American Economic Journal: Applied Economics 2022, 14(4): 1–34
https://doi.org/10.1257/app.20190297

Mortality Change among Less Educated Americans†

By Paul Novosad, Charlie Rafkin, and Sam Asher*

Measurements of mortality change among less educated Americans can be biased because the least educated groups (e.g., dropouts) become smaller and more negatively selected over time. We show that mortality changes at constant education percentiles *can be bounded with minimal assumptions. Middle-age mortality increases among non-Hispanic Whites from 1992 to 2018 are driven almost entirely by the bottom 10 percent of the education distribution. Drivers of mortality change differ substantially across groups. Deaths of despair explain most of the mortality change among young non-Hispanic Whites, but less among older Whites and non-Hispanic Blacks. Our bounds are applicable in many other contexts.* (*JEL* I12, I26, J15)

Mortality rates among non-Hispanic Whites without college degrees have increased substantially over the last 20 years (Meara, Richards, and Cutler 2008; Cutler and Lleras-Muney 2010b; Cutler et al. 2011; Olshansky et al. 2012; Case and Deaton 2015, 2017). While widely publicized, this fact by itself is difficult to interpret because overall education levels have also risen during this time period; the share of 50–54-year-old women without a college degree, for example, was 63 percent in 1992 and 36 percent in 2018. The average person without a college degree occupies a lower position in both the educational and the socioeconomic distribution today than in the past. It is therefore not necessarily surprising that people at a fixed low level of education are less healthy today compared with those at the same level in earlier decades (see Figure 1). If education levels are rising, it is theoretically possible for the mortality rate to be lower at every percentile in the education distribution, but to be higher at every education level.[1]

* Novosad: Dartmouth College (email: paul.novosad@dartmouth.edu); Rafkin: MIT (email: crafkin@mit.edu); Asher: Imperial College London (email: sam.asher@ic.ac.uk). Ilyana Kuziemko was coeditor for this article. Novosad and Rafkin share lead authorship. Asher was affiliated with Johns Hopkins University at the time of acceptance. We are grateful for feedback on earlier versions of this paper from Alberto Abadie, Patty Anderson, Leila Agha, John Beshears, Emily Blanchard, Raj Chetty, James Choi, Eric Edmonds, Shahe Emran, Jim Feyrer, Francisco Ferreira, Amy Finkelstein, Nate Hilger, David Laibson, JoAnna Leyenaar, Ethan Ligon, Erzo Luttmer, Ellen Meara, Nina Pavnik, Jim Poterba, Bruce Sacerdote, Na'ama Shenhav, Forhad Shilpi, Jon Skinner, Chris Snyder, Gary Solon, Bob Staiger, Doug Staiger, Michael Stepner, and Elie Tamer. Toby Lunt and Ryu Matsuura provided excellent research assistance. All errors are our own. The authors have no competing interests. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant no. 1122374. This manuscript includes material that previously circulated in a working paper titled "Getting Signal from Interval Data: Theory and Applications to Mortality and Intergenerational Mobility."

† Go to https://doi.org/10.1257/app.20190297 to visit the article page for additional materials and author disclosure statements or to comment in the online discussion forum.

[1] A similar phenomenon is described by the well-known college swipe, "If the worst student at college X went to (inferior college) Y, it would raise the average intelligence of both schools."

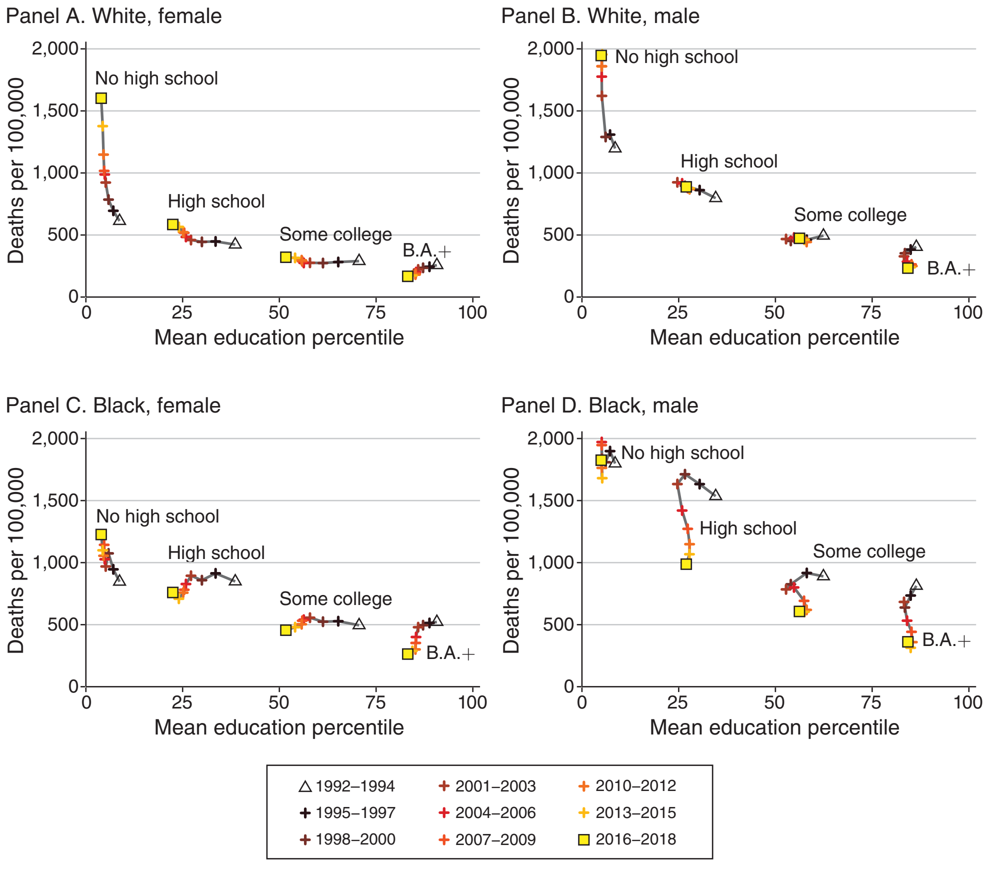


Figure 1. Mortality versus Education Rank, Age 50–54, 1992–1994 to 2016–2018

Notes: "White" refers to non-Hispanic White and "Black" to non-Hispanic Black. The figure shows change in mortality and average education rank for individuals aged 50–54 at different levels of education, from 1992–1994 to 2016–2018. Each point represents the average number of deaths per 100,000 people among people with one of four levels of education: No High School, High School, Some College, and a B.A. or Higher. The x coordinate of each point represents the average education percentile among people with the given level of educational completion. For example, a 50-year-old White woman with a high school education was at the thirty-ninth percentile of the education distribution in 1992–1994 and at the twenty-second percentile in 2016–2018.

Sources: ACS, CPS, NCHS

There are three possible interpretations of rising mortality among non-Hispanic Whites without college degrees. Each has a substantially different policy implication. First, this result could be nothing more than an artifact of shifts in the education distribution, with no changes in the underlying relationship between education percentile and mortality. Second, mortality could be rising uniformly among individuals in the bottom half of the education distribution. Third, mortality could be rising substantially at the very bottom of the education distribution, with fewer changes or even improvements in the percentiles reflecting high school graduates. In this paper, we develop new methods to distinguish between these scenarios and show that the third interpretation is the one most supported by the evidence.

This selection bias in estimates of mortality change at fixed education levels has been a major barrier to the study of disparities in death rates, not least because education is one of the only measures of socioeconomic status that is recorded in publicly available vital statistics data. Some researchers have argued that the bias is so large that estimates of mortality change by education level are effectively meaningless (Dowd and Hamoudi 2014, Bound et al. 2015, Currie 2018). Other researchers have limited analysis to population subsets where education has not substantially changed (Case and Deaton 2015, 2017). Similar challenges arise in studies of educational gradients in fertility, birth outcomes, and disability, as well as in the study of assortative mating and intergenerational mobility (Cutler and Lleras-Muney 2010a; Aizer and Currie 2014; Greenwood et al. 2014; Bertrand et al. 2021; Asher, Novosad, and Rafkin 2022).

In principle, the selection bias can be addressed by studying mortality in fixed percentile ranges of the education distribution, for example, in the bottom 10 percent. This approach holds constant the size and relative rank of each education bin over time. While calculating mortality in fixed education percentiles has been suggested before (Bound et al. 2015), doing so is not trivial, because education levels are inherently lumpy, especially as reported in standard mortality datasets. For example, if education is bottom-coded at the twentieth percentile (as in 1992, where 20 percent of women in some cohorts are high school dropouts), the mortality rate at the tenth education percentile cannot be point-estimated without strong assumptions.

This paper introduces a new partial identification method that addresses this concern. We show that outcomes conditional on arbitrary education ranks can at best be bounded. We treat the measurement of mortality y at a given education rank x as an interval data problem, where the education rank is only observed to lie within some bin $[x_k, x_{k+1}]$ of the rank distribution. Extending the approach of Manski and Tamer (2002), we show that $E(y|x \in [a,b])$ can be sharply and meaningfully bounded for arbitrary values of a and b.[2] Our approach requires only two assumptions. First, we assume that there exists a latent education rank, which is only coarsely observed in the education data; this assumption follows directly from a standard human capital model. Second, we assume that the mortality rate is weakly decreasing in the latent education rank; this assumption is supported by theory and empirical evidence. We show that bounds can be further tightened by disallowing kinks or discrete jumps in the education-rank function; this third assumption also makes it possible to loosen the monotonicity assumption.[3]

Using this partial identification approach, we document changes in mortality from 1992–1994 to 2016–2018 among the US population aged 25–69, in constant education percentile bins. We focus in particular on two domains where researchers have noted deteriorating outcomes: (i) mortality change in the bottom half of the education distribution; and (ii) changes in deaths from poisoning, suicide

[2] Our key innovation to the setup in Manski and Tamer (2002) is that we develop bounds on $E(y|x \in [a,b])$ when the latent distribution of x is known. In the case of education rank data, the latent distribution is uniform by construction. We also develop general bounds on $E(y|x \in [a,b])$ when x is not necessarily uniform; these bounds may be useful in other cases, e.g., top-coded income data that are assumed to follow a Pareto distribution.

[3] A curvature constraint is not central to our results. We show in online Appendix Section D that our central results hold using the first two assumptions alone; however, adding plausible structural assumptions yields tighter bounds. Allowing discrete jumps or kinks at major education boundaries (like high school or college completion) also has no material effect on the results.

and chronic liver disease, described by earlier researchers as "deaths of despair" (Case and Deaton 2015, 2017).

We have three primary findings. First, among middle-aged non-Hispanic White (hereafter referred to as White) men and women, the group most widely discussed in the recent literature, mortality increases are driven almost entirely by the bottom 10 percent of the own-gender education distribution (the part of the distribution represented by high school dropouts in 2018).[4] From 1992–1994 to 2016–2018, age-adjusted mortality for Whites in the least educated 10 percent has risen by 69–112 percent for women and 47–67 percent for men (2.2–3.2 percent and 1.6–2.2 percent per year, respectively). Mortality in percentiles 10–45 (approximately high school completers in 2018) is rising for both White men and women under age 50, but is flat or declining at higher ages where most deaths occur. The mortality increases described by Case and Deaton (2015, 2017) are thus both more severe and more focused in a narrow population subgroup than has previously been recognized.

Second, non-Hispanic Blacks have experienced large improvements in mortality in all education groups *except* for the least educated 10 percent. In the least educated 10 percent, Black women's mortality has risen 9–17 percent from 1992–1994 to 2016–2018, while Black men's mortality change has been very close to zero. This has led to a substantial convergence between Black and White outcomes at the bottom of the education distribution. Conditional upon being in the least educated 10 percent of the national distribution, White men over the age of 50 in 2016–2018 have higher mortality than similarly-aged Black men. White women in the least educated 10 percent have higher mortality than similarly-educated Black women in 2016–2018. In nearly all other education-age groups, White men and women have lower mortality than Black men and women.

A single proximate cause cannot explain these divergent death rates. The change in deaths from despair, which has been widely discussed in prior research and in the media, accounts for a large share of mortality increases for young Whites, but a very small share of rising mortality among older Whites and very little of the divergent mortality rates of Blacks. Further, deaths of despair have increased more uniformly across the education distribution than deaths from other causes. The least educated middle-aged Whites, in particular, are now at higher risk of dying from cancer, heart diseases and respiratory diseases, among other causes, even as mortality from these causes has declined sharply for those outside of the bottom 10 percent. Note that earlier unadjusted estimates of these mortality changes were particularly difficult to interpret for women, for whom education has risen considerably more than among men, creating a larger possible selection bias.

A long prior literature relies upon education as a proxy of socioeconomic status to study mortality change, both because of its wide availability in the data, and because it is a marker of permanent rather than transitory socioeconomic status. Olshansky et al. (2012) noted rising mortality rates among high school dropouts

[4] Throughout the paper, we take "Whites in the bottom 10 percent" to mean "Whites who are in the bottom 10 percent of the own-gender national education distribution" (where ranks pool across all races). We discuss in Section III why this is a more useful categorization than "Whites in the bottom 10 percent of the education distribution of Whites." Nevertheless, when we employ the latter definition, we find results that are broadly consistent with our findings here (online Appendix Figure D2).

from 1990 to 2008, but this work attracted debate because it did not adjust for the substantial increase in the negative selection associated with being a dropout over the sample period. Case and Deaton (2015, 2017) justified ignoring the selection bias in mortality change by focusing on population subgroups for whom education levels had not changed substantially; however, they did not look specifically at outcomes among high school dropouts exactly because of the selection bias addressed in our paper. Meara, Richards, and Cutler (2008); Bound et al. (2015); Hendi (2015); and Leive and Ruhm (2021) use an adjustment for selection bias that is implicitly based on stricter (and in our view, less plausible) assumptions that underestimates mortality change at the bottom of the distribution.[5]

Our finding of dramatically rising mortality in the bottom 10 percent broadly supports the earlier selection-unadjusted findings of Olshansky et al. (2012) and Sasson (2016): the mortality increases at the bottom of the education distribution prove to be large, even after removing substantial selection bias. We find a larger decline at bottom of the distribution than Hendi (2015, 2017), both because of our approach to selection bias, and because we use the much larger vital statistics data which are better suited to detect mortality changes in small groups like high school dropouts (Sasson 2017).

Several other recent papers document the relationship between socioeconomic status and mortality. Currie and Schwandt (2016a, b) study differences in mortality across counties, finding, like us, that changes in mortality inequality are highly heterogeneous across age, race, and place. They show that mortality inequality across space is falling between Blacks and Whites and among younger individuals (especially children), but rising among older adults. They also document dramatic declines in mortality among Black men. Our findings confirm this result among Black men in the most educated 90 percent. But we find that middle-aged Black men in the least educated 10 percent have experienced mortality increases, although these increases are minor compared with similarly educated Whites.

Chetty et al. (2016) use deaths as reported in tax records to describe changes in mortality throughout the income distribution. While the study of mortality using tax records is an important innovation, vital statistics are likely to remain valuable as sources of information on mortality because they record cause of death in detail and because they are publicly available. Our work makes it possible to use education as a marker of socioeconomic status in the vital statistics data, which is important given that so few other predictors of socioeconomic status are recorded.

In a related approach, Goldring, Lange, and Richards-Shubik (2016) derive a one-tailed statistical test to examine whether the mortality gradient in education is changing over time. Like us, they assume that: (i) there exists a latent education rank distribution; and (ii) mortality is monotonically decreasing in the latent education rank. They conclude that the education gradient is getting steeper (as do we), but their approach does not generate estimates of mortality change.

In addition to the empirical findings, this paper introduces a new methodology to tighten the CEF bounds of Manski and Tamer (2002) in contexts with known

[5] We compare their approach to ours in Section II and online Appendix Section C.3.

conditioning distributions (like ranks, which are uniform by construction). In the simplest case without curvature constraints, we provide analytical bounds that are readily calculated. We also provide a numerical framework for tightening bounds with arbitrary structural constraints, such as the curvature constraint we employ in our main results.[6] Our methodology may be of use in interval-censoring contexts as diverse as bond ratings, top-coded incomes, and Likert scales. It is especially useful when studying education, because education data remain interval-censored in rank terms even as granular administrative data become available for other variables, such as income. As a result, these bounds may be applied whenever the researcher wishes to study trends in a given outcome over time by education group.

We have posted both unadjusted and constant-percentile mortality estimates for all ages and groups with the manuscript, which we hope will be useful for other researchers interested in studying US mortality change. Code to calculate bounds on mortality in constant percentile groups given raw education data is also posted online.[7]

I. Data Sources

We briefly summarize the data construction process and provide more details in online Appendix Section B.

Death records from 1992–2018 were obtained from the US National Vital Statistics System of the NCHS (National Center for Health Statistics 1993–2019). Mortality rates (deaths per 100,000 people) were obtained by dividing the number of deaths in each age, race, gender and education cell by the population total from the Current Population Survey (CPS; United States Census Bureau 1993–2019b). We code ages in 5-year bins to mitigate bias from changing age within bins over time (Gelman and Auerbach 2016; Case and Deaton 2017). Education could be consistently matched across datasets in four groups: (i) less than a high school degree, (ii) high school degree/GED, (iii) some college, and (iv) a bachelor's degree or more.[8] Annual estimates were pooled into three-year bins. Following earlier work, estimates are presented separately for men and women, and for non-Hispanic Blacks and Whites. Results are not shown for Hispanics, because their higher in- and out-migration over the sample period make mortality change among Hispanics more difficult to interpret (Markides and Eschbach 2005). Mortality rates closely match those in other recent studies (Case and Deaton 2015, 2017).

Causes of death were partitioned into the following subgroups: cancers, heart diseases, deaths of despair, injuries, and other diseases. Deaths of despair are deaths from poisoning, suicide, and alcoholic liver diseases and cirrhosis (Kochanek, Arias, and Bastian 2016; Case and Deaton 2017); we exclude suicides from injuries. More detail on the distribution of deaths is reported in online Appendix Table A1.

[6] Our method is also easily generalized to measure other conditional parameters, like a median or other percentile of the outcome distribution.

[7] Stata and Matlab code for the bounding algorithm and replication code for this paper is available on GitHub at https://github.com/devdatalab/paper-nra-mortality.

[8] We aggregate the small share of people who attain no high-school education with people who attain some high-school education but do not drop out. See online Appendix Section B for details.

The strength of the NCHS data is its large number of observations and precision. The weakness is that mortality rates can only be measured in NCHS by dividing deaths by the population in a different dataset, creating risk of bias if the datasets have different biases in covariate measurement. We address many of these potential biases in the robustness section of the paper, but concerns cannot be ruled out entirely. The best alternative measures of US mortality come from the National Health Interview Survey (NHIS), which matches individuals to mortality records, eliminating division bias. The weakness of the NHIS is that it has too small a sample to measure mortality change among less educated groups with much precision, as we show in online Appendix Section D.6.

II. Methods: Bounding Mortality in Constant Education Percentile Bins

The Selection Problem.—When the level of education in the population rises, individuals at each level of education mechanically occupy a lower set of ranks in the education distribution. For example, among 50–54-year-old women, dropouts were approximately the bottom 19 percent in 1992 and the bottom 8 percent in 2018. The people who drop out of high school in 2018 may be more negatively selected than in 1992. If mortality rises among dropouts from 1992 to 2018, one might worry that such negative selection, rather than worsening health outcomes, drives the mortality increase. In fact, mortality can rise at each *level* of education even if the mortality rate in the population is constant. This statistical paradox is known as the Will Rogers Phenomenon or "stage migration" in the medical literature (Feinstein, Sosin, and Wells 1985).

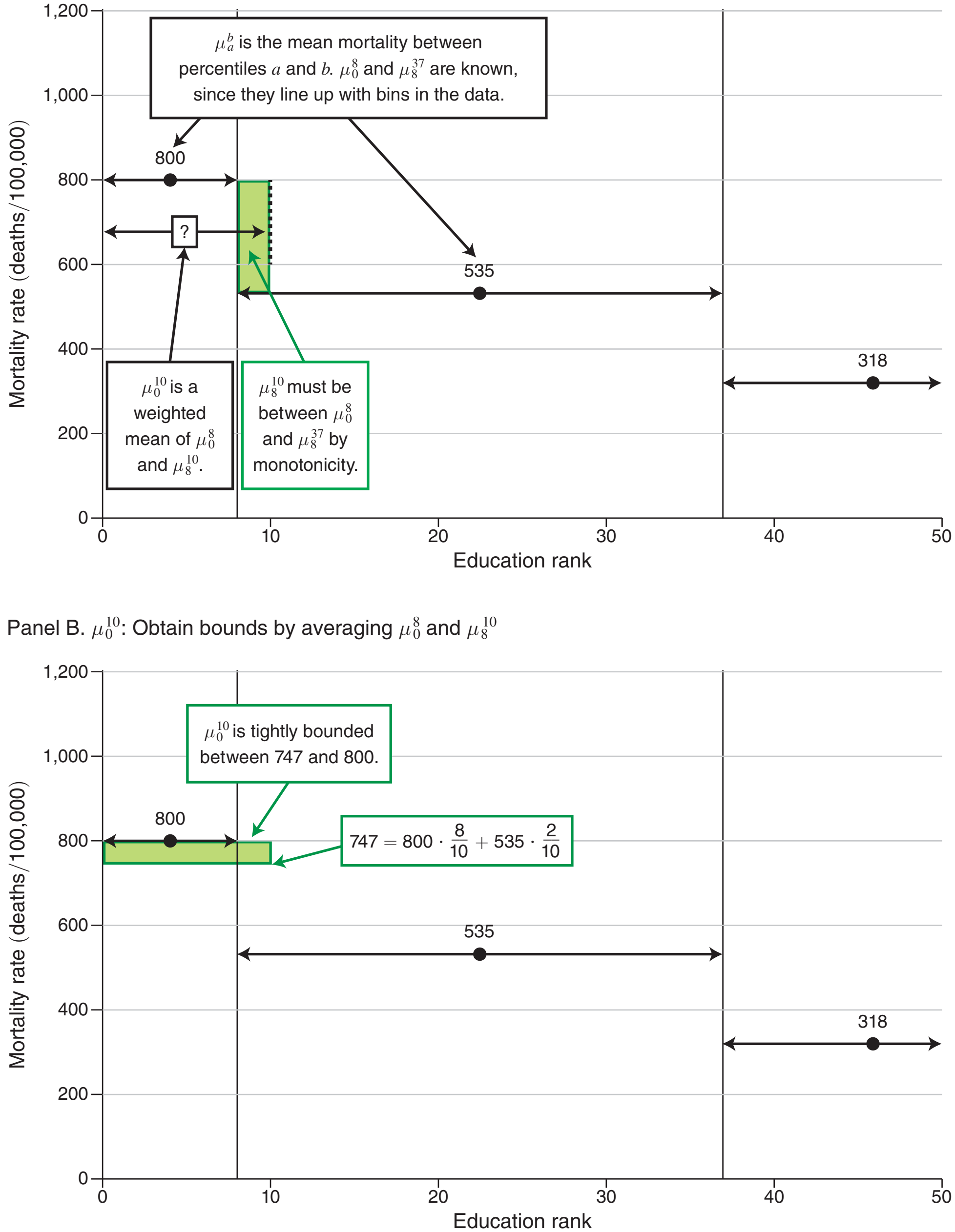
One can resolve this problem by measuring mortality within a constant range of education percentiles (e.g., percentiles 0–10 or 0–50) instead of at fixed education levels (Bound et al. 2015). Using education ranks holds the relative size of the group constant over time; the bottom 10 percent is no more negatively selected (in relative terms) in 1992 than in 2018. But it is not trivial to implement this solution because education is typically observed in coarse categories that cover many percentiles. How does one calculate mortality among the least educated 10 percent, if the bottom 15 percent are bottom-coded as high school dropouts?

We treat this as an interval data problem, where the latent education rank is only observed to lie within a set of coarse bins. We present a method that bounds the conditional expectation of mortality at a given percentile and in percentile ranges (e.g., average mortality rates in percentiles 0–10). We introduce and discuss these new bounds in the context of our empirical application, but they are valid in many other contexts with interval-censored conditioning variables.

We first describe the method intuitively and then formalize it. Figure 2 presents a graphical example, continuing to focus on women ages 50–54. For these women, mortality in 2016–2018 is known to be 800 deaths per 100,000 in percentiles 0–8 (high school dropouts) and 535 deaths in percentiles 8–37 (high school completers). Suppose that we wish to bound the mortality rate in percentiles 0–10 (panel A). Our key assumption, formalized below, is that mortality is weakly decreasing in the latent education rank.

Panel A. μ_0^{10}: Observe that μ_0^8 is point identified and μ_8^{10} is bounded between 535 and 800

Panel B. μ_0^{10}: Obtain bounds by averaging μ_0^8 and μ_8^{10}

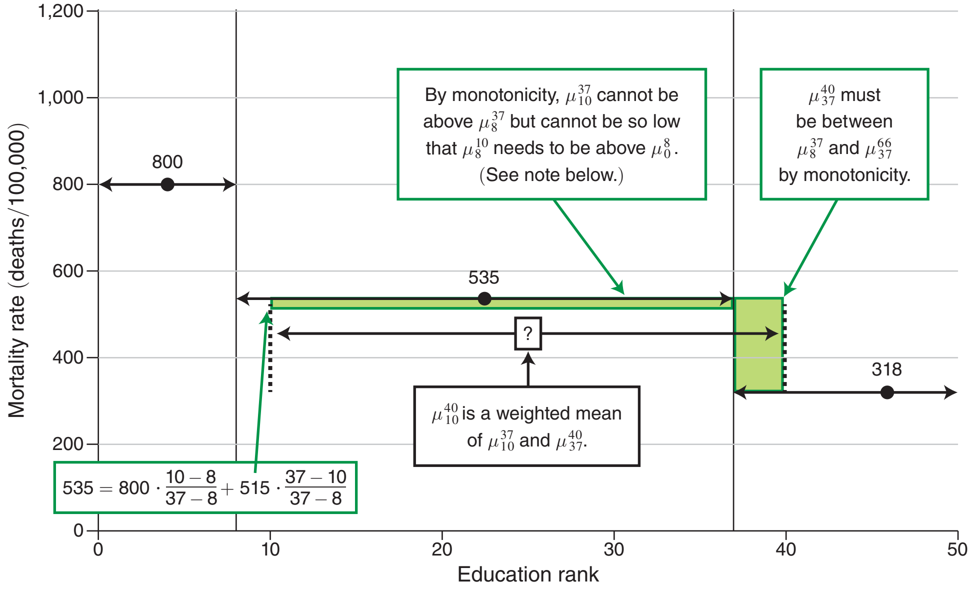


(*continued*)

Figure 2. Calculating the CEF of Mortality Given Education Rank

Define μ_a^b as the average mortality between ranks a and b. Mortality in percentiles 0–10 (i.e., μ_0^{10}) is a weighted mean of mortality in percentiles 0–8 $\left(\mu_0^8\right)$, which is known, and mortality in percentiles 8–10 $\left(\mu_8^{10}\right)$, which is unknown. We can bound μ_8^{10} from above: it must be weakly lower than μ_0^8 $(= 800)$, or else monotonicity

Panel C. μ_{10}^{40}: Obtain the lowest possible value of μ_{10}^{37} (= 515)



Panel D. μ_{10}^{40}: Using the lowest value of μ_{10}^{37}, average with the lowest value of μ_{37}^{40} (= 318)

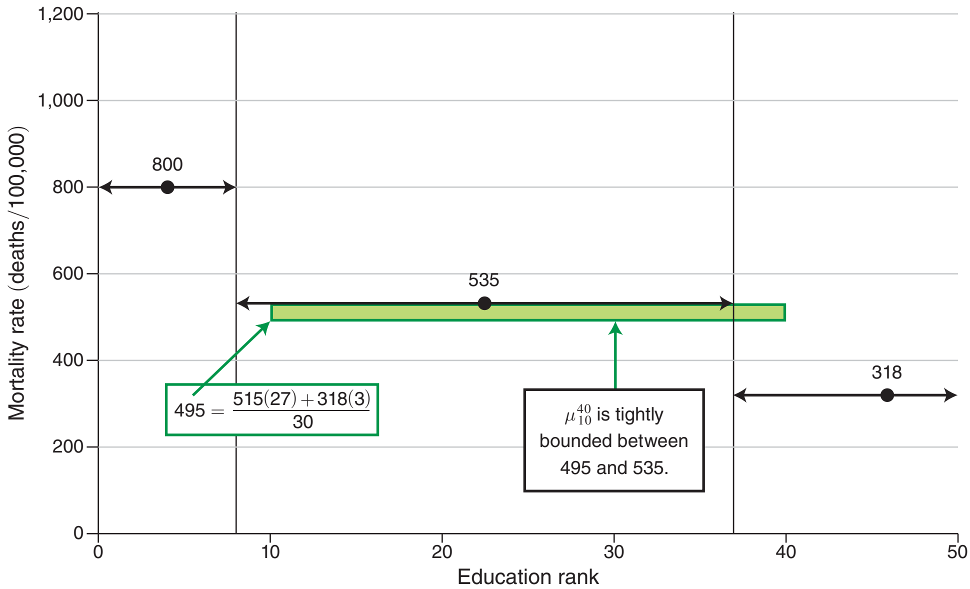


FIGURE 2. CALCULATING THE CEF OF MORTALITY GIVEN EDUCATION RANK (*continued*)

Notes: Figure 2 provides a graphical description of the calculation of the bounds on μ_a^b in two scenarios. The data are from women aged 50–54 in 2016–2018. The vertical lines show the rank bin boundaries for each education bin for this group. The points show the mean mortality in each bin. The first two panels show the calculation of μ_0^{10} and the following two panels show the calculation of μ_{10}^{40}. In panel C, the upper bound of μ_{10}^{37} cannot exceed the value of μ_8^{37}, because that would require $\mu_8^{10} > \mu_{10}^{37}$. The lower bound cannot be below 515, or else μ_8^{10} would need to be higher than μ_0^8 to fit the bin mean, thus violating monotonicity.

Source: NCHS

would be violated. We can also bound μ_8^{10} from below: μ_8^{10} must be weakly larger than μ_8^{37} $(= 535)$, or else monotonicity cannot hold between μ_8^{10} and μ_{10}^{37}. Taking the weighted mean of μ_0^8 and the bounds on μ_8^{10}, we can infer that mortality in the least educated 10 percent (μ_0^{10}) must be in the interval $[747, 800]$ (panel B of Figure 2).[9]

The previous example describes the simplest possible case. There are more complex cases where the bounds are *not* simply weighted averages of the adjacent bin means. Panels C and D of Figure 2 demonstrate such a case, in the setting of calculating μ_{10}^{40}. Here, the bounds also take into account the following additional logic. We discuss the case of forming lower bounds. Because μ_8^{37} is known $(= 535)$, μ_{10}^{37} must be relatively tightly bounded. The *highest* value of μ_8^{10} permits the *lowest* value of μ_{10}^{37} while still meeting the constraint that $\mu_8^{37} = 535$. Monotonicity implies that μ_8^{10} cannot exceed 800 (since that is the bin mean in ranks 0 to 8). Then, in order that $\mu_8^{37} = 535$, it must be the case that the lower bound for μ_{10}^{37} is 515.[10] Applying similar logic to obtain bounds on μ_{37}^{40}, we take the weighted mean of the lower bounds on μ_{10}^{37} and μ_{37}^{40} to obtain a lower bound on μ_{10}^{40} of 495. Similar logic gives upper bounds.

The key intuition behind these bounds is that mortality in an arbitrary rank range is a weighted mean of known and partially identified values. If the weight on known values is high or the partially identified values are tightly bounded, then mortality in the rank range can be tightly bounded. The bounds can be tightened further with additional structural assumptions if desired, as we demonstrate below.

A. *Assumptions*

Assumption 1 (Latent Education Ranks).—We have implicitly assumed in the narrative thus far that there exists a continuous latent education rank distribution, which is partitioned into discrete intervals by the observed education levels. For instance, if 10 percent of people are high school dropouts, then these people occupy distinct (continuous) ranks 0 through 10. This assumption arises out of a standard human capital investment model where schooling costs are convex and individual educational attainment is determined by individual-specific cost and benefit shifters (Card 1999), and is required for the selection adjustments used in the prior mortality literature (Hendi 2015; Cutler et al. 2011; Bound et al. 2015; Goldring, Lange, and Richards-Shubik 2016). A person who is highly ranked within her bin (for instance, the highest-ranked high school dropout) is a person who would have attained a higher level of education if the cost were only marginally lower or the benefit to them only marginally higher. Consider an example where two individuals *A* and *B* are identical except *A* has a lower discount rate, which raises her demand for education. *A* and *B* may obtain the same level of education because years of education are lumpy. However, *A* may be right at the margin of attaining a higher level of education and *B* may be right at the margin of attaining a lower level of education. If the

[9] The upper bound of μ_0^{10} is 800. The lower bound of μ_0^{10} is $0.8 \times 800 + 0.2 \times 535 = 747$.

[10] That value for the lower bound satisfies the equation $\frac{2}{29} \times 800 + \frac{27}{29} \times \mathit{LowerBound} = 535$.

discount rate also affects health-seeking behavior, then we would expect A to have lower mortality risk than B, even though their levels of education are the same.[11]

Assumption 2 (Monotonicity).—We assume that mortality rates are non-increasing in latent education percentile. This assumption is suggested by the standard human capital model above, in that many factors correlated with socioeconomic status are expected to raise educational attainment and improve health; the direct effect of education on health is also expected to be positive. This assumption has been made either implicitly or explicitly by other researchers attempting to control for the rank change problem that we address in this paper (Cutler et al. 2011; Bound et al. 2015; Hendi 2015; Goldring, Lange, and Richards-Shubik 2016). The assumption is supported by empirical evidence that mortality and health are consistently decreasing in levels and in years of education in the United States and Europe (Pappas et al. 1993; Mackenbach et al. 2003; Meara, Richards, and Cutler 2008; Cutler and Lleras-Muney 2010b). These papers provide evidence of monotonically decreasing mortality across education bins; our assumption further imposes that mortality is non-increasing in rank within education bins. Further corroborating evidence comes from Chetty et al. (2017), who show that mortality is monotonically decreasing in granular income ranks.[12]

Importantly, while we invoke this assumption in the main results, our qualitative findings are similar if we loosen this restriction or replace it with an alternative structural assumption (online Appendix Section D).

B. *Formalization of Bounds on the Interval-Censored CEF*

Our approach extends Manski and Tamer (2002), who provide bounds on an interval-censored CEF with an unknown distribution. We show that (i) the Manski and Tamer (2002) bounds can be improved upon substantially in our context by recognizing the distribution of the conditioning variable; and (ii) the bounds on the mean value of the CEF in some interval may be much tighter than the bounds on the CEF at a given point. Finally, we present a numerical framework that permits the inclusion of arbitrary structural assumptions which may further tighten the bounds. The reader who is not interested in the details of the formalization may skip to Section IID.

Consider random variables y and x. In our setting, the variable y is the binary variable indicating whether an individual survives (survival $= 1$, death $= 0$), and x is the latent education rank. Although we generalize x, in the setting where x corresponds to ranks, we can think of x as belonging to the interval $[0, 100]$, the set

[11] Note that we are *not* making causal claims about the relationship between education rank and mortality. Rather, like the prior literature, we use education as a proxy for socioeconomic status that is readily available in mortality data. Our exercise is analogous to measuring mortality at a given income percentile, which is understood to be a meaningful measure even though the income level at that percentile may change over time. If the education *level* has a causal effect on health (e.g., through knowledge gain), then we might expect survival to improve at education ranks which reflect higher levels of education in 2018 than in 1992; our framework allows for this possibility.

[12] We discuss some empirical exceptions to this general monotonicity in groups that we study in Section III.

of education ranks. Define the average survival rate $Y(x = i) = E(y|x = i)$ for particular latent education rank i.[13]

Assume that x is only observed to lie in one of K closed intervals that are non-overlapping (except at endpoints) and cover the distribution of x. Each interval (or "bin") is indexed by $k \in \{1, \ldots, K\}$ and we write that interval k is the set $[x_k, x_{k+1}]$. For instance, $[x_1, x_2]$ represents the set of education ranks corresponding to the lowest education level in the data.

Our goal is to estimate $Y(x = i)$ for some i (e.g., $E(y|x = 10)$ is the survival rate at the tenth percentile), or $E(y|x \in [a,b])$ (e.g., $E(y|x \in [0,10])$ is the average survival rate in the least educated 10 percent).[14] Define the expected value of y in bin k as

$$r_k := E(y|x \in [x_k, x_{k+1}]).$$

Thus, r_1 is the average survival rate for people in the lowest education bin, e.g., high-school dropouts.[15]

Because x represents education ranks, ranks are uniform by construction:[16]

(Condition U) $$x \sim U(0,100).$$

We formalize the monotonicity assumption:

(Assumption M) $E(y|x = i)$ is weakly increasing in i.

Restate the following assumptions from Manski and Tamer (2002):

(Assumption I) If x is observed to lie in bin k, then $P(x \in [x_k, x_{k+1}]) = 1$.

(Assumption MI) If x is observed to lie in bin k, then $E(y|x, x_k, x_{k+1}) = E(y|x)$.

Assumptions I and MI are regularity conditions about interval censoring. Assumption I yields that, if x is interval censored, it truly lies within its given bin, and assumption MI states that the fact of interval censoring yields no additional information about x.[17]

From Manski and Tamer (2002), we have:

(Manski-Tamer bounds) $$r_{k-1} \leq E(y|x) \leq r_{k+1}.$$

[13] For consistency with the literature (Manski and Tamer 2002), we frame the problem in terms of the survival rate, which is monotonically increasing in education rank, rather than the mortality rate, which is decreasing in rank.

[14] Note that a and b need not correspond to points x_k that demarcate bins.

[15] In the case of survival rates, let $r_0 = 0$ and $r_{K+1} = 1$; these are the upper and lower bounds for the well-defined survival probability.

[16] We label U as a "condition" rather than an "assumption" because it is guaranteed to hold with ranks.

[17] These always hold in our case, because all data are interval censored. We label them as assumptions for consistency with Manski and Tamer (2002).

Intuitively, with no information on the distribution of the conditioning variable, the CEF is sharply bounded by its mean value in the prior and subsequent bin. Recognizing the uniform distribution of ranks yields the following proposition.

PROPOSITION 1: *Let x be in bin k. Under Condition U and Assumptions M, I, and MI (Manski and Tamer 2002), and without additional information, the following bounds on $E(y|x)$ are sharp:*

$$\begin{cases} r_{k-1} \le E(y|x) \le \frac{1}{x_{k+1}-x}\big((x_{k+1}-x_k)\,r_k - (x-x_k)\,r_{k-1}\big), & x < x_k^* \\ \frac{1}{x-x_k}\big((x_{k+1}-x_k)\,r_k - (x_{k+1}-x)\,r_{k+1}\big) \le E(y|x) \le r_{k+1}, & x \ge x_k^* \end{cases}$$

where

$$x_k^* = \frac{x_{k+1}\,r_{k+1} - (x_{k+1}-x_k)\,r_k - x_k\,r_{k-1}}{r_{k+1}-r_{k-1}}.$$

We refer to these as NRA bounds; the proof is in online Appendix Section C.1. Figure 3 shows the Manski and Tamer (2002) bounds and the NRA bounds on the mortality function above describing women aged 50–54 in 2018. The NRA bounds, which use the distribution of the data, are substantially tighter than the Manski-Tamer bounds.

In online Appendix Section C.1, we generalize Proposition 1 to the case with an arbitrary (but known) conditioning distribution. This generalization may be useful in settings where variables are commonly modeled with parametric distributions. For instance, in a setting with interval-censored income, this method could be applied under the assumption of a log-normal or Pareto income distribution. Online Appendix Section C.1 describes an additional proposition providing analytical bounds on the average value of the CEF between percentiles a and b, which we call $\mu_a^b = E(y|x \in [a,b])$ for interval-censored x.

As demonstrated in Figure 2, bounds on percentile ranges can be very tight. An important case is given by $\mu_{a'}^{b'} = E(y|x \in [a',b'])$ where a' and b' are particular rank boundaries in the education data (i.e., they correspond to x_k for some k). In that case, $\mu_{a'}^{b'}$ can be point identified; it is exactly the value of r_k in the observed bin (the bin mean), or the weighted average of the bin means across the bins that a' and b' span. In contrast, $E(y|x = i)$ is not generically point identified at any value of i.

Table 1 presents an illustrative comparison of bounds on $E(y|x = i)$ and μ_a^b for different intervals. Bounds on μ_a^b are generally (but not universally) tighter than bounds on $E(y|x = i)$, and in some cases they are much tighter. Bounds on μ_a^b are tightest when a and b are close to bin boundaries in the data; we use this fact to select the objects of our analysis in the results below.[18]

[18] For instance, in our context, μ_0^{10} can be tightly bounded for most groups in most years, but μ_0^{25} cannot. This is a limitation of the information contained in the data; if an analyst views μ_0^{25} as a much more important object than μ_0^{10}, they can tighten bounds on the mortality function by making additional structural assumptions (see below).

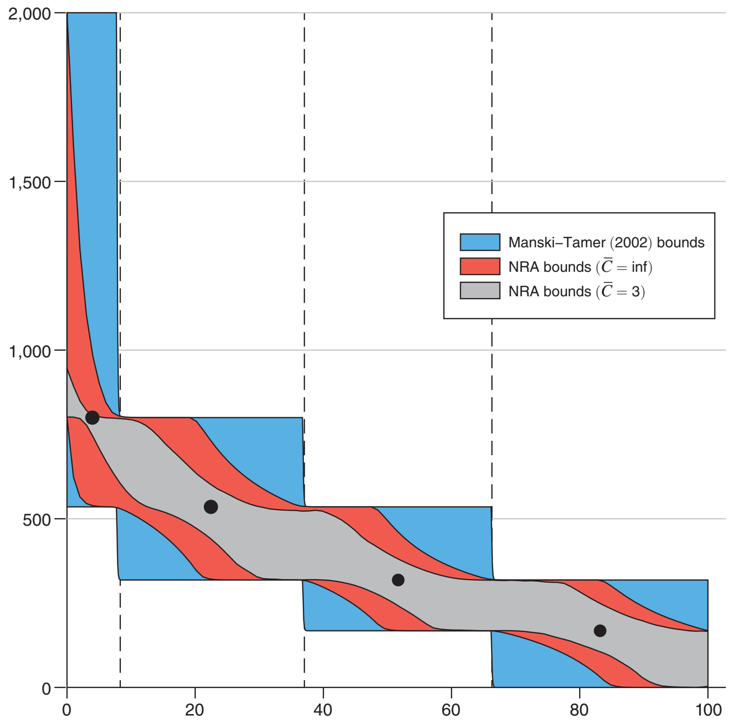


Figure 3. Change in Total Mortality of US Women, Age 50–54
Bounds on Conditional Expectation Functions

Notes: Figure 3 shows bounds on the conditional expectation function of mortality as a function of latent educational rank. The sample consists of US women aged 50–54; mortality is measured in deaths per 100,000 women. The points in the graph show the mean education rank and mortality in each year for individuals with (i) less than high school; (ii) high school; (iii) some college; and (iv) a B.A. or higher. The curves show the bounds on expected mortality at each latent parent rank ($E(y|x = i)$ in the text). The outer (blue) bounds are calculated following Manski and Tamer (2002). In the bottom bin, the blue bounds are truncated at 2,000 for visual clarity but actually extend to 100,000 (since the procedure cannot reject a mortality rate of 1 up to the first bin cut). The middle (red) bounds are calculated following our method with unrestricted curvature. The tightest (gray) bounds are calculated restricting the curvature to 3 percent of mean mortality across every percentile bin (2 times the largest curvature found in US income rank-rank data (Chetty et al. 2016)). Education rank is measured relative to the set of all women aged 50–54.

Source: NCHS

C. *A Numerical Framework for Arbitrary Constraints*

An advantage of the partial identification approach is that we can transparently leverage plausible structural assumptions to obtain tighter bounds. For cases where analytical solutions may be unavailable, we develop a numerical optimization framework for calculating $E(y|x)$ and other functions of the CEF. The numerical optimization generates identical results to Proposition 1 under Assumptions 1 and 2 only, but allows us to impose arbitrary additional structural constraints. In particular, we consider a constraint on the curvature of the CEF, which prevents large discrete changes in the level or slope of the CEF at a single point in the rank distribution.

Table 1—Bounds on Mortality throughout the Education Rank Distribution, 50–54-Year-Old Women, All Races

Statistic	Monotonicity only ($\bar{C} = \infty$)	Curvature only ($\bar{C} = 3$)	Monotonicity and curvature $\bar{C} = 3$
Panel A. 1992–1994			
$Y(x = 10)$: first quintile median	[455.9, 682.1]	[343.3, 793.8]	[456.1, 614.7]
$Y(x = 25)$: bottom half median	[427.7, 587.2]	[0.0, 1,163.1]	[436.9, 586.7]
$Y(x = 8)$: median $\leq$ high school (1992–94)	[455.9, 738.0]	[453.1, 726.2]	[485.9, 638.8]
$Y(x = 4)$: median $\leq$ high school (2016–18)	[455.9, 1,013.2]	[263.6, 972.2]	[573.1, 730.5]
μ_0^{20}: first quintile mean	[570.2, 587.2]	[539.0, 607.1]	[567.6, 586.2]
μ_0^{50}: bottom half mean	[501.6, 530.7]	[431.3, 582.1]	[504.3, 529.5]
μ_0^{16}: mean $\leq$ high school (1992–94)	[587.2, 598.7]	[585.3, 595.2]	[588.1, 595.1]
μ_0^{8}: mean $\leq$ high school (2016–18)	[587.2, 741.5]	[259.7, 1,041.2]	[587.5, 725.6]
Panel B. 2016–2018			
$Y(x = 10)$: first quintile median	[516.0, 799.9]	[284.9, 1,074.7]	[534.3, 799.8]
$Y(x = 25)$: bottom half median	[318.5, 685.4]	[208.2, 775.5]	[349.0, 600.5]
$Y(x = 8)$: median $\leq$ high school (1992–94)	[535.3, 799.9]	[417.0, 1,009.8]	[535.4, 799.8]
$Y(x = 4)$: median $\leq$ high school (2016–18)	[535.3, 1,046.3]	[737.3, 831.1]	[733.8, 816.3]
μ_0^{20}: first quintile mean	[640.1, 799.9]	[476.9, 903.0]	[641.2, 783.0]
μ_0^{50}: bottom half mean	[520.8, 570.1]	[455.5, 553.0]	[521.3, 551.2]
μ_0^{16}: mean $\leq$ high school (1992–94)	[666.2, 799.9]	[551.7, 952.2]	[667.7, 793.0]
μ_0^{8}: mean $\leq$ high school (2016–18)	[797.2, 799.9]	[799.9, 799.9]	[799.9, 799.9]

Notes: The table shows bounds on mortality in 1992–94 (panel A) and 2016–2018 (panel B) at various ranks or rank ranges in the education distribution. The notation $Y(x = i) = E(y|x = i)$ describes mortality at education percentile i, and μ_a^b describes average mortality between education percentiles a and b. $\overline{C}$ is the maximum percentage change in mortality function curvature allowed in any one percentile that does not correspond to an education bin boundary.

Sources: ACS, CPS, and NCHS

The numerical framework identifies bounds on a target parameter (e.g., $E(y|x = i)$ or $E(y|x \in [a,b])$) by identifying a pair of CEFs which respectively maximize and minimize that parameter, subject to matching the observed bin means in the data and meeting a set of restrictions, like a curvature constraint or monotonicity restriction. The framework is very flexible: arbitrary assumptions and outcome measures can be considered. The numerical optimization approach is described in online Appendix Section C.2.

Our main results below use the numerical optimization, imposing a structural assumption that the second derivative of the underlying CEF cannot exceed some constant $\bar{C}$. This curvature constraint prevents marginal changes in the latent education rank from being associated with discrete jumps or kinks in the CEF. The intuition for this assumption is that a marginal increase in education rank should not yield a discrete benefit for health.[19] Naturally, adding structural assumptions yields

[19] Alternatively, we can allow sharp kinks or jumps in the mortality function at ranks corresponding to major education bin boundaries, such as high school completion. This would be motivated by the possibility of sheepskin effects (Hungerford and Solon 1987), wherein completing high school (say) gives a discrete benefit for

(weakly) tighter bounds. Figure 3 shows the effect of adding the curvature constraint to bounds on mortality of the sample group above; Table 1 shows the effect on mortality at various points and ranges of the education distribution.

While we impose the curvature constraint in our primary results, our findings are robust to imposing a weaker curvature constraint or none at all (online Appendix Section D). This highlights a key advantage of the partial identification approach: we can clearly present how each assumption affects the bounds.

Our publicly available code permits the researcher to apply a flexible set of structural assumptions. Different assumptions may be more or less plausible in different applications; the code allows the researcher to adjust these assumptions depending on the context.

D. *Illustrating the Bias in Naïve Mortality Estimates*

Figure 4 presents an illustrative comparison of the difference between naïve estimates of mortality change (dots) at education levels with bounded estimates of mortality change (lines) at constant education percentiles, i.e., μ_a^b. The graph plots mortality rates for women aged 50–54 with less than a high-school degree (panel A) and with a high-school degree (panel B), showing percentage changes from 1992–1994 through 2016–2018.[20]

Panel A shows μ_0^{17} (i.e., the bottom 17 percent), and panel B shows μ_{17}^{60} (i.e., the women between ranks 17–60 in the own-gender education distribution). We choose these ranks because they are approximately the share of women in 1992–1994 with less than a high-school degree or exactly a high school degree, allowing the bounds to be very tight in the starting period.

The naïve estimate of mortality change for high school dropouts in this age-gender group is 36 percent. The comparable constant percentile estimates are bounded between 13 percent and 34 percent. The naïve estimate is thus unambiguously biased upward, but it is close to the upper bound on mortality change.

The bias in the high school completer group is substantial and reverses the sign of mortality change. Here, the naïve analysis suggests that mortality has risen by 17 percent from 1992–1994 to 2016–2018. Holding ranks fixed, however, we conclude that mortality has in fact *fallen* by 5–14 percent in percentiles corresponding to high school in 1992–1994.

Online Appendix Figure A2 shows similar graphs split by race and gender. The bias in the naïve estimates depends on the mortality-education gradient and the magnitude of the shift in bin boundaries. The examples above show that the bias can vary, even within the same age-gender group. There is no simple rule of thumb for adjusting naïve estimates, but our paper provides a measure that corrects for the shift in the education distribution.

health. Note that sheepskin effects would only affect mortality through the causal effect of education on health (including through any mediating channel like income); any part of the relationship between socioeconomic status and mortality that is not driven by education would not be affected by sheepskin effects. We show below that allowing for these effects does not widen the bounds appreciably.

[20] Note that women of all races/ethnicities are pooled in this example, so the point estimates of mortality change are not directly comparable to the race-specific estimates in the results section.

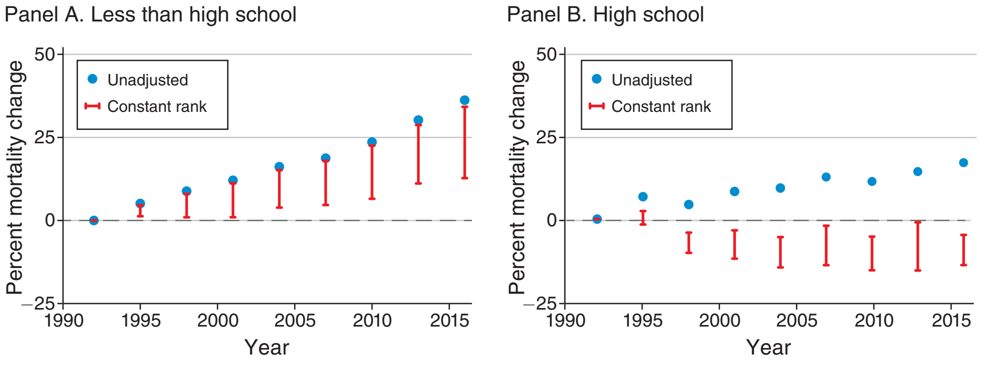


Figure 4. Changes in US Mortality, Women Age 50–54, 1992–1994 to 2016–2018: Naïve and Constant Rank Interval Estimates

Notes: Figure 4 shows mortality changes for 50–54-year-old women from 1992–1994 to 2016–2018 (all races combined), calculated under different methods. The points show unadjusted estimates for women at constant education levels—dropouts in panel A and high school graduates in panel B. Both of these population groups have shrunk as proportions of the population during the sample period. The vertical bars show bounds on mortality change in constant rank bins—ranks 0–17 in panel A and ranks 17–60 in panel B. These ranks are chosen because they are close to the share of women in 1992–1994 with less than a high school degree or exactly a high school degree, allowing the bounds to be very tight in the starting period.

E. *Comparison with Alternative Methods of Correcting for Selection Bias*

We are aware of three approaches that other researchers have taken when faced with this selection problem (other than ignoring it completely). First, researchers have reassigned individuals across bins at random to obtain constant percentile mortality estimates (Meara, Richards, and Cutler 2008; Bound et al. 2015; Hendi 2015; Leive and Ruhm 2021)—for instance, reassigning high school graduates to the dropout bin. This approach implicitly implies that the CEF of mortality given education rank is a highly discrete step function with constant mortality in each bin.

This function is unlikely to be a plausible description of reality for several reasons. First, this function implies enormous sheepskin effects in education, because it suggests that the individual who just barely completed high school (percentile 8.0 for 50–54-year-old women in 2016–2018) has far lower mortality than the individual who was right at the margin of completing high school but then dropped out (percentile 7.9). Second, it implies that the high school completer right at the margin of dropping out (percentile 8.0) has the same expected mortality as the high school completer right at the margin of going on to college (percentile 36.9). A standard human capital model rejects this function: individuals at the margin of completing college would be expected to have higher socioeconomic status than those at the margin of dropping out, and thus lower mortality risk.

Moreover, the implicit assumption of this step function can introduce downward bias in mortality change estimates. To construct mortality in a fixed percentile range at the bottom of the distribution, the researcher increasingly adds randomly-selected high schoolers over time (because the percentile threshold for high school is falling as education rises). But a randomly chosen high schooler likely has lower average mortality than a high schooler who is at the margin of dropping out. This approach thus leads to underestimates of mortality change in the least educated group.

This step function is the edge case in the NRA bounds, if we permit discrete jumps in the mortality CEF at education boundaries (as in online Appendix Section D). Our partial identification strategy thus accommodates the approaches of this prior work, but also permits a range of more realistic CEFs that would reject this function.

Second, Cutler et al. (2011) reassign individuals across education bins based on additional data. For instance, to reassign college completers to high school, they use a regression approach to reassign the individuals who would be most likely to have been high school completers in an earlier time period, based on age, region, marital status, and income. This approach is an improvement over the random reassignment used in the other papers above, but is not available in vital statistics data which report few markers of socioeconomic status other than race and education.

Finally, researchers have avoided the problem of selection bias by focusing on cohorts, subgroups or sample periods where education levels have not changed very much, such as Case and Deaton (2015, 2017). While valid, this strategy must constrain analysis to subgroups for whom education levels have not changed substantially. In many cases, researchers would like to study groups like high-school dropouts where relative ranks have changed over time. Our methods allow us to study such groups—and indeed, we show that disaggregating the least educated is important for understanding US mortality change. Second, it is not clear when education levels have changed "too much" such that selection bias becomes an important concern. Our approach provides a principled way of quantifying the possible selection bias. If the bias is small, then the qualitative conclusions may be unchanged from a naïve approach. Online Appendix Section C.3 examines the similarities and differences between the results that arise from the use of these different methodological choices in the prior literature.

III. Results

A. *Applying the Methodology*

We begin by noting some details of the application of NRA bounds (Section II) to our specific mortality setting.

Selecting a Curvature Constraint.—Our primary results are computed numerically, under the assumptions of monotonicity and constrained curvature. To choose a conservative curvature constraint, we require the curvature to be less than 50 percent higher than the largest value of the curvature of the US income rank-mortality function reported by Chetty et al. (2016) (online Appendix Section C.2).[21]

[21] To generate a comparable $\bar{C}$ across all age-year-gender CEFs, we construct a "normalized" $\bar{C}$ which is the absolute value of the second derivative for the CEF, divided by the mean across all percentiles. This procedure accounts for the potential concern that, e.g., CEFs with higher mortality (for instance, in older groups) may have larger (unnormalized) $\bar{C}$ without having larger curvature.

Choosing Percentile Ranges.—Mortality can be most tightly bounded in rank intervals that are close to rank bin boundaries in the data. We select percentile bins for analysis by matching the education levels of 50–54-year-olds in 2003, a group that is approximately the middle cohort in our sample and an age group emphasized by the prior literature.[22] We calculate mortality for the following four education groups: (i) the bottom 10 percent (the share of the age 50–54 population who were high school dropouts in 2003); (ii) percentiles 10 to 45 (those with high school degrees only in 2003); (iii) percentiles 45 to 70 (2-year college degrees in 2003); and (iv) the top 30 percent (Bachelor's degrees or higher in 2003). Mortality estimates in education quartiles or deciles would be useful, but given the existing rank bin boundaries in the data, they cannot be bounded as tightly and are thus less informative. We do not present these, but they can be readily calculated from the shared code and data. Because the data include the universe of deaths and statistical uncertainty regarding the population totals is very small, we follow the previous literature in omitting confidence intervals.

Own-Gender, Across-Race Ranks.—As in Chetty et al. (2016), we rank men and women in each age-year group against members of their own gender, estimating mortality for a given percentile group of men or women. For instance, when we examine the bottom 10 percent of the education distribution, we mean "the least educated 10 percent of women," rather than "women in the bottom 10 percent of the entire population education distribution." We chose own-gender reference points because (i) women's and men's labor market opportunities and choices are often different, and (ii) women and men often share households and incomes, making population ranks misleading. We construct ranks *across* all racial groups (including other races, e.g., Hispanics, that we do not analyze in the paper).

Note that our method assumes that latent ranks are uniformly distributed within education rank bins; this assumption does not necessarily hold within racial groups. For instance, among the bottom 10 percent of women, the education ranks of *White* women (ranked against all women) may not be uniform. The assumption of uniformity is not integral to our approach; the analytical formulas we provide permit the imposition of arbitrary parametric assumptions about the ranks. In Section IIIE, we provide several pieces of evidence that the uniformity assumption does not bias our results on mortality changes. For parsimony, we therefore proceed with the assumption of uniformity but acknowledge that it does not hold exactly.

We present an alternative modeling choice that guarantees the assumption holds: in online Appendix Section D, we present results when we generate education ranks within own-race and own-gender groups. The advantage of the own-race approach is that reranking people within race and gender recovers uniformity of ranks within race-gender cells, by construction. The disadvantage is that doing so departs from the convention in the literature of comparing outcomes among all people within each gender, rather than within each race-gender group.

[22] For any other age group, a different set of percentile bins might yield tighter bounds, but we chose the same percentile bins for all groups to maximize comparability.

Nonmonotonicity.—For a small number of population subgroups, we do not observe monotonicity in the data. In the majority of these cases, the mortality rate in the higher education group is within 5 percent of that in the lower education group, so the monotonicity violation is not substantive in comparison with the width of the bounds. For Black cohorts over the age of 55 in 1992–1994, there are more substantial violations: high school graduates often have higher mortality than dropouts, and B.A. recipients often have higher mortality than individuals with some college. Such nonmonotonic means are isolated to the oldest Black cohorts.[23]

Because our constrained optimization imposes monotonic CEFs on these groups, we may *overstate* health improvements for the oldest Black age groups, because monotonicity makes mortality among dropouts look *worse* for these cohorts in the period 1992–1994. Given that our primary finding is divergence by education group, allowing for nonmonotonic mortality among older Black cohorts would only strengthen our results. We also show in online Appendix Section D that all our results are robust if we loosen either the monotonicity or constrained curvature assumptions.

B. *Unadjusted Mortality Changes by Education Levels*, *Ages 50–54*

Figure 1 presents the raw data for 50–54-year-olds in 3-year bins, showing total mortality (deaths/100,000) from 1992–1994 to 2016–2018, separately for each education level and by race and gender.

The four groups of points on each graph represent individuals with (i) less than high school education; (ii) high school education; (iii) some college; and (iv) a Bachelor's degree or higher. The mean education percentile for individuals in a given education category is plotted on the *x*-axis. In 1992–1994, 17.4 percent of women aged 50–54 had less than a high school education. The average percentile rank for someone in this group is $17.4/2 = 8.7$; mortality for this group is therefore plotted (with a Black triangle in panel A) at 8.7 on the *x*-axis. In 2016–2018, 8.0 percent of women had less than a high school education; their mean education percentile was 4.0 (yellow square). Intermediate points show the transition path between these years.

Among White women (panel A), 50–54-year-old high school dropouts had mortality rates 161 percent higher in 2016–2018 than in 1992–1994, suggesting an annualized mortality increase of 4.1 percent per year. Unadjusted mortality rose 38 percent for 50–54-year-old high-school-educated White women, rose 11 percent for women with some college, and fell by 35 percent for White women with Bachelor degrees or higher. Panels B through D present unadjusted estimates for White men, Black women, and Black men.

The points systematically shift to the left over time, because education for all race and gender groups rose steadily over the sample period. The decreasing average rank over time implies that unadjusted mortality changes at given education levels are biased upward by selection (Dowd and Hamoudi 2014; Bound et al.

[23] This pattern could arise from some form of positive selection, such as survival of the 1980s crime waves or HIV epidemic. Examination of this hypothesis is beyond the scope of this study.

2015; Currie 2018). The next section adjusts these estimates for changes in the size and relative rank of each group by studying constant percentile education groups rather than constant levels of education.

C. *Mortality Changes in Constant Education Percentile Bins, Ages 50–54*

We now turn to estimates of mortality for the same age group in constant percentile bins, which are displayed in Figure 5.[24] The top-left panel shows mortality rates for White women, with one series for each education percentile group. Mortality among the least educated 10 percent rose steadily from [614, 738] deaths per 100,000 in 1992–1994 to over 1,475 in 2016–2018, an increase of 100–150 percent, or about 2.9–3.9 percent per year. As expected, this is a smaller increase than the unadjusted mortality change shown in Figure 1; the original point estimate for dropouts is outside the bounds for the constant rank group, and the bias in the naïve estimate could be as large as 61 percent. However, even the lower bound on mortality change (+100%) implies a stark increase in mortality, and a change much higher than that in the next constant rank group. In percentiles 10–45, White women's mortality change is bounded between $[-6\%, 19\%]$. Among the most educated 30 percent, White women's mortality change was bounded between $[-45\%, -36\%]$, an annualized decline of about 2 percent per year.

Turning to White men, we find a similar divergence of the least educated 10 percent. Mortality increased by [36%, 62%] in the bottom 10 percent, while the group from percentiles 10–45 saw mortality changes in $[-6\%, 3\%]$. White men in the top 30 percent experienced mortality declines of at least 43 percent. As above, the naïve estimates from Section IIIB are outside of these bounds, but they are not far from the upper bound estimates.

The remaining panels of Figure 5 show estimates for 50–54-year-old Black women and men, respectively. Mortality rates among Blacks also diverged by education group, but less so than among Whites. Among 50–54-year-old Black women, mortality rose by 34–41 percent for the bottom 10 percent, but declined among all groups in the top 90 percent. For Black men aged 50–54, mortality change was close to zero in the bottom 10 percent but declined by at least 30 percent in all other groups.

D. *Constant Education Percentile Changes in Mortality at Other Ages*

This subsection expands the analysis to all age groups and presents our primary results. Figure 6 shows bounds on mortality change from 1992–1994 to

[24]The four constant education percentile groups correspond to education percentile bins in 2003. To obtain bounds on mortality changes when mortality in each year is interval-censored, we first estimate bounds on total mortality in 1992–1994 and 2016–2018, respectively denoted $\left[t^l_{2016}, t^u_{2016}\right]$ and $\left[t^l_{1992}, t^u_{1992}\right]$. We obtain mortality changes in percent terms as

$$\text{lower bound on mortality change} = 100 \times \left(t^l_{2016}/t^u_{1992}\right)$$

$$\text{upper bound on mortality change} = 100 \times \left(t^u_{2016}/t^l_{1992}\right).$$

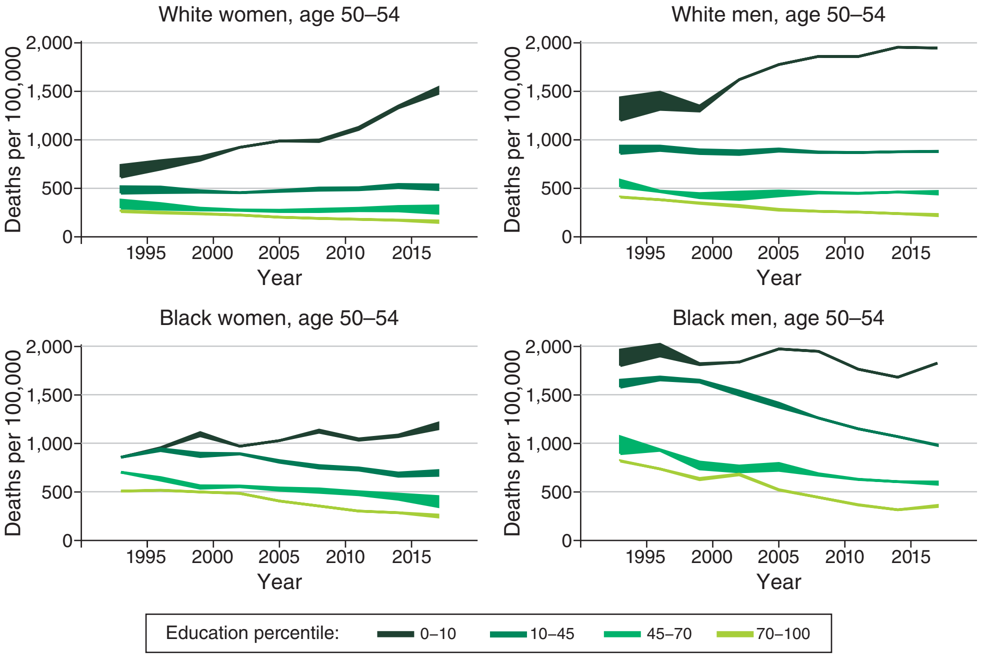


Figure 5. All-Cause Mortality Change in Constant Education Percentiles: Age 50–54, 1992–1994 to 2016–2018

Notes: "White" refers to non-Hispanic White and "Black" to non-Hispanic Black. Each interval represents the bounded set containing the number of deaths per 100,000 people in a given time period, among people in the education percentiles specified in the legend. The education percentiles correspond to the percentile bins describing four levels of education for the median age group in 2003: No High School, High School, Some College, and a B.A. or Higher. Bounds are computed as described in Section II. The sample consists of people ages 50–54.

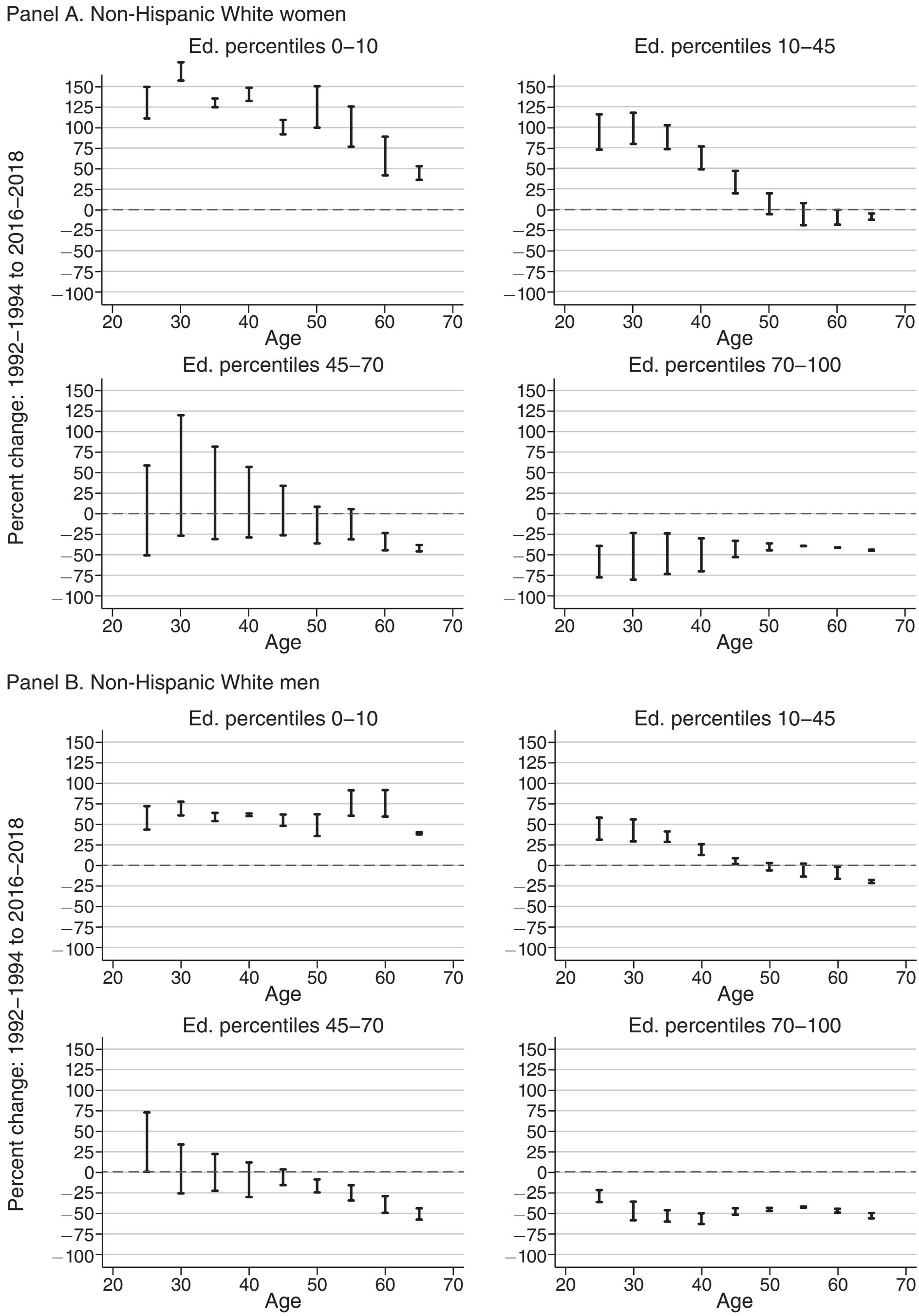
Sources: ACS, CPS, NCHS

2016–2018, separated by race, gender, age bin, and the four constant education percentile categories described above. We draw attention to three findings.

First, the mortality increases among White men and women (panels A and B of Figure 6) are principally driven by the bottom 10 percent. Trends in mortality rates among White men and women of all ages are similar to the trends for 50–54-year-olds discussed in Section IIIC. Among percentiles 0–10, White men and women experienced large mortality increases—larger than 50 percent for most age groups.

In education percentiles 10–45, mortality is largely flat or declining among Whites over the age of 50. At younger ages, trends are stark, paralleling those in the bottom ten percentiles: mortality has risen 30–61 percent among 25-year-old White men and 73–116 percent among similarly aged White women. However, since most deaths occur at older ages, the all-age death rate in percentiles 10–45 is relatively flat among Whites. While we lack precision for percentiles 45–70, we observe decisive *declines* in mortality in the top 30 percentiles of the education distribution for both men and women.

Second, we observe divergence of mortality by education among Black men and women, but without the large rise in mortality in the bottom 10 percent (panels C



(*continued*)

Figure 6. Mortality Change in Constant Education Percentiles (1992–1994 to 2016–2018, All Ages)

and D of Figure 6). In particular, all percentiles *except* 0–10 exhibited substantial mortality reductions among Blacks, while mortality change hovered around zero in the bottom 10 percent. Some middle-aged Black cohorts in the bottom 10 percent

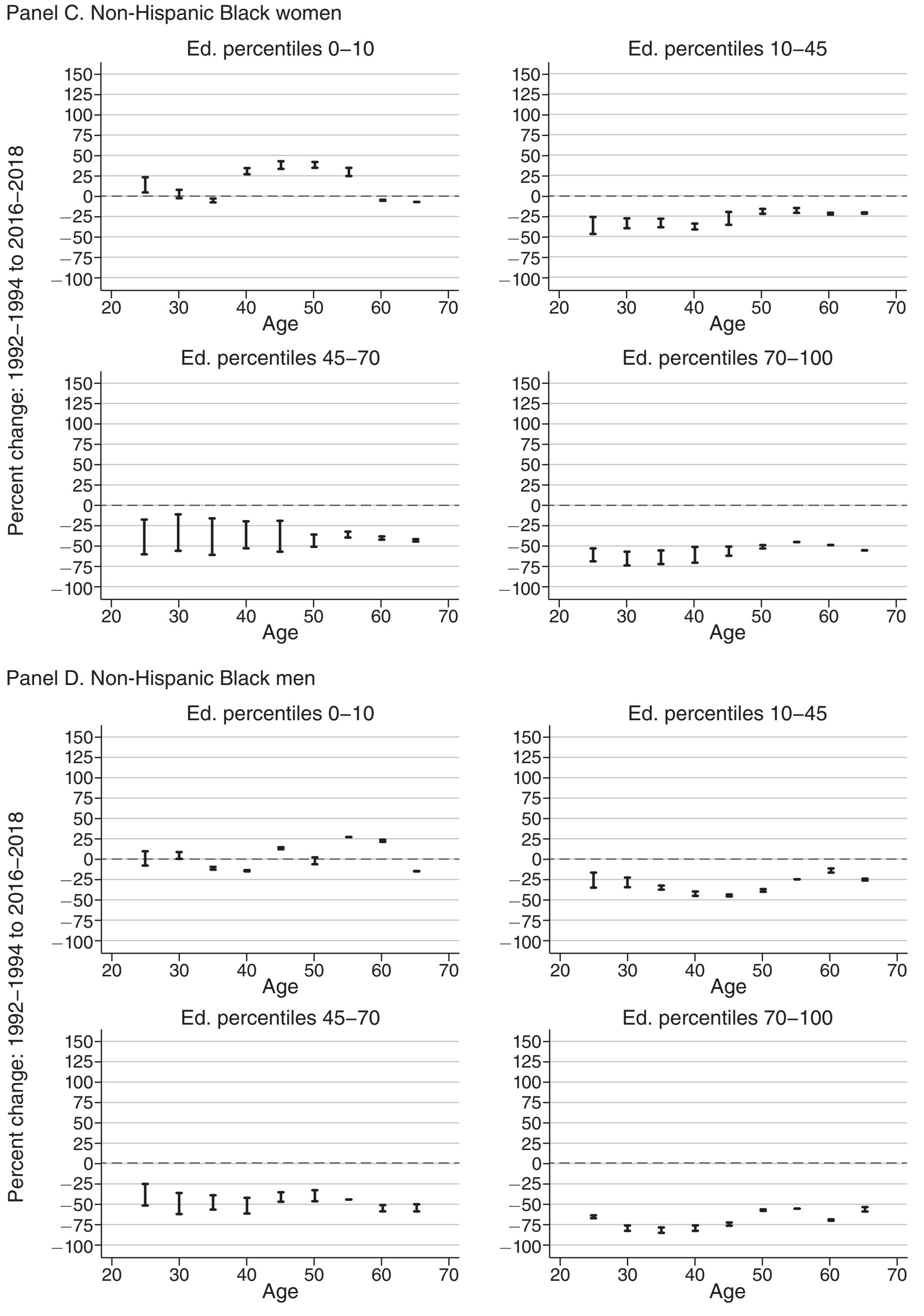
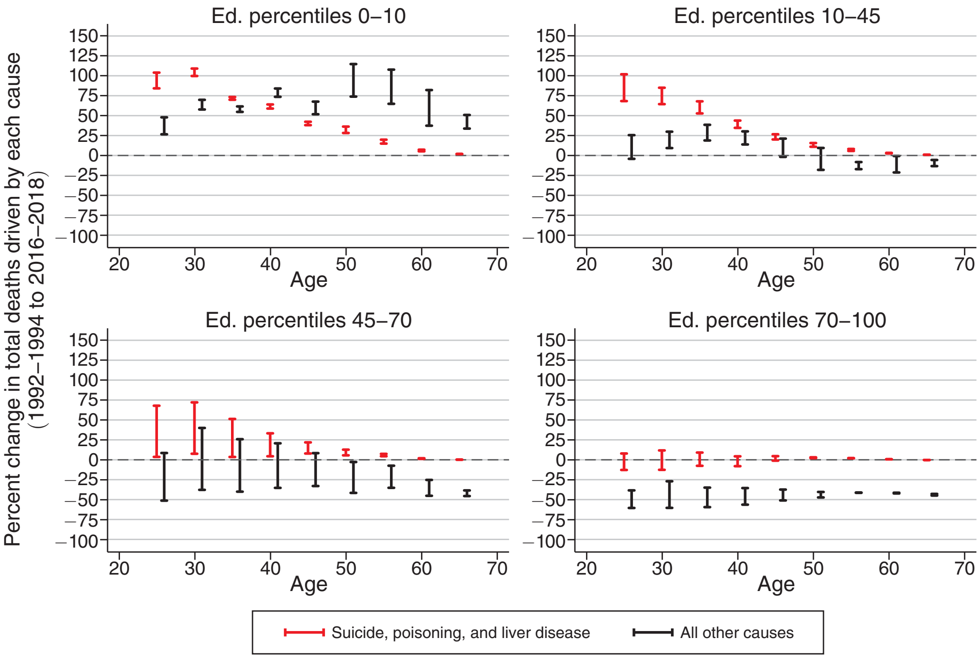


Figure 6. Mortality Change in Constant Education Percentiles (1992–1994 to 2016–2018, All Ages) (*continued*)

Notes: The graph shows changes in mortality by age, sex, race, and constant percentile education bin. The vertical lines show the bounded set containing the percentage change in the mortality rate from 1992–1994 to 2016–2018 for the given age group. Bounds are computed as described in Section II.

Sources: ACS, CPS, NCHS

Panel A. Non-Hispanic White women



(*continued*)

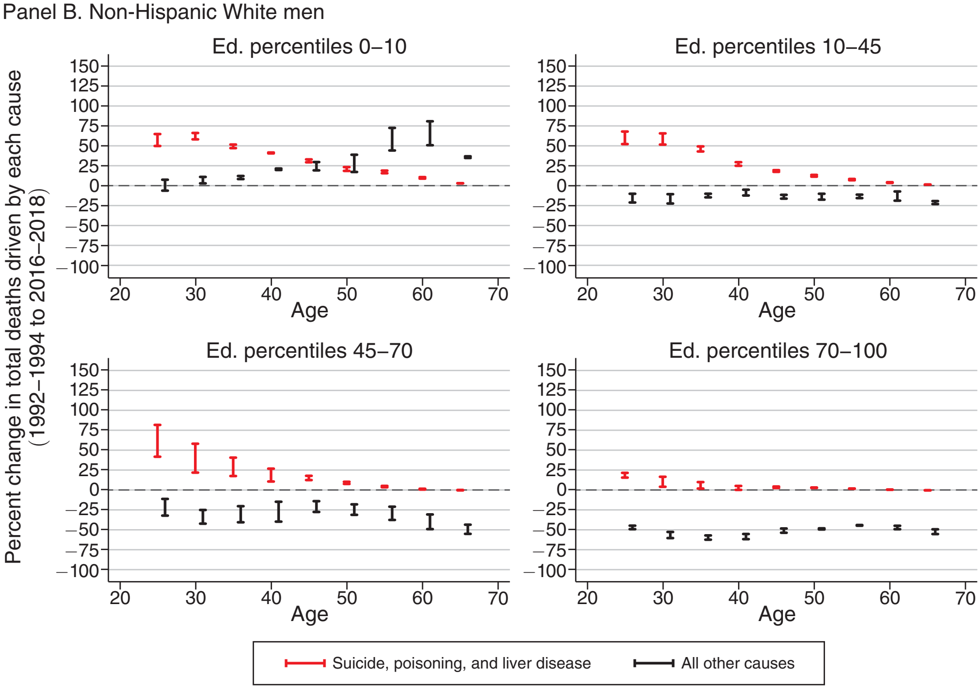
FIGURE 7. DECOMPOSITION OF MORTALITY CHANGE FROM 1992–1994 TO 2016–2018: CONTRIBUTION OF DEATHS OF DESPAIR

saw mortality increases (especially Black women aged 40–59), while other ages saw only small positive mortality increases or reductions. Black men in the most educated 30 percent had the largest reductions in middle-age mortality out of all groups, closing some of the mortality gap with White men.

Third, we find that the proximate causes of these mortality changes vary substantially by race and age. Figure 7 decomposes the mortality changes above into deaths from suicide, poisoning, and liver disease (what Case and Deaton 2017 call "deaths of despair," orange bars) and all other deaths (black bars). We present the percent change in *total deaths* driven by each of the two causes, so that adding the two bars produces the total percentage mortality change (as displayed in Figure 6).

For example, consider the first group displayed in panel A: White women aged 25–29 in the least educated 10 percent. Mortality from deaths of despair for this group increased by 616–750 percent. However, because they started from a low base, this change in deaths of despair mechanically caused total mortality to rise by 84–104 percent (the orange bar on the graph). Deaths from all causes *other* than despair increased by 30–55 percent for this group, causing total mortality to rise by 27–48 percent (the black bar on the graph).[25]

[25] More precisely, let total deaths in year y be in interval $\left[t_y^l, t_y^u\right]$. Let deaths of despair d be in interval $\left[d_y^l, d_y^u\right]$. Index years 2016–2018 and 1992–1994 by 2016 and 1992, respectively. The lower bound for the orange bar is



(*continued*)

FIGURE 7. DECOMPOSITION OF MORTALITY CHANGE FROM 1992–1994 TO 2016–2018: CONTRIBUTION OF DEATHS OF DESPAIR

Among White men and White women below 45, deaths of despair play an important role in driving overall mortality increases. In particular, the gains in deaths of despair are responsible for the majority of the mortality increase of White men below 40. At older ages, deaths of despair are still rising but play a minor role in overall mortality increases. Deaths of despair are also rising in percentiles 45–70 at all ages for both White men and women, and for White men under 50 even in the highest educated group (albeit at much lower rates).

One way of summarizing these results is to aggregate mortality changes across all ages, though it masks some of the important heterogeneity and emphasizes changes at older ages where most deaths occur. To aggregate mortality rates across ages while holding constant the change in the population age distribution over time, we weight the age-specific mortality rates in the data with the standardized US

given by

$$\textit{lower bound} \ = \ 100 \times \left(d^{l}_{2016} - d^{u}_{1992}\right)/t^{u}_{1992},$$

whereas the upper bound is given by

$$\textit{upper bound} \ = \ 100 \times \left(d^{u}_{2016} - d^{l}_{1992}\right)/t^{l}_{1992}.$$

Mortality changes for other deaths are given similarly.

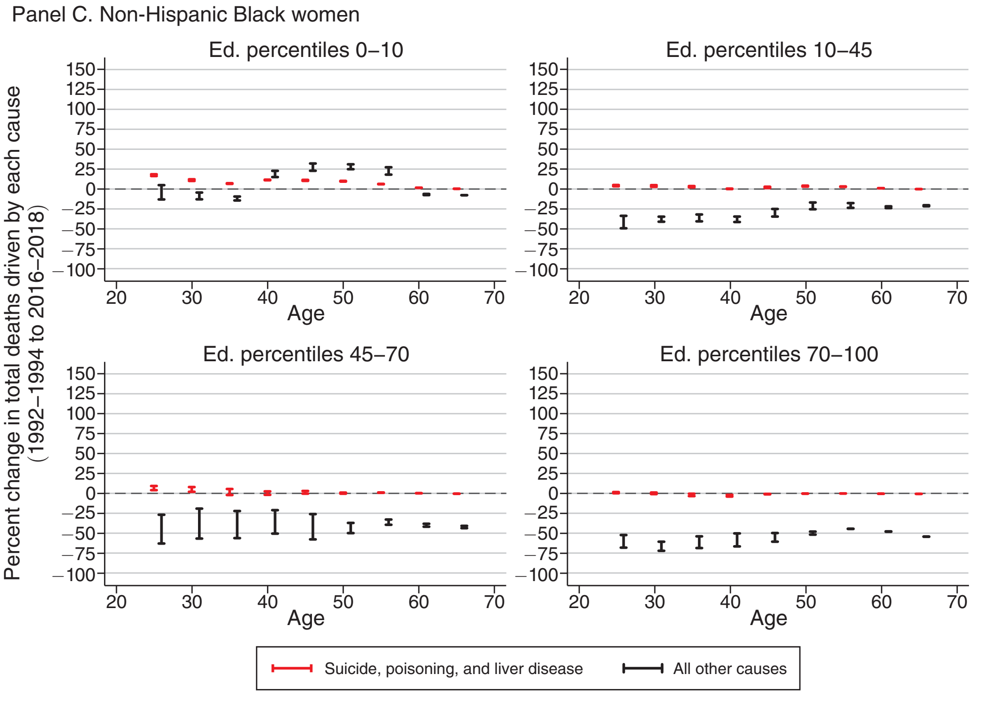


Figure 7. Decomposition of Mortality Change from 1992–1994 to 2016–2018: Contribution of Deaths of Despair (*continued*)

population distribution for ages 25–69.[26] Using these aggregates, Table 2 presents age-adjusted bounds on mortality change from 1992–1994 to 2016–2018 for each constant education percentile group.

The table highlights the substantial divergence of mortality rates between high and low education groups in all four gender and race groups. For Whites in the least educated 10 percent, mortality rose substantially, by 69–112 percent for women and 47–67 percent for men. Among Blacks in the least educated 10 percent, mortality rose for women by 9–17 percent and was close to unchanged for men (−1 percent to +3 percent). For the most educated 30 percent of individuals across all race and gender groups, mortality rates fell by over 35 percent, with the largest gains for the most educated Black men.[27]

Online Appendix Table A2 shows the percent increase in deaths of despair, as well as heart disease, cancer, injuries and other causes for all sex/race groups, combining all ages. It is notable that among middle-aged Whites in the bottom

[26] The standardized US population distribution was obtained from https://seer.cancer.gov/stdpopulations/.

[27] These numbers, along with mortality levels by age, education bin, race and sex for all groups are reported in the accompanying data files. As noted earlier, online Appendix Figure A1 plots these estimates against the naïve estimates of mortality change at four *levels* of education: dropouts, high school completion, some college, and B.A. or higher. We also disaggregate these over time, by race, in online Appendix Figure A2.

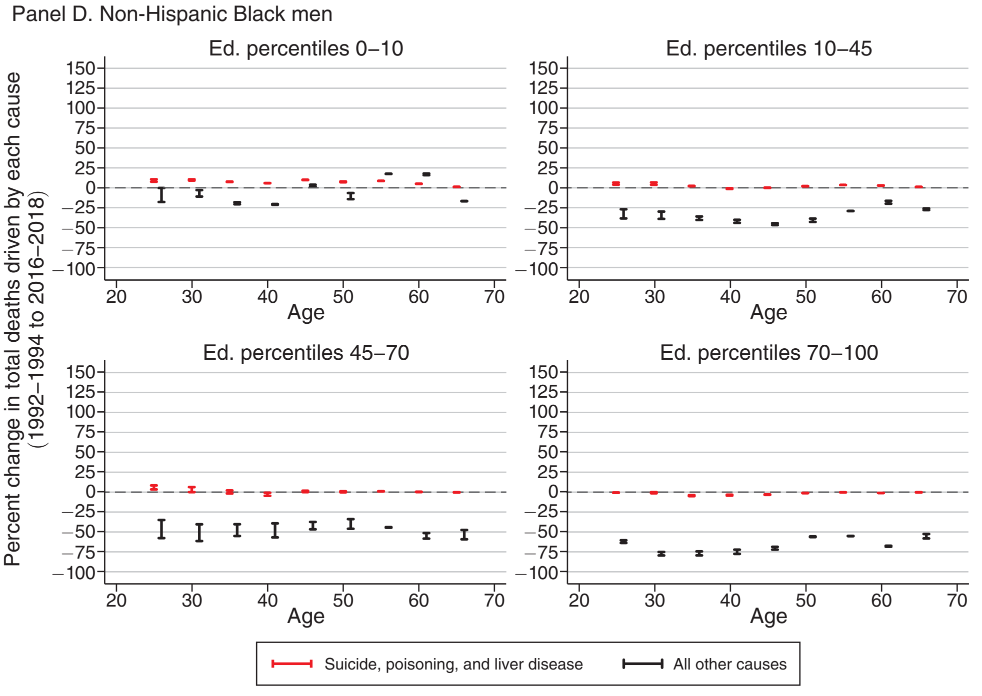


Figure 7. Decomposition of Mortality Change from 1992–94 to 2016–18: Contribution of Deaths of Despair (*continued*)

Notes: "White" refers to non-Hispanic White. The panels decompose the change in total mortality from 1992–1994 to 2016–2018 into two parts: the change in total deaths driven by deaths of despair, and the change in total deaths driven by all other causes. Estimates are disaggregated by age, sex, race, and constant percentile education bin. The orange lines show bounds on the contribution to total mortality change driven by changes in deaths of despair. The value on the *y*-axis is the amount that total mortality for each group would have changed if the rates of all deaths *other* than deaths of despair were unchanged. The black lines show the contribution to total mortality change driven by all causes of death *other* than deaths of despair. Deaths of despair are deaths from suicide, poisoning, and chronic liver disease. Bounds are computed using the set identification methods described in Section II.

Sources: ACS, CPS, NCHS

10 percent, mortality from cancer, heart disease, and other diseases all rose over the sample period.[28]

E. *Robustness*

Online Appendix D presents a range of robustness checks on the primary results. Online Appendix Section D.1 explores how the bounds change as the monotonicity restriction is loosened; for the least educated group, loosening monotonicity does not lead to dramatically different results because the empirical monotonicity across bins is so strong.

[28] Note that in online Appendix Table A2, we show the percentage changes for each cause, while Figure 7 shows the contribution of deaths of despair to total mortality. For deaths of despair, the percentage changes are very large because they begin from a small base; but for many cohort groups, they contribute only a small amount to changes in total mortality.

TABLE 2—AGE-ADJUSTED CHANGES IN ALL-CAUSE MORTALITY BY EDUCATION PERCENTILE, 1992–94 TO 2016–18

	0–10th	10th–45th	45th–70th	70th–100th
White women	(+77%, +111%)	(−0%, +21%)	(−39%, −4%)	(−48%, −39%)
White men	(+50%, +68%)	(−6%, +5%)	(−40%, −18%)	(−52%, −45%)
Black women	(+11%, +17%)	(−28%, −21%)	(−47%, −33%)	(−55%, −51%)
Black men	(−0%, +3%)	(−33%, −29%)	(−54%, −44%)	(−67%, −64%)

Notes: "White" refers to non-Hispanic White and "Black" refers to non-Hispanic Black. The table shows the percent change in all-cause mortality, defined as total deaths in a year divided by population. To hold the population distribution constant, we weight the age-specific mortality rates from the data with the standardized US population distribution for ages 25–69. We use age-specific mortality rates from each period, but a single set of weights for all periods.

Online Appendix Section D.2 shows that results are robust to alternate assumptions on the bounding methodology. We show that results are similar when: (i) bin boundaries are based on education levels in 1992–1994; (ii) education percentiles are defined relative to members of the same race and gender, rather than just the same gender; (iii) we permit sheepskin effects in education (allowing the CEF to have discrete jumps at bin boundaries); and (iv) we remove the curvature constraint and permit CEFs with unconstrained curvature.

As noted above, defining education percentiles relative to the same race and gender, as in (ii), guarantees that the uniformity assumption holds, so it is a strong confirmation that non-uniform distributions within each bin do not drive our results. We further probe this assumption in online Appendix Section D.3, where we demonstrate that our findings cannot be explained by changes in the relative distribution of Black-White education ranks *within* education bins. Together, these imply that our assumption of uniform ranks within bins for each race group is not leading to bias.

Division Bias.—One concern with our mortality estimates may be that they are calculated by dividing the number of deaths (from NCHS data) by the population (from the CPS). If ethnic status or education is misreported in one of the two datasets, our mortality estimates could be biased. Note that for our mortality *change* estimates to be biased, the extent of misreporting would have to change differentially across datasets. We show that the division bias from any such misreporting is unlikely to be large enough to spuriously generate the large changes in mortality that we find among the least educated Whites.

We would be most concerned if death records increasingly overstate the number of White high school dropouts among the deceased, and/or the CPS increasingly understates the population of White high school dropouts. Either of these situations would cause our mortality change estimates to be biased upward. One way that this could happen would be if individuals who are White increasingly report Hispanic identity in the CPS, but not in the death records.[29] A second way would be

[29] Note that if Hispanics increasingly report as White, that would cause our mortality change estimates to be biased down (i.e., our reported estimates are *conservative*), because Hispanics generally have lower mortality than Whites.

if individuals who are dropouts increasingly inflate their education when responding to the CPS (thus lowering the population count of dropouts), but their education is correctly reported on death certificates.[30] However, if there is a constant rate of differential misreporting between CPS and death certificates, our change estimates are not biased.

We address division bias in three ways. First, in online Appendix Section D.4, we show that measurement error in ethnicity or changes in reporting patterns of Hispanic identity cannot explain our results. We simulate systematic measurement error in Hispanic identity and show that our results are sustained even with highly implausible changes in patterns of Hispanic reporting. This exercise rules out that, say, a greater propensity among economically successful Hispanics to identify as White could yield our results.

Second, in online Appendix Section D.5, we bound the error that could arise from false reporting of education or ethnicity in the CPS by examining the size of synthetic CPS dropout cohorts over time. If CPS respondents increasingly overreport their education, or if White respondents are increasingly reporting themselves as Hispanic, then the synthetic cohort of non-Hispanic White dropouts will shrink in size more than can be explained by the death rate and the rate of continuing adult education. We show that under the worst-case assumptions for our hypothesis, misreporting of education in the CPS could potentially account for less than 8 percent of the mortality change of the least educated White women in the 1950–1954 birth cohort and less than 24 percent in the 1960–1964 birth cohort. The worst-case bias for White men and at higher education groups is even smaller. As we discuss in online Appendix Section D.5, this worst-case bias scenario is very unlikely to be true; it is therefore implausible that erroneous population counts in the CPS are driving our findings.

Third, in online Appendix Section D.6, we calculate mortality rates and other health measures using the NHIS (Centers for Disease Control and Prevention 1998–2019). The NHIS makes it possible to measure mortality in a sample of individuals without any division bias, because survey respondents are followed up for many years and any deaths are recorded. The NHIS broadly supports the notion of divergent outcomes between high-school dropouts and high-school completers (Hendi 2015, 2017; Sasson 2016, 2017), but the very small samples of dropouts lead to very imprecise estimates.[31]

In online Appendix Section D.6, we also examine self-reported health status in the NHIS, which is measured more precisely than mortality. We find that self-reported health status declines more for White female dropouts than for White women with all higher levels of education, with the difference concentrated among 40–60 year olds—the same age group that had the highest differential mortality change between dropouts and high school completers in our main analysis

[30] It is also possible that true Hispanic identity is decreasingly reported on death certificates, or that death certificates increasingly report dropout status either among dropouts or those with high school. We view these circumstances as less likely, but the tests below address them as well.

[31] Hendi (2015) finds that mortality is not rising for the male dropouts in the NHIS. Our NHIS analysis is consistent with Sasson (2017), who argues that the NHIS sample of White male dropouts is too small to distinguish between zero mortality change and our reported effects of 1.6–2.2 percent growth per year in the bottom 10 percent.

(Figure 6). Among men, health changes are similar between dropouts and high school graduates for young men, but among older men, dropouts experience substantially more health deterioration than high school completers, again consistent with the results in Figure 6. Changes in self-reported health status are thus consistent with our finding that mortality changes among Whites are driven by those in the least educated 10 percent.

Finally, in online Appendix Section D.7, we replicate the analysis after pooling dropouts and high school completers into a single education group. This eliminates most of the division bias and misreporting concerns because: (i) the synthetic cohort analysis above shows that the size of the less than or equal to high school (LEHS) population in the CPS cannot be biased by more than 10 percent either for men or for women; and (ii) the group size is much larger, so a small amount of misreporting cannot substantially shift the population size and bias the estimated mortality rate. The disadvantage of pooling these groups is that we can no longer tightly bound mortality among the bottom 10 percent. However, we can decisively reject the hypothesis that mortality change among the bottom 45 percent of Whites is driven by selection alone. For cohorts under the age of 45, we continue to find that mortality rates in the bottom 45 percent of the education distribution have risen by more than 50 percent for White women and 25 percent for White men from 1992–1994 to 2016–2018. These numbers are lower than the estimates for mortality increases in the bottom 10 percent in the main part of the paper, because they pool the high mortality increases among the bottom 10 percent with the smaller mortality increases among percentiles 10–45.

To conclude, while there is undoubtedly some measurement error in education and ethnicity in both the vital statistics and the CPS data, it is very unlikely that measurement error can explain the substantial increase in mortality among the least educated non-Hispanic Whites. It is worth noting that other measures of socioeconomic status also have their limitations; for example, studies using income as a measure of socioeconomic status often exclude those reporting zero income, and do not consider all transfers or illicit income, which may be important at the bottom of the income distribution.

IV. Conclusion

This paper makes two primary contributions. Methodologically, we introduce new bounds on conditional expectation functions with interval-censored conditioning data. Our approach is particularly useful for bounding CEFs with education data. In many cases, one wishes to present trends in a given outcome over time by education group (e.g., wages, fertility, or marriage rates over time for people with a BA), an analysis that is subject to similar concerns about selection raised in this paper. Our method addresses these concerns by making it feasible to track outcomes in constant education ranks over time.

The method is broadly applicable to other contexts as long as researchers are willing to assume some parametric distribution for conditioning data. Other settings where it could be useful applied include the study of CEFs with top-coded or interval-censored income data, Likert scales, or bond ratings.

Empirically, this paper studies US mortality change at different points in the education distribution. The postwar era has been characterized by improving health and survival of nearly all demographic groups in all developed countries. Rising mortality among White non-Hispanic Americans represents a major deviation from this trend, and understanding the factors behind this change is a central policy concern.

While there has been substantial interest in education as a risk factor for mortality change, the selection bias inherent in earlier estimates of mortality among the less educated has made it difficult to study. Our approach generates estimates of mortality change in constant education percentiles that quantify the uncertainty from changing education bin boundaries over time. Our findings point to large increases in mortality for White men and women in the bottom 10 percent of the education distribution, indicating a public health crisis among the least educated.

Our findings reconcile several views previously expressed in the literature. We confirm that earlier estimates of mortality at constant levels of education did overstate mortality increases due to selection bias. However, the mortality change due to selection bias is swamped by the actual mortality change at constant education percentiles. Death rates for the least educated have dramatically diverged from death rates of other groups in virtually all middle-age race and gender groups. These mortality increases have a range of causes beyond the widely discussed increases in deaths of despair.

These findings are consistent with the mortality divergence across education groups from 1981 to 2000 (Meara, Richards, and Cutler 2008); we show that this divergence has continued through 2018 and cannot be explained by selection bias from rising education. Our findings also support the Case and Deaton (2015, 2017) findings that rising middle-age mortality is concentrated among less educated Whites. But our analysis focuses on more disaggregated population subgroups where education levels have changed substantially over time (e.g., younger cohorts and women)—subgroups where unadjusted estimates were thought to be substantially biased.

These results provide a new perspective to recent analyses of changing mortality at different percentiles in the income distribution (Cristia 2009, Chetty et al. 2016). Like the poor, the least educated experience a range of socioeconomic disadvantages, such as high unemployment, low insurance coverage, poor nutrition, and exposure to harmful environmental factors. Our estimates imply that recent middle-age mortality increases among the least educated 10 percent are worse than those among the poorest 10 percent. This could be because low income is more transitory than low education or because education is a marker of early life disadvantage and reflects low socioeconomic status in the present as well as in past years.

To better understand the causes of these mortality increases, researchers have searched for factors that predict rising mortality (Cutler et al. 2011; Case and Deaton 2017; Ruhm 2018). Pinpointing the causes of mortality increase will require taking into account the fact that education is a key predictor of mortality change and that the proximate drivers of mortality change differ substantially across different groups.

REFERENCES

Aizer, Anna, and Janet Currie. 2014. "The Intergenerational Transmission of Inequality: Maternal Disadvantage and Health at Birth." *Science* 344 (6186): 856–61.

American Council on Education. 1993–2014. *GED Testing Program Statistical Reports*. Washington, DC: American Council on Education. https://nces.ed.gov/programs/digest/d15/tables/dt15_219.60.asp (accessed March 22, 2019).

Asher, Sam, Paul Novosad, and Charlie Rafkin. 2022. "Intergenerational Mobility in India: New Measures and Estimates across Time, Space, and Communities." Unpublished.

Bertrand, Marianne, Patricia Cortés, Claudia Olivetti, and Jessica Pan. 2021. "Social Norms, Labour Market Opportunities, and the Marriage Gap between Skilled and Unskilled Women." *Review of Economic Studies* 88 (4): 1936–78.

Bound, John, Arline T. Geronimus, Javier M. Rodriguez, and Timothy A. Waidmann. 2015. "Measuring Recent Apparent Declines in Longevity: The Role of Increasing Educational Attainment." *Health Affairs* 34 (12): 2167–73.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *The Handbook of Labor Economics*, Vol. 3A, edited by Orley C. Ashenfelter and David Card, 1801–63. Amsterdam: Elsevier.

Case, Anne, and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century." *Proceedings of the National Academy of Sciences* 112 (49): 15078–83.

Case, Anne, and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity* 2017 (1): 397–476.

Centers for Disease Control and Prevention. 1998–2019. "National Health Interview Survey." https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm (accessed May 28, 2019).

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang. 2017. "The Fading American Dream: Trends in Absolute Income Mobility since 1940." *Science* 356 (6336): 398–406.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The Association between Income and Life Expectancy in the United States, 2001–2014." *JAMA* 315 (16): 1750–66.

Cristia, Julian P. 2009. "Rising Mortality and Life Expectancy Differentials by Lifetime Earnings in the United States." *Journal of Health Economics* 28 (5): 984–95.

Currie, Janet M. 2018. "Inequality in Mortality over the Life Course: Why Things Are Not as Bad as You Think." *Contemporary Economic Policy* 36 (1): 7–23.

Currie, Janet, and Hannes Schwandt. 2016a. "Inequality in Mortality Decreased among the Young While Increasing for Older Adults, 1990–2010." *Science* 352 (6286): 708–12.

Currie, Janet, and Hannes Schwandt. 2016b. "Mortality Inequality: The Good News from a County-Level Approach." *Journal of Economic Perspectives* 30 (2): 29–52.

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

Cutler, David, and Adriana Lleras-Muney. 2010a. "The Education Gradient in Old Age Disability." In *Research Findings in the Economics of Aging*, edited by David A. Wise. Chicago: University of Chicago Press.

Cutler, David M., and Adriana Lleras-Muney. 2010b. "Understanding Differences in Health Behaviors by Education." *Journal of Health Economics* 29 (1): 1–28.

Dowd, Jennifer B., and Amar Hamoudi. 2014. "Is Life Expectancy Really Falling for Groups of Low Socio-economic Status? Lagged Selection Bias and Artefactual Trends in Mortality." *International Journal of Epidemiology* 43 (4): 983–88.

Feinstein, Alvan R., Daniel M. Sosin, and Carolyn K. Wells. 1985. "The Will Rogers Phenomenon: Stage Migration and New Diagnostic Techniques as a Source of Misleading Statistics for Survival in Cancer." *New England Journal of Medicine* 312 (25): 1604–08.

Gelman, Andrew, and Jonathan Auerbach. 2016. "Age-Aggregation Bias in Mortality Trends." *Proceedings of the National Academy of Sciences* 113 (7): E816–17.

Goldring, Thomas, Fabian Lange, and Seth Richards-Shubik. 2016. "Testing for Changes in the SES-Mortality Gradient When the Distribution of Education Changes Too." *Journal of Health Economics* 46: 120–30.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review* 104 (5): 348–53.

Hendi, Arun S. 2015. "Trends in US Life Expectancy Gradients: The Role of Changing Educational Composition." *International Journal of Epidemiology* 44 (3): 946–55.

Hendi, Arun S. 2017. "Trends in Education-Specific Life Expectancy, Data Quality, and Shifting Education Distributions: A Note on Recent Research." *Demography* 54 (3): 1203–13.

Hungerford, Thomas, and Gary Solon. 1987. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 69 (1): 175–77.

Kochanek, Kenneth D., Elizabeth Arias, and Brigham A. Bastian. 2016. "The Effect of Changes in Selected Age-Specific Causes of Death on Non-Hispanic White Life Expectancy between 2000 and 2014." *NCHS Data Brief* 250: 1–8.

Leive, Adam A., and Christopher J. Ruhm. 2021. "Has Mortality Risen Disproportionately for the Least Educated?" *Journal of Health Economics* 79: 102494.

Mackenbach, Johan P., Vivian Bos, Otto Andersen, Mario Cardano, Giuseppe Costa, Seeromanie Harding, Alison Reid, Örjan Hemström, Tapani Valkonen, and Anton E. Kunst. 2003. "Widening Socioeconomic Inequalities in Mortality in Six Western European Countries." *International Journal of Epidemiology* 32 (5): 830–37.

Manski, Charles F., and Elie Tamer. 2002. "Inference on Regressions with Interval Data on a Regressor or Outcome." *Econometrica* 70 (2): 519–46.

Markides, Kyriakos S., and Karl Eschbach. 2005. "Aging, Migration, and Mortality: Current Status of Research on the Hispanic Paradox." *Journals of Gerontology: Series B* 60 (2): 68–75.

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

National Center for Health Statistics. 1993–2019. "Mortality: All County." https://www.cdc.gov/nchs/index.htm (accessed May 28, 2019).

Novosad, Paul, Charlie Rafkin, and Sam Asher. 2022. "Replication data for: Mortality Change among Less Education Americans." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E14796V1.

Olshansky, S. Jay, Toni Antonucci, Lisa Berkman, Robert H. Binstock, Axel Boersch-Supan, John T. Cacioppo, Bruce A. Carnes, et al. 2012. "Differences in Life Expectancy Due to Race and Educational Differences Are Widening, and Many May Not Catch Up." *Health Affairs* 31 (8): 1803–13.

Pappas, Gregory, Susan Queen, Wilbur Hadden, and Gail Fisher. 1993. "The Increasing Disparity in Mortality between Socioeconomic Groups in the United States, 1960 and 1986." *New England Journal of Medicine* 329 (2): 103–09.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "IPUMS USA." https://www.ipums.org (accessed September 19, 2017).

Ruhm, Christopher J. 2018. "Deaths of Despair or Drug Problems?" NBER Working Paper 24188.

Sasson, Isaac. 2016. "Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990–2010." *Demography* 53 (2): 269–93.

Sasson, Isaac. 2017. "Reply to Trends in Education-Specific Life Expectancy, Data Quality, and Shifting Education Distributions: A Note on Recent Research." *Demography* 54 (3): 1215–19.

United States Census Bureau. 1993–2019a. "Current Population Survey." https://ceprdata.org/cps-uniform-data-extracts/march-cps-supplement/march-cps-data/ (accessed May 28, 2019).

United States Census Bureau. 2007–2019b. "American Communities Survey." https://ceprdata.org/acs-uniform-data-extracts/ (accessed May 8, 2020).