicates the last non-college regression of **Table 1**, and column 2 adds state fixed effects to this regression.[8] Columns 3 and 4 repeat the first two regressions but adds the 1992 non-college mortality rate. This initial rate enters with a high level of statistical significance, and its inclusion has strong effects on some of the other coefficients. In particular, the coefficient on the obesity rate, which is negative in the first two columns of the table, becomes strongly positive and significant. Additionally, the coefficients on the NCHS indicator variables decline by about ½ or more.

Columns 4 through 8 repeat the analysis for changes in college mortality rates. The most important development is that the strong evidence for beta convergence in college mortality rates seen earlier results in a very large coefficient on the 1992 mortality rate (about -0.9 with or without state fixed effects). Here again, the inclusion of the initial rate has a strong effect on other variables; the smoking and obesity variables both become positive and strongly significant.

Bor et al. (2024) found that a higher share of college-educated county residents resulted in lower levels of mortality within narrowly defined demographic groups. Our regressions, which model mortality changes, indicate that the 1992 college share is correlated with higher non-college mortality changes when state fixed effects and the 1992 mortality rate are left of the

[8] Because the state-policy variable is determined at the state level, it is dropped from the regression with state fixed effects.

regression. This counter-intuitive result is eliminated when either the state fixed effects or the initial mortality rate is included. A possible reconciliation of these results with those in Bor et al. (2024) is that the college share does affect mortality levels, but this share either has little effect on changes or has an effect that runs through other variables in the model (such as the smoking rate).

Table 4 presents what we call our baseline regressions, which include the initial mortality rates but leave out state fixed effects. The baseline regressions also leave out the employment-to-population variables, which were insignificant or wrong-signed in the regressions of the previous table. Finally, we also omit the manufacturing and college shares, which were insignificant in the non-college regressions. Looking across the two columns of this table indicates that most covariates affect non-college and college mortality in ways that accords with intuition—higher smoking and obesity rates are associated with increases in mortality rates for both educational groups, while higher levels of 1992 income and higher changes in income are associated with lower changes in mortality. Yet quantitatively, the correlations are larger with non-college mortality changes than they are with college mortality changes.

5.4 Robustness checks. [additional checks to be added] One concern with our approach is that we include only smoking and obesity as measures of health behaviors. Yet, as is well known, there are many other unmeasured social health determinants that may be correlated with county smoking rates. To test for this, we use the County Health Rankings data for 2007 (released in 2010) and 2019 (released in 2022).[9] As shown in Appendix Figure A.2, the increasingly close association between smoking and mortality between 2007 and 2019 that we find using the

[9] The advantage of these data is that they include a variety of different measures at the county level. The disadvantage is that they use random effects models (which conditions on poverty rates) to create measures for the smaller counties (where sample sizes tend to be modest), so there may be a mechanical correlation among the independent variables.

smaller set of covariates is largely unaffected even after conditioning on an extensive list of covariates related to social determinants of health.[10]

While we have not completed a full sensitivity analysis by sex, race, and ethnicity, we note that the basic pattern of a widening educational gradient since 1992 is found when considering Black and Hispanic populations separately (Appendix Figure A.3). Finally, we are concerned that sharp declines in mortality for the non-college population in (e.g.) New York and California may be the consequence of the increasing fraction of the population comprising immigrants, who may be healthier than non-immigrants. Because more than two-thirds of immigrants during the past several decades identified as Hispanic and Asian,[11] we can contrast state-level mortality rates limited to Black and White populations (with a much smaller fraction of immigrants) with the corresponding mortality rates for the entire population; as shown in Appendix Figure A.4 the spatial distribution of mortality is nearly identical to that found in the full sample used in our analysis.

6. Discussion and Conclusion

The rise in the education gradient in mortality has been documented in previous literature (e.g., Meara et al., 2008; Galama et al., 2018), with the most recent (Case and Deaton, 2023) showing an accelerating rise through 2019 (and another jump in the gap during the Covid-19 pandemic). Less well understood, however, are two additional worrisome trends in midlife

[10] Controls include county population rates of obesity, fair-poor health, teen birth, low birthweight, fatal crashes, homicide, being uninsured, preventable hospital admissions, child poverty, single parent households, and county rural status.

[11] https://www.pewresearch.org/short-reads/2024/07/22/key-findings-about-us-immigrants

mortality during the same period. The first is a 32 percent increase in regional mortality inequality (e.g., Couillard et al., 2022), while the second is the emergence of a strong rural health penalty in midlife. In this paper, we attempt to show that the three trends are closely related and explained by a fundamental divergence in college and non-college patterns of mortality. For college graduates, mortality has continued to decline and exhibited both sigma and beta convergence. While college-educated county-level mortality rates in Georgia and California may have differed in 1992, by 2019 they have converged substantially. A similar story holds for health behaviors; smoking rates exhibited wide variations across counties for college graduates in 1992; by 2019 they had converged to an extremely low average rate.

By contrast, mortality for non-college graduates barely budged on average between 1992 and 2019, while the spatial variance of mortality grew. A similar pattern of increasing spatial inequality in smoking was also observed for non-college graduates. This resulted in an increasing concentration of counties with high non-college mortality and poor health behaviors. Indeed, 25 percent of counties (weighted by population) accounted for 62 percent of the increasing education gap, or 81 percent of the change in mortality excluding deaths of despair. Analogous to the Chetty et al. (2016) study of income, our findings support the view that health effects related to "place" matter far more at the bottom than the top of the distribution. In this study, place matters much more for non-college graduates than for college graduates.

We acknowledge that our results are all associations rather than causation, and so finding (for example) that the manufacturing share in the 1990s does not predict subsequent increases in mortality rates does not mean that a causal effect is not present. Similarly, the extraordinarily strong predictive impact of the county smoking rate may not entirely reflect causal effects. In a counterfactual in which all non-college populations are assigned the average smoking rate of

college graduates, the increase in the education-mortality gradient since 1992 vanishes. Of course, smoking itself is a major health hazard, but its predictive power could also mean that it is correlated (more so than for obesity) with other unmeasured factors. And while our sensitivity analysis in the appendix included natural candidates for social determinants of health as child poverty, violent crimes, single parenting, and preventable hospitalizations, they did not diminish the rising importance of smoking rates in explaining health.

Montez et al. (2020) and others have suggested that the divergence in state policies could explain the widening spatial patterns of mortality across states. By using county-level mortality data, we can test the importance of state-level policies against county-level factors such as health behaviors and income. And while the regression analysis using the state level policy index did not suggest an influential role in predicting mortality changes, we acknowledge that the index may be imperfect. Including state-level fixed effects suggests a much larger impact for state policy, implying that the upper bound estimate of state policies could still be quite large.

We did not address the dramatic widening in the education gradient that occurred during the Covid-19 pandemic, as our primary interest is in the long-term evolution of spatial and educational differences in health behaviors and health outcomes in the U.S. We anticipate, however, that the trends we have identified in the data through 2019 will only continue to worsen even as Covid-19-related mortality continues to subside.

References

Bor, Jacob H., David M. Cutler, Edward L. Glaeser, and Ljubica Ristovska. “Human Capital Spillovers and Health: Does Living Around College Graduates Lengthen Life?” National Bureau of Economic Research No. 32346 (2024).

Case, Anne, and Angus Deaton. "Broken Down by Work and Sex: How Our Health Declines." *Analyses in the Economics of Aging* (2007): 185.

Case, Anne, and Angus Deaton. "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century." *Proceedings of the National Academy of Sciences* 112, no. 49 (2015): 15078-15083.

Case, Anne, and Angus Deaton. "Accounting for the Widening Mortality Gap between American Adults with and without a BA." *Brookings Papers on Economic Activity* (2023).

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong et al. "Social capital I: measurement and associations with economic mobility." *Nature* 608, no. 7921 (2022): 108-121.

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong et al. "Social capital II: determinants of economic connectedness." *Nature* 608, no. 7921 (2022): 122-134.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. "The association between income and life expectancy in the United States, 2001-2014." *Jama* 315, no. 16 (2016): 1750-1766.

Couillard, Benjamin K., Christopher L. Foote, Kavish Gandhi, Ellen Meara, and Jonathan Skinner. "Rising geographic disparities in US mortality." *Journal of Economic Perspectives* 35, no. 4 (2021): 123-146.

Galama, Titus J., Adriana Lleras-Muney, and Hans Van Kippersluis. "The effect of education on health and mortality: a review of experimental and quasi-experimental evidence." (2018).

Hoxby, Caroline, Comment on "Accounting for the Widening Mortality Gap between American Adults with and without a BA." *Brookings Papers on Economic Activity* (2023).

Meara, Ellen R., Seth Richards, and David M. Cutler. "The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000." *Health affairs* 27, no. 2 (2008): 350-360.

Montez, Jennifer Karas, Jason Beckfield, Julene Kemp Cooney, Jacob M. Grumbach, Mark D. Hayward, Huseyin Zeyd Koytak, Steven H. Woolf, and Anna Zajacova. "US state policies, politics, and life expectancy." *The Milbank Quarterly* 98, no. 3 (2020): 668-699.

Novosad, Paul, Charlie Rafkin, and Sam Asher. "Mortality change among less educated Americans." *American Economic Journal: Applied Economics* 14, no. 4 (2022): 1-34.

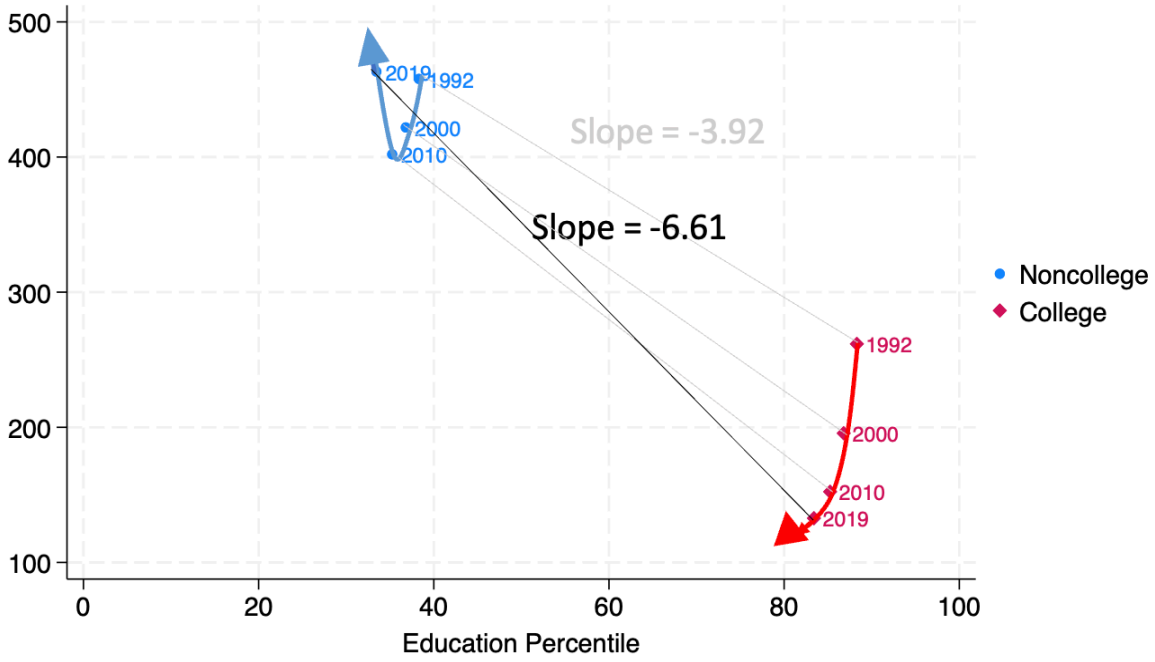


Figure 1: Midlife Mortality Rates by Education and Education Percentile

College = 4-year degree or more. Noncollege = less than 4-year college degree.
Mortality rate = deaths per 100,000 adults aged 25 to 64 age-adjusted to year 2000 population.

\

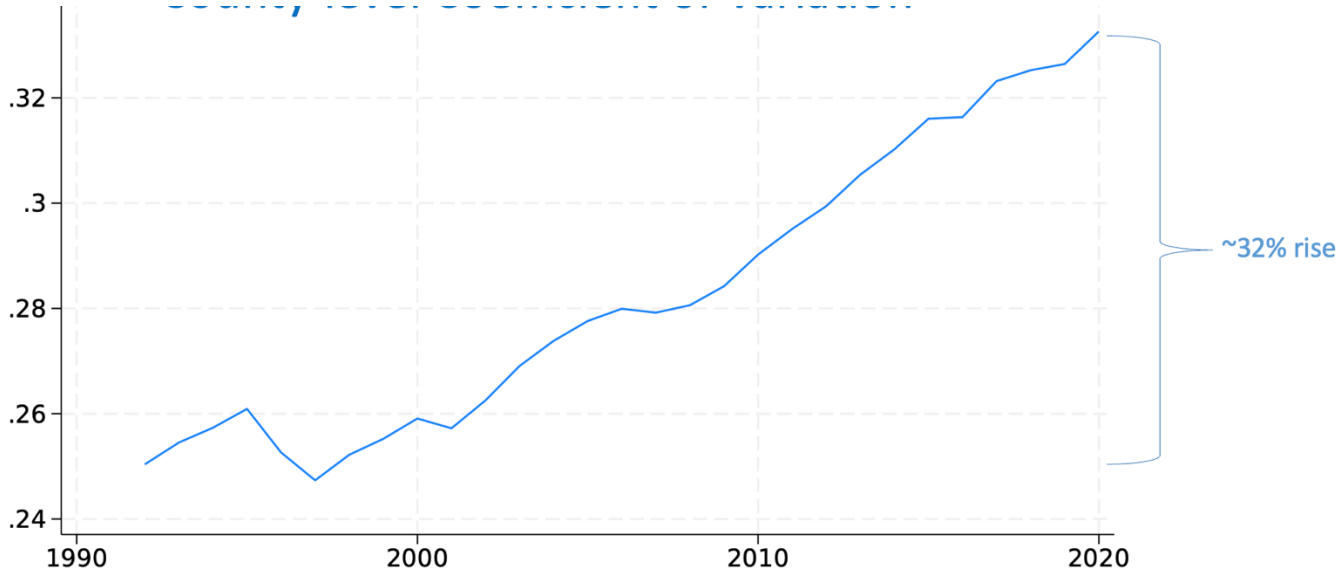


Figure 2: The County-Level Coefficient of Variation in the U.S., 1992-2019

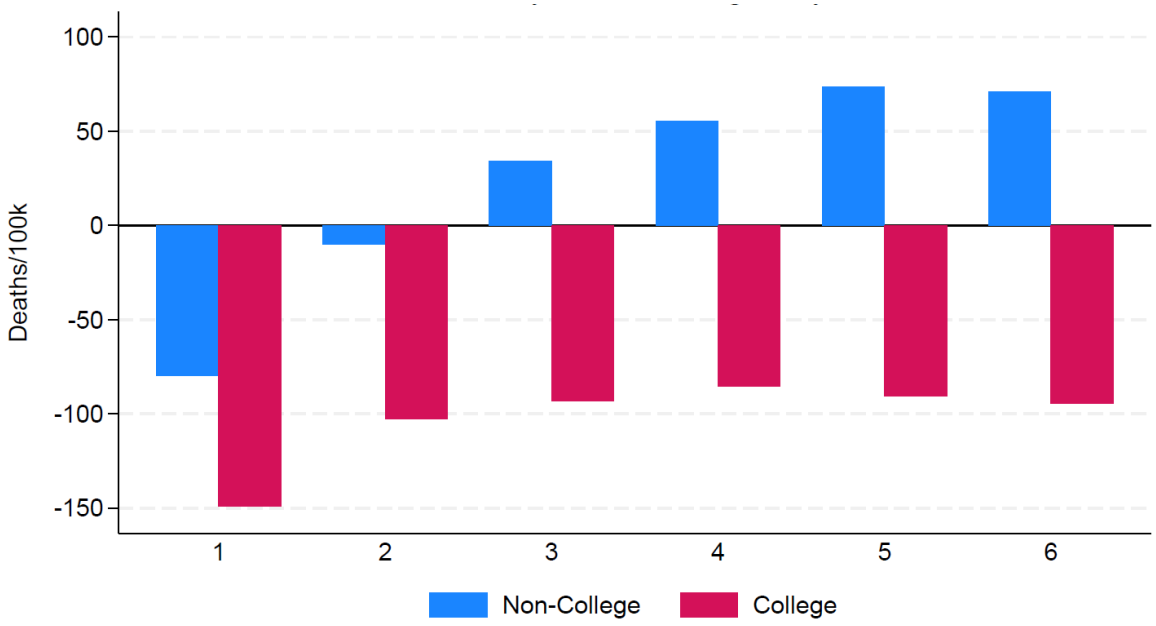


Figure 3: Urban-Rural Designations from the National Center for Health Statistics and Midlife Mortality-Rate Changes by Education, 1992-2019

1 = Large central metro, 2 = Large fringe metro, 3 = Medium metro, 4 = Small metro, 5 = Micropolitan 6 = Non-metropolitan

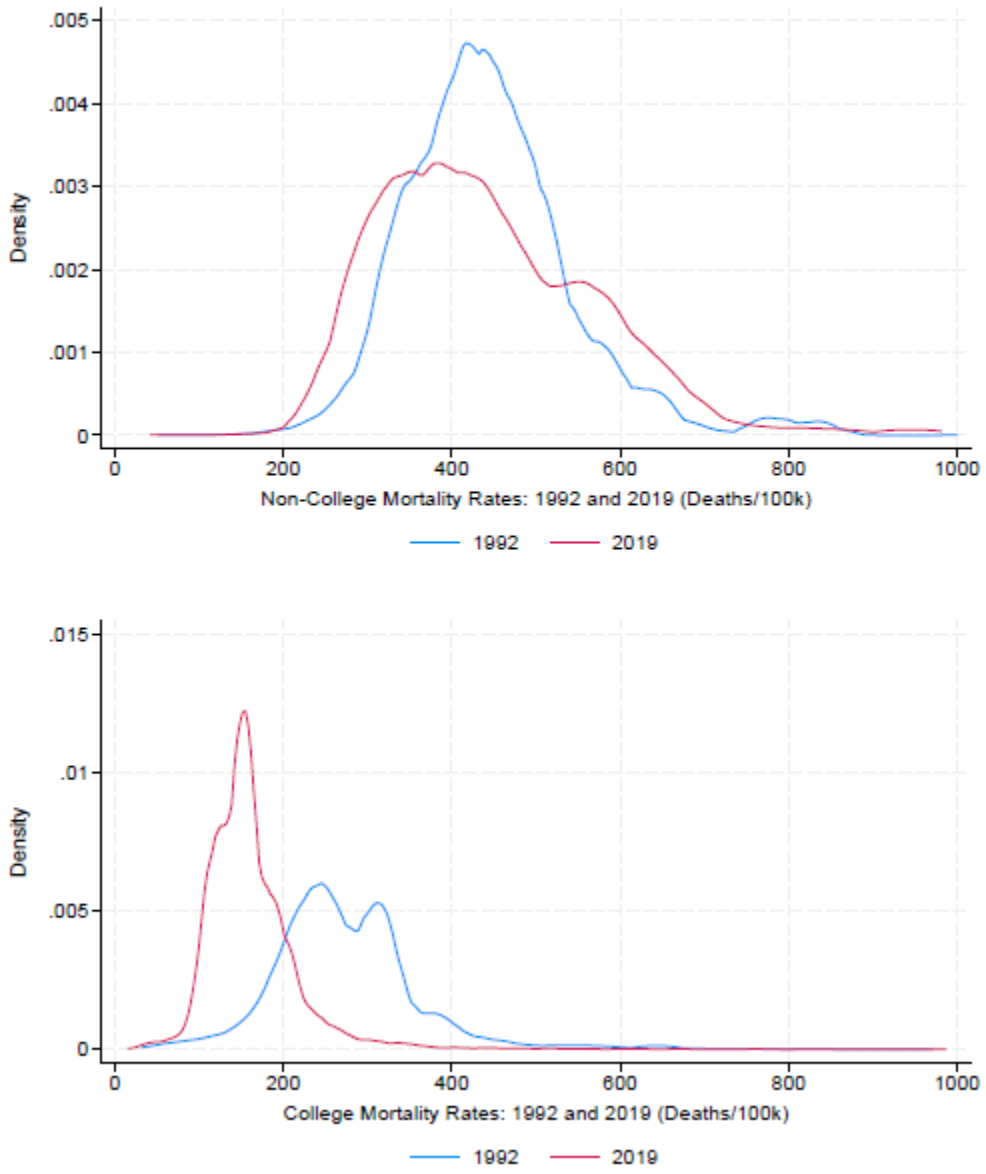


Figure 4: County-Level Densities of Mortality Rates per 100,000 Population in 1992 and 2019: Non-College Populations (Top Panel) and College-Graduate Populations (Bottom Panel). Densities are weighted by population.

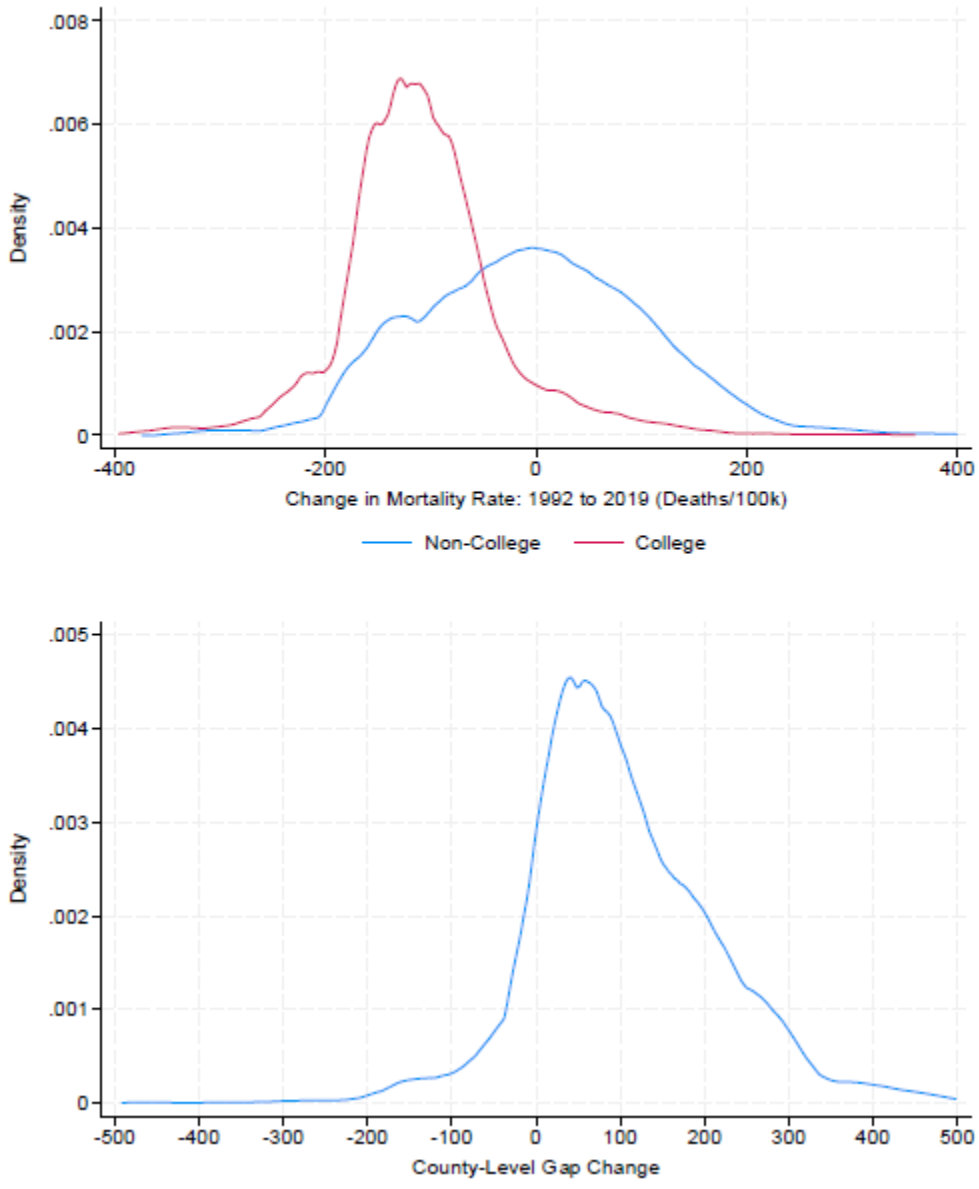


Figure 5: County-Level Densities of Changes in Mortality Rates from 1992 to 2019 by Educational Attainment (Top Panel) and Density of County-Level Changes in Mortality-Rate Gaps (Bottom Panel). Densities are weighted by population. Gap changes depicted in the lower panel are defined as the change in a county's non-college mortality rate minus the change in its college mortality rate.

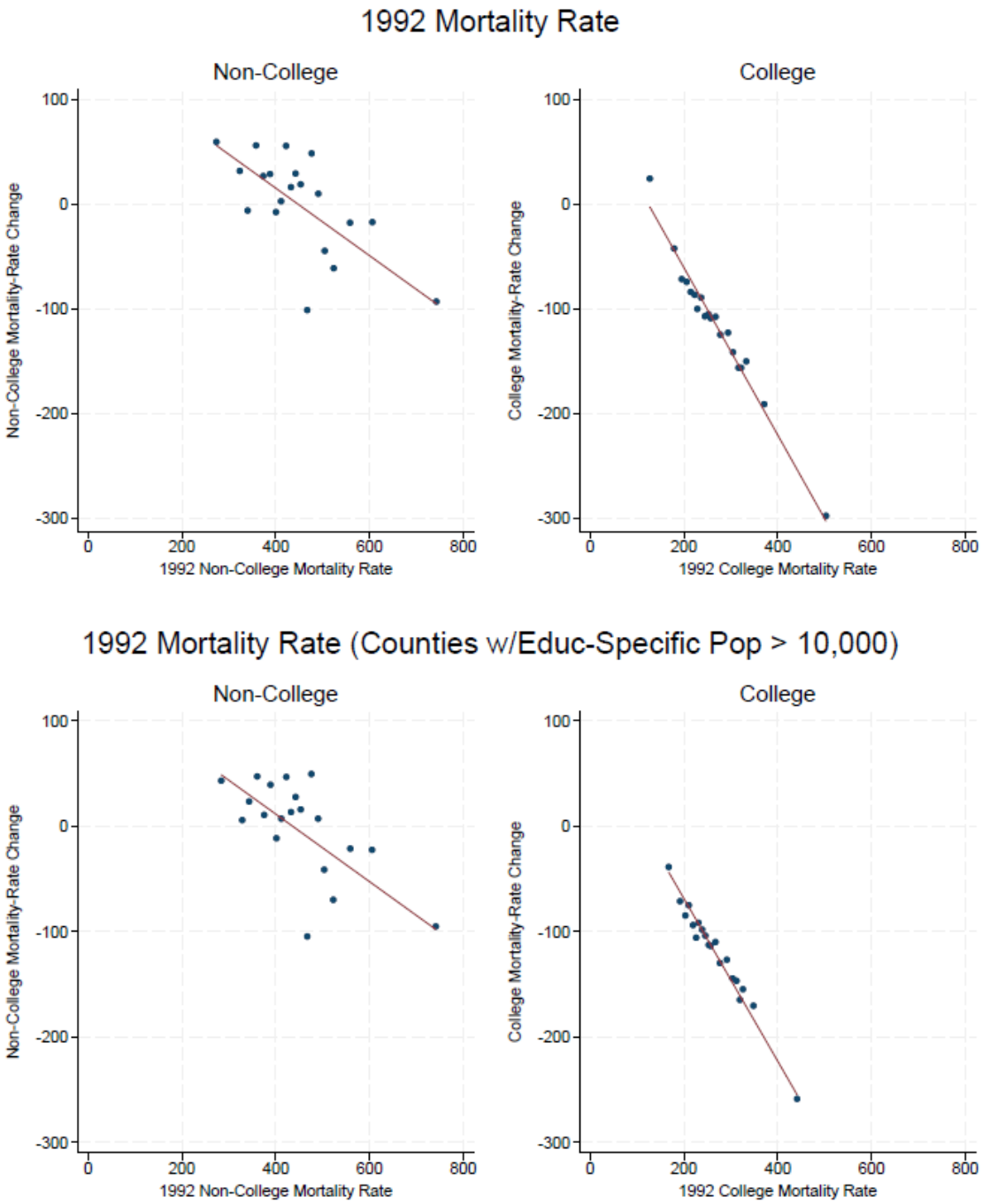


Figure 6: Correlation of Mortality-Rate Changes with Education-Specific Mortality Rates in 1992. The lower panels exclude any county-education observations with a 1992 population of less than 10,000 people. Data weighted by county population.

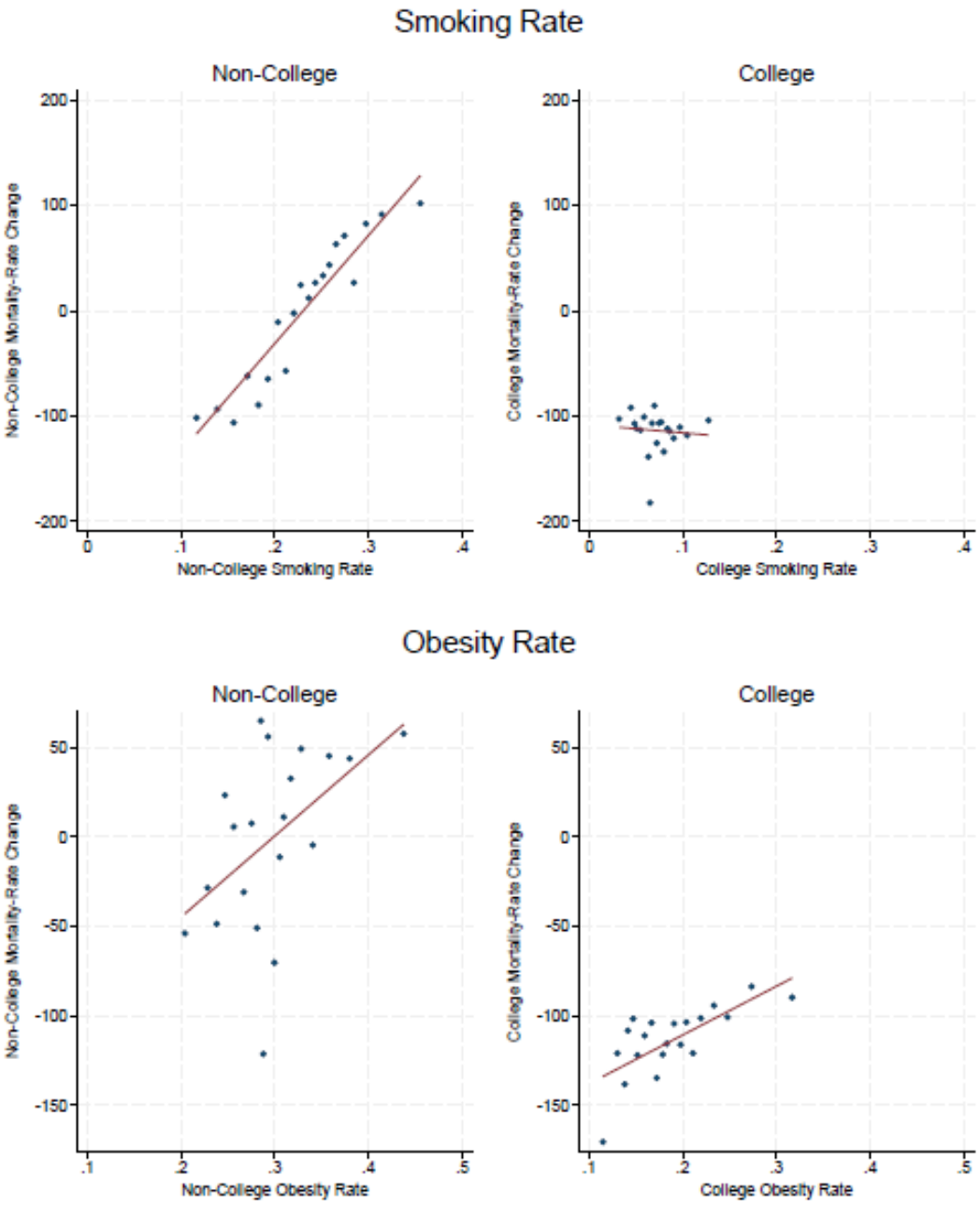


Figure 7: Correlations of Mortality Rate Changes with Education-Specific Smoking and Obesity Rates. Rates are derived from individual-level BRFSS data and correspond to averages within the county and educational group from 1996 to 2010. Data are weighted by population. Vertical and horizontal scales for the top panels are different than those for the bottom panels.

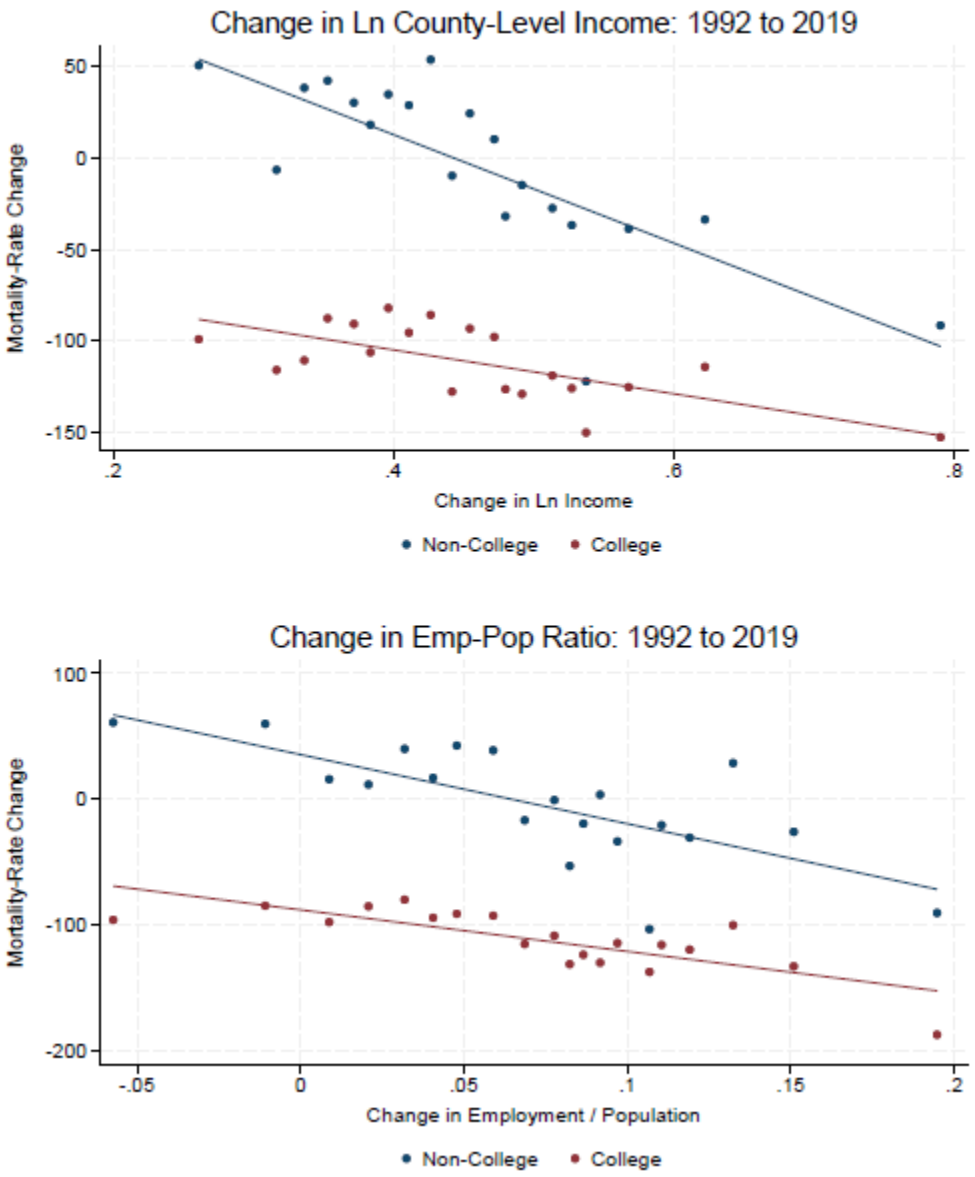


Figure 8: Correlations of Mortality-Rate Changes with Changes in Per-Capita County Level Income (Top Panel) and County Employment-Population Ratios (Bottom Panel). Changes are calculated between 1992 and 2019. Data weighted by population

Table 1. Non-College Mortality-Change Regressions

	(1) Δ Mort	(2) Δ Mort	(3) Δ Mort	(4) Δ Mort	(5) Δ Mort	(6) Δ Mort	(7) Δ Mort	(8) Δ Mort	(9) Δ Mort	(10) Δ Mort
Smoking Rate		984.9***								718.4***
		(105.4)								(92.97)
Obesity Rate		218.9**								–87.43
		(67.00)								(60.71)
Ln 1992 County Income			–196.5***							–108.8**
			(29.48)							(39.72)
Chg Ln Cnty Income 92-19				–290.3***						–124.4***
				(58.36)						(28.13)
1992 College Share					–464.8***					148.8*
					(65.53)					(64.56)
State Policy Index						–37.67**				–7.881
						(13.36)				(5.090)
1992 Emp-Pop Ratio							–159.0***			34.21
							(37.47)			(54.00)
Chg Emp-Pop 92-19								–551.6***		129.2
								(145.7)		(101.5)
1992 Manufacturing Share									321.1***	75.13+
									(53.19)	(39.77)
NCHS Urban-Rural Code=2	71.64***									59.77***
	(17.95)									(15.00)
NCHS Urban-Rural Code=3	113.8***									75.29***
	(11.88)									(12.52)
NCHS Urban-Rural Code=4	135.3***									89.27***
	(19.89)									(14.80)
NCHS Urban-Rural Code=5	153.8***									105.4***
	(20.39)									(12.34)
NCHS Urban-Rural Code=6	149.2***									97.15***
	(21.48)									(13.50)
Constant	–78.68***	–291.8***	2041.3***	130.5***	104.5***	9.677	83.96***	37.94**	–47.67*	922.8*
	(21.24)	(37.04)	(298.2)	(21.97)	(14.01)	(12.83)	(17.25)	(11.51)	(20.94)	(412.3)
R^2	0.263	0.322	0.137	0.092	0.119	0.113	0.028	0.083	0.053	0.471
Observations	1956	1956	1956	1956	1956	1956	1956	1956	1956	1956

Notes: Regressions are weighted by 1992 population. The dependent variable in each regression is the change in the non-college mortality rate per 100k population between 1992 and 2019. Standard errors are clustered by state. + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 2. College Mortality-Change Regressions

	(1) Δ Mort	(2) Δ Mort	(3) Δ Mort	(4) Δ Mort	(5) Δ Mort	(6) Δ Mort	(7) Δ Mort	(8) Δ Mort	(9) Δ Mort	(10) Δ Mort
Smoking Rate		–302.1**								–120.5
		(104.5)								(144.1)
Obesity Rate		315.0**								46.85
		(115.1)								(89.39)
Ln 1992 County Income			–40.20							–43.03*
			(25.09)							(17.95)
Chg Ln Cnty Income 92-19				–144.6**						–84.51*
				(44.36)						(36.47)
1992 College Share					–28.39					273.4***
					(57.62)					(34.84)
State Policy Index						–7.373				–3.733
						(7.145)				(5.230)
1992 Emp-Pop Ratio							–107.4+			–106.2
							(53.91)			(63.90)
Chg Emp-Pop 92-19								–371.5**		–214.1*
								(130.2)		(93.08)
1992 Manufacturing Share									141.5**	91.29*
									(43.45)	(38.89)
NCHS Urban-Rural Code=2	46.66***									28.26***
	(7.769)									(6.179)
NCHS Urban-Rural Code=3	55.85***									32.28***
	(9.367)									(7.632)
NCHS Urban-Rural Code=4	66.95***									46.13***
	(9.380)									(7.671)
NCHS Urban-Rural Code=5	59.11***									38.27***
	(10.34)									(7.973)
NCHS Urban-Rural Code=6	59.71***									33.81***
	(12.18)									(9.399)
Constant	–148.6***	–151.4***	307.5	–45.74*	–106.1***	–110.7***	–53.01*	–83.74***	–132.7***	348.9+
	(8.882)	(26.55)	(258.0)	(18.34)	(11.80)	(5.535)	(24.92)	(6.734)	(11.41)	(183.3)
R^2	0.116	0.040	0.013	0.057	0.001	0.009	0.028	0.072	0.016	0.199
Observations	1953	1953	1953	1953	1953	1953	1953	1953	1953	1953

Notes: Regressions are weighted by 1992 population. The dependent variable in each regression is the change in the college mortality rate per 100k population between 1992 and 2019. Standard errors are clustered by state. + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 3. Non-College (N-C) and College (COL) Mortality-Change Regressions, with and without State Fixed Effects

	(1) N-C-noFE	(2) N-C-w/FE	(3) N-C-noFE	(4) N-C-w/FE	(5) COL-noFE	(6) COL-w/FE	(7) COL-noFE	(8) COL-w/FE
1992 Mortality Rate			–0.442***	–0.482***			–0.889***	–0.911***
			(0.0395)	(0.0448)			(0.0146)	(0.0146)
Smoking Rate	718.4***	616.3***	847.6***	757.9***	–120.5	–93.66	309.8***	246.9***
	(92.97)	(125.7)	(82.83)	(113.4)	(144.1)	(190.7)	(49.28)	(40.46)
Obesity Rate	–87.43	–1.862	170.2**	226.4***	46.85	9.046	88.79*	107.6**
	(60.71)	(50.88)	(62.52)	(52.20)	(89.39)	(101.2)	(35.39)	(35.97)
Ln 1992 County Income	–108.8**	–32.72	–114.2**	–66.74+	–43.03*	–24.02	–20.57*	–9.921
	(39.72)	(44.92)	(33.01)	(34.63)	(17.95)	(21.56)	(8.920)	(9.875)
Chg Ln Cnty Income 92-19	–124.4***	–128.2***	–114.5***	–106.2***	–84.51*	–80.78*	–39.62***	–44.23***
	(28.13)	(26.72)	(20.07)	(20.19)	(36.47)	(34.04)	(9.003)	(9.000)
1992 College Share	148.8*	82.39	49.43	–46.80	273.4***	283.3***	–114.3***	–128.9***
	(64.56)	(66.26)	(65.52)	(57.34)	(34.84)	(37.37)	(29.71)	(24.22)
State Policy Index	–7.881		–11.30*		–3.733		–4.080*	
	(5.090)		(5.541)		(5.230)		(1.931)	
1992 Emp-Pop Ratio	34.21	–29.07	81.10*	69.01+	–106.2	–168.6*	15.99	24.96
	(54.00)	(58.01)	(31.38)	(37.23)	(63.90)	(72.29)	(14.98)	(16.40)
Chg Emp-Pop 92-19	129.2	92.59	109.5	63.53	–214.1*	–220.9*	–1.625	–0.249
	(101.5)	(99.17)	(71.24)	(72.93)	(93.08)	(103.2)	(27.07)	(23.93)
1992 Manufacturing Share	75.13+	44.37	–15.87	–61.16	91.29*	67.43+	–64.73***	–52.47***
	(39.77)	(50.61)	(36.21)	(46.57)	(38.89)	(37.02)	(14.63)	(14.45)
NCHS Urban-Rural Code=2	59.77***	53.83**	21.92*	10.80	28.26***	16.05+	–2.399	–3.792
	(15.00)	(18.64)	(10.65)	(9.389)	(6.179)	(8.322)	(3.266)	(3.766)
NCHS Urban-Rural Code=3	75.29***	74.66***	34.21**	25.45*	32.28***	27.34**	4.606	3.392
	(12.52)	(14.11)	(10.44)	(10.55)	(7.632)	(8.421)	(3.825)	(3.971)
NCHS Urban-Rural Code=4	89.27***	81.70***	38.00*	26.06+	46.13***	34.82***	9.823*	10.58*
	(14.80)	(15.61)	(14.94)	(13.75)	(7.671)	(8.720)	(4.222)	(4.274)
NCHS Urban-Rural Code=5	105.4***	96.92***	50.43***	37.34**	38.27***	29.43**	11.54+	11.95*
	(12.34)	(13.77)	(12.84)	(11.83)	(7.973)	(8.909)	(5.746)	(5.241)
NCHS Urban-Rural Code=6	97.15***	91.46***	49.03***	38.20**	33.81***	22.94*	6.096	8.189
	(13.50)	(15.74)	(13.44)	(12.87)	(9.399)	(10.79)	(5.702)	(5.881)
Constant	922.8*	216.9	1110.3**	739.4*	348.9+	203.2	348.1***	268.2*
	(412.3)	(446.4)	(348.1)	(363.5)	(183.3)	(194.9)	(95.12)	(103.1)
R^2	0.471	0.553	0.588	0.654	0.199	0.258	0.800	0.815
Observations	1956	1956	1956	1956	1953	1953	1953	1953

Notes: Regressions are weighted by 1992 population. The dependent variable in each regression is the change in the education-specific mortality rate per 100k population between 1992 and 2019. Columns 2, 4, 6 and 8 include state fixed effects. Standard errors are clustered by state. + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 4. Baseline Mortality-Change Regressions

	(1) Non-College	(2) College
1992 Mortality Rate	–0.452***	–0.864***
	(0.0404)	(0.0146)
Smoking Rate	858.1***	357.1***
	(78.86)	(55.95)
Obesity Rate	146.5*	107.0**
	(67.77)	(34.29)
Ln 1992 County Income	–61.27**	–51.38***
	(22.03)	(9.372)
Chg Ln Cnty Income 92-19	–90.36***	–43.30***
	(14.85)	(7.346)
State Policy Index	–14.41*	–4.043+
	(5.628)	(2.399)
NCHS Urban-Rural Code=2	9.611	–4.442+
	(9.460)	(2.572)
NCHS Urban-Rural Code=3	27.25**	2.399
	(9.615)	(3.902)
NCHS Urban-Rural Code=4	31.81*	8.890+
	(14.57)	(4.767)
NCHS Urban-Rural Code=5	42.13**	10.42+
	(12.93)	(5.898)
NCHS Urban-Rural Code=6	40.26**	2.860
	(13.52)	(6.451)
Constant	625.7*	630.6***
	(253.5)	(105.4)
R^2	0.599	0.745
Observations	1956	1953

Notes: Regressions are weighted by 1992 population. The dependent variable in each regression is the change in the education-specific mortality rate per 100k population between 1992 and 2019. State fixed effects are not included. Standard errors are clustered by state. + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 5. Means and Standard Deviations of Counterfactual Predicted Values from Baseline Mortality-Change Regressions

Scenario	Non-College		College		
	Mean (1)	Std. Dev. (2)	Mean (3)	Std. Dev. (4)	Δ Gap (5)=(1)-(3)
Actual Changes in Non-College Mortality Rates	–5.0	110.8	–111.8	86.1	106.8
Predicted Values from Baseline Mortality-Change Regressions	–5.0	85.8	–111.8	74.3	106.8
C1: Change All Non-College Smoking Rates to Mean College Smoking Rate	–138.0	59.5	–111.8	74.3	–26.3
C2: Change All Non-College Obesity Rates to Mean College Obesity Rate	–20.4	83.8	–111.8	74.3	91.4
C3: Assign All Counties the 90th percentile of 1992 Income Level	–23.3	80.3	–127.1	73.6	103.8
C4: Assign All Counties the 90th percentile of Change in Income: 1992–2019	–17.0	81.9	–117.5	73.5	100.5
C5: Assign All Counties the 90th percentile of State Policy	–25.7	80.5	–117.6	74.1	91.9

Note: Counterfactuals are based on baseline mortality-change regressions in Table 4.

Appendix

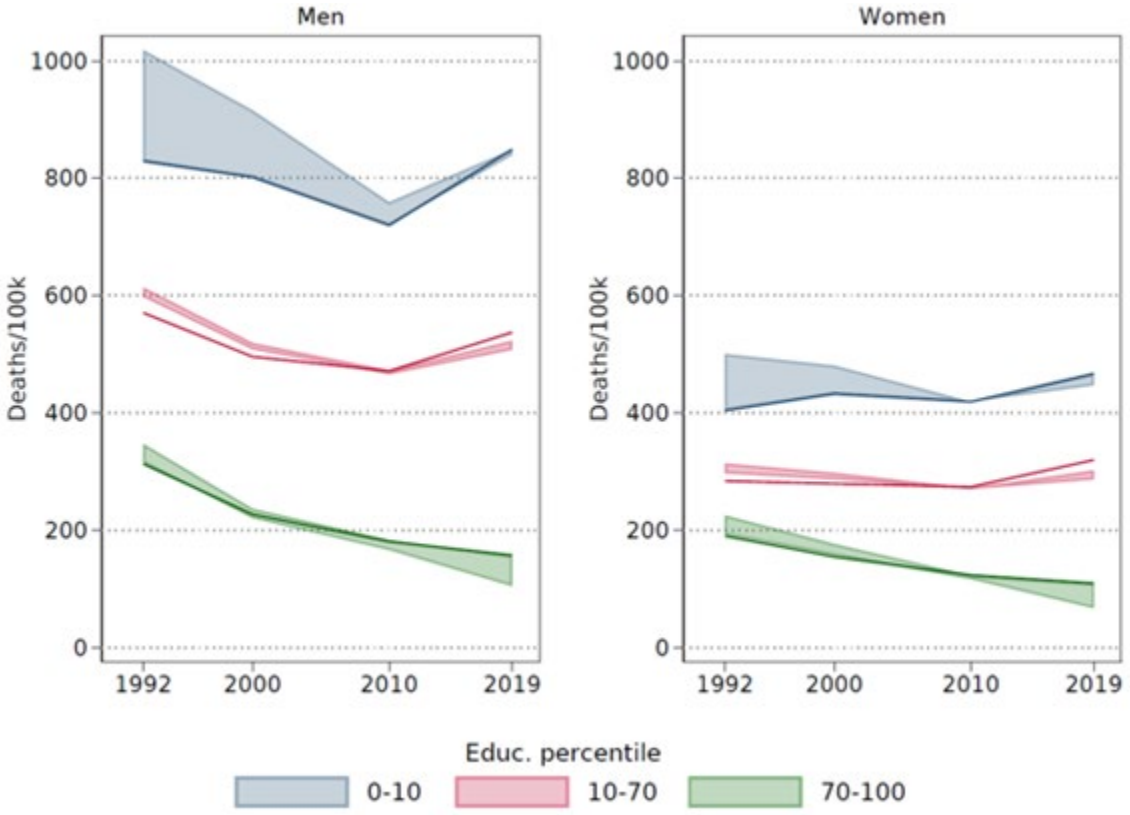


Figure A.1: Midlife Mortality by Education and year with Novosad et al. (2022) Selection Correction

The selection model in Novosad et al. (2022) develops bounds on the education-mortality gradient that are anchored on the 10th and 70th percentiles; assuming monotonicity they are able to estimate upper and lower bounds on the percentile estimates with wider bounds when the actual rate of (say) high school dropouts is different from the 10th percentile (as for men in 1992).

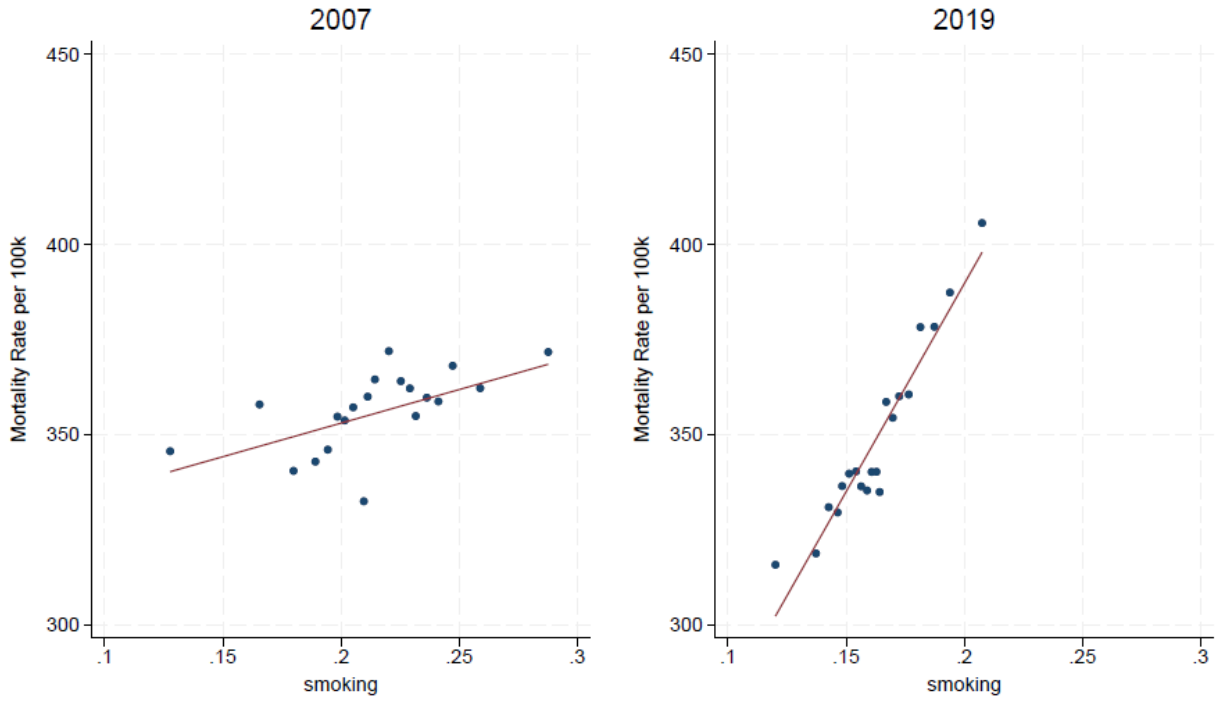
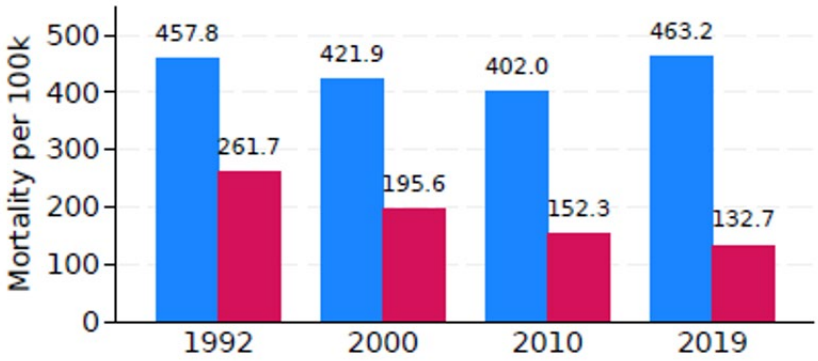


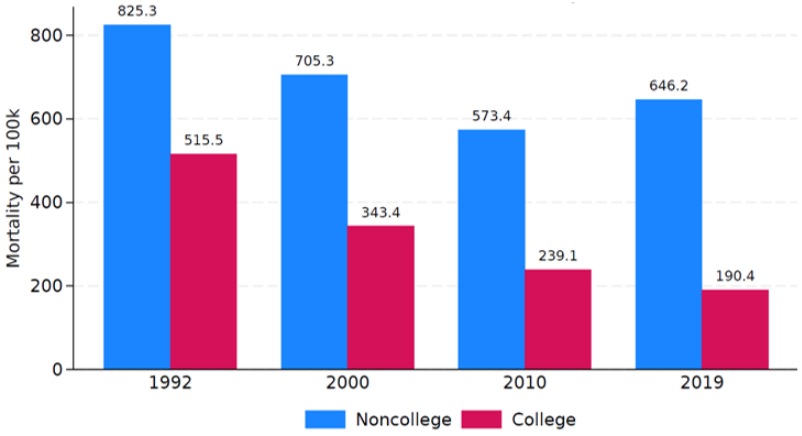
Figure A.2: Binscatter graph of county level smoking rates (all education groups) and midlife mortality rates: Conditional on Other Determinants of Health

Notes: Data from County Health Rankings; mortality rates from NCHS. Covariates which are controlled for include obesity, fair-poor health, teen births, low birthweight, rate of fatal crashes, murder rate, fraction uninsured, preventable hospital admissions, child poverty, fraction single parents, and rural status.

Panel A: Midlife Mortality (Entire Sample) by Education



Panel B: Black Midlife Mortality by Education



Panel C: Hispanic Midlife Mortality by Education

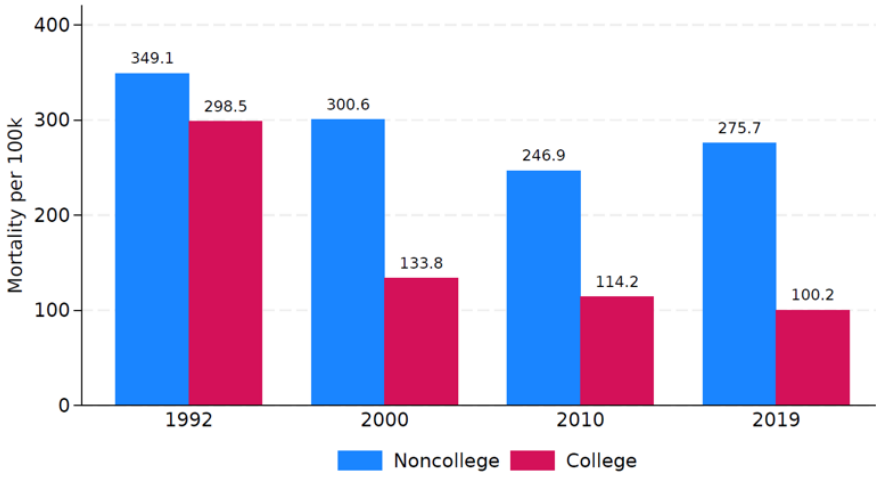
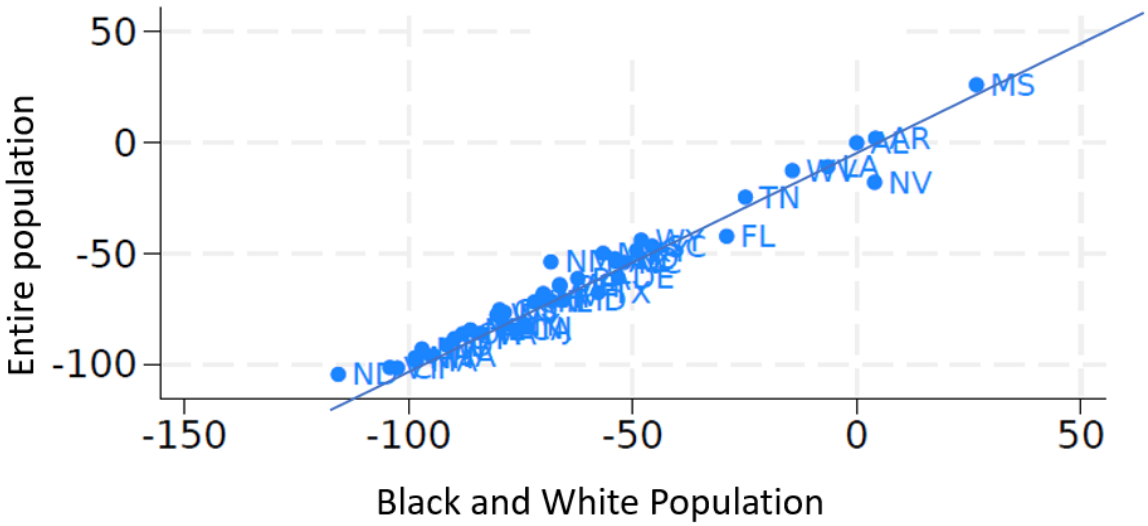


Figure A.3: Black and Hispanic Midlife Mortality by Education and Year



Mortality rate is deaths/100,000 among age 25-64 year-olds. It is age-adjusted to year 2000 population.

Figure A.4: State-Level Mortality for Black and White Populations Compared to the Entire Population

The mortality estimates are centered on Alaska. This graph shows that the mortality patterns for largely non-immigrants (Black and White) are nearly identical to the overall population; thus our results are unlikely to be biased by states with large immigrant populations with low mortality rates.