Journal of Health and Social Behavior

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT

to article in

Journal of Health and Social Behavior, 2019, Vol. 60, Issue 2

Weathering, Drugs and Whack-a-mole: Fundamental and Proximate Causes of Widening Educational Inequity in U.S. Life Expectancy by Sex and Race, 1990–2015

Arline T. Geronimus
University of Michigan

John Bound
University of Michigan

Timothy A. Waidmann
Urban Institute

Javier M. Rodriguez
Claremont Graduate University

Brenden Timpe
University of Michigan

METHODOLOGICAL APPENDIX

Data

Mortality data are drawn from the public-use version of the National Center for Health Statistics' Multiple Cause of Death (MCD) files, accessed through the NBER (National Bureau of Economic Research n.d.). We use the MCD files' cause of death variables to assign individuals to one of 17 mutually exclusive categories, as laid out in Table A5. The population at risk of death is constructed using the 1990 and 2000 decennial Census, the 2010 ACS 3-year sample, and the 2015 ACS accessed via IPUMS (Ruggles et al. 2017).

While the Census bureau has imputed educational attainment for all individuals not reporting it, this information is missing for a portion of vital records and must be imputed. There are two sources of missing education data in vital records. First, in two of the years examined, several states reported no education information on the death certificate in the first two years studied (7 in 1990 and 3 in 2000). To improve consistency over time, we obtained a restricted version of the MCD that includes geocoding for state and county of occurrence and limited our analyses to deaths occurring in the 43 states (and the District of Columbia) that reported education in all 4 years studied. This restriction dropped Georgia, Louisiana, New York, Oklahoma, Rhode Island, South Dakota and Washington. We experimented with the alternative of including states when they became available, and found doing so had virtually no effect on our results. Second, in the remaining states, a small share of death certificates had missing information. In the states where only some data were missing, we used the distribution of educational attainment in non-missing cases, by cause of death, age, sex and race, to impute education to the missing cases, assuming that education was missing randomly. Thus if in 1990, among death certificates with non-missing education 20% of black women who died of opioid abuse between ages 35 and 39 were in the lowest quartile of education (as determined by the distribution of census observations), then we randomly assigned 20% of the death certificates for black women age 35–39 who died of opioid abuse and had missing education to be in the lowest quartile.

A second concern is with the quality of next-of-kin reporting on education in vital records. Indeed, there is evidence from linked survey and vital records data that death certificates tend to over-report high school graduation relative to surveys (Rostron, Boies, and Arias 2010). Reassuringly, in studying research questions where, unlike in our case, it is possible to analyze various data sources, comparisons between tabulations based on standard methods we use, combining vital statistics and census data, and analyses based on linked administrative and survey data show similar trends (Sasson 2017).

Alternative data

Researchers studying the changes over time in the mortality gradient have either analyzed matched vital statistics and census data (as we do) (Bound et al. 2015; Case and Deaton 2015, 2017; Olshansky et al. 2012; Sasson 2016) or have used survey data linked with vital statistics data (Hendi 2015; Ho 2017; Ho and Fenelon 2015; Meara, Richards, and Cutler 2008). There are potential limitations with both sources of data. There is evidence that educational attainment is sometimes misreported on death certificates used in standard life table construction methods, with, for example, individuals more likely to be identified as high school graduates on death

certificates than in linked survey data (Rostron et al. 2010; Sorlie and Johnson 1996). The linked survey data, however, excludes the institutionalized population, such as those incarcerated or in nursing homes (National Center for Health Statistics, Office of Analysis and Epidemiology 2018), important population segments for studying mortality and social inequities in LE that, themselves, are socially patterned. An additional complication is that not all survey records are successfully linked to the vital statistics mortality data. Analysis of the linking procedure and comparisons between aggregate mortality statistics and those derived from linked data suggest that the linked data miss the more vulnerable segments of the population (National Center for Health Statistics, Office of Analysis and Epidemiology 2009; Sasson 2017). Furthermore, as survey non-response has become more common, an increasing share of respondents are linked probabilistically or declared ineligible for linking at all (National Center for Health Statistics, Office of Analysis and Epidemiology 2009). We also note that in contexts in which linked data include sufficiently large samples to support studies, the linked and matched data yield similar results (Sasson 2017).

To explore the implications of these different data sources for our estimates, we compare our estimates of drug-related death to those reported in Ho (2017). These estimates provide the most closely related comparison for our figures. Even so, they differ in several important ways: The definition of drug-related deaths, the age range, the method used to calculate YLLs, and the division of education by level rather than within-cohort rank. To facilitate a direct comparison, we use cohort/race/sex-specific distributions of educational attainment to convert her estimates by education level to estimates by bottom 25% and top 75% of educational attainment. We present these transformed estimates for 30–60-year-olds in columns 1 and 4 of Table A1.

In columns 2 and 5, we present our estimates of drug-related deaths by education rank for ages 25–60. Our estimated YLLs are generally higher and grow more over the study period. This is partly due to our more expansive definition of drug-related deaths, suggesting that the conclusions we draw about the role of DODs in growing educational inequity may be somewhat conservative. When we adopt Ho's definition of drug-related deaths and reproduce her estimates using our data and methodology, our results in columns 3 and 6 are slightly reduced, though still higher than Ho's, especially in later years. While Ho's education gradient grows from .04 in 1992–1996 to .10 in 2007–2011 (column 1 minus column 4), our estimates start at .05 in 1990 and grow to .36 in 2010 (column 3 minus column 6). The results for white women are somewhat more consistent: our estimated gradients of .02 in 1990 lines up exactly with Ho's estimates while our estimate of .16 in 2010 exceeds Ho's estimate of .09. We are comforted by the fact that Ho's estimates, based on linked data, and ours are broadly consistent. The limitations associated with the linked data she uses are quite different from the limitations associated with the matched data we use. The fact that her estimates line up with ours as well as they do would seem to confirm that neither her results nor ours are an artifact of the data being used.

Table A2 shows the results of a similar exercise using Ho's broader age range of 25+. Our estimates of YLLs for this age range are smaller than those reported in Ho (2017), largely because we use a narrower age range of 25–84. Despite this discrepancy, our estimates of drug-related inequality are quite similar; despite our differing age ranges, methodologies, and data sources, our estimates of the change in the gradient between 1990 and 2010 are .13 and .14 for white men and women, respectively, while Ho's estimates are .21 and .18.

Table A1 Estimates of years of life lost between ages 30–60 by education rank from Ho (2017) and Authors' manuscript

	Low Education			High Education		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ho (2017)	Authors' manuscript	Ho definition	Ho (2017)	Authors' manuscript	Ho definition
Panel A: White men						
1990		.11	.08		.04	.03
1992–1996	.09			.05		
2000		.22	.17		.09	.07
1997–2001	.10			.05		
2002–2006	.14			.08		
2010		.42	.38		.18	.16
2007–2011	.20			.10		
2015		.64	.58		.24	.22
Panel B: White women						
1990		.04	.04		.02	.02
1992–1996	.05			.03		
2000		.10	.08		.04	.04
1997–2001	.05			.03		
2002–2006	.11			.06		
2010		.27	.25		.10	.09
2007–2011	.15			.06		
2015		.38	.34		.12	.11

Notes: Estimates in column 1 and 4 report drug-related YLLs between ages 30 and 60 by level of education from Table 1 of Ho (2017). Figures have been converted to YLLs for the bottom quartile and top three quartiles of the education distribution using the race/cohort/sex-specific distribution of educational attainment from Census data. Columns 2 and 5 report YLLs between ages 25 and 60 using Vital Statistics data combined with Census population data and the methodology Authors' manuscript Columns 3 and 6 report YLLs between ages 25 and 60 using Vital Statistics data matched to Census population data, the methodology of JHSB-18-0280, and the definition of drug-related deaths used by Ho (2017).

Table A2 Estimates of years of life lost after age 25 by education rank from Ho (2017) and Authors' manuscript.

	Low Education			High Education		
	(1) Ho (2017)	(2) Authors' manuscript	(3) Ho definition	(4) Ho (2017)	(5) Authors' manuscript	(6) Ho definition
Panel A: White men						
1990		.29	.17		.13	.08
1992–1996	.22			.16		
2000		.56	.39		.25	.17
1997–2001	.24			.18		
2002–2006	.38			.26		
2010		1.04	.89		.46	.38
2007–2011	.61			.34		
2015		1.57	1.37		.63	.53
Panel B: White women						
1990		.12	.09		.07	.06
1992–1996	.16			.12		
2000		.24	.19		.12	.10
1997–2001	.19			.14		
2002–2006	.35			.22		
2010		.68	.62		.28	.25
2007–2011	.51			.29		
2015		.96	.86		.34	.30

Notes: Estimates in column 1 and 4 report drug-related YLLs after age 25 by level of education from Table 1 of Ho (2017). Figures have been converted to YLLs for the bottom quartile and top three quartiles of the education distribution using the race/cohort/sex-specific distribution of educational attainment from Census data. Columns 2 and 5 report YLLs between ages 25–84 using Vital Statistics data combined with Census population data and the methodology of Authors' manuscript. Columns 3 and 6 report YLLs between ages 25 and 84 using Vital Statistics data combined with Census population data, the methodology of JHSB-18-0280, and the definition of drug-related deaths used by Ho (2017).

Classification of Education

To address the increasingly selected composition of individuals at different levels of education, we categorize educational attainment by rank within the distribution of a birth cohort. Specifically, for each race- and sex-specific cohort born between 1905 and 1990 we calculated the fraction of the population who completed at least 5, 9, 10, 11 and 12 years of education by age 25–34. We categorized education by relative ranks (bottom quartile versus top three quartiles) in the overall distribution, rather than by credentials or years of education, to adjust for distributional changes in educational attainment across cohorts. The sex and race-specific distributions of education for each birth cohort were calculated using the distributions of school completion as reported in the first round of Census data collected after a cohort reached age 25.

In general, the 25th percentile occurs within a category rather than precisely between two categories, even if that category is a single year of education. Thus, to handle the problem that these "border" categories contain persons in both the bottom quartile and second quartile, we divide both the population and death records in this category proportionately.

Years of life lost

The principal measure of mortality used in this paper, years of life lost between ages 25 and 85, is the complement of life expectancy between those two ages. That is YLL=60– ${}_{60}e_{25}$. Life expectancy is calculated using standard demographic methods (Chiang 1984). Life expectancy and years of life lost are standard measures. Some authors work instead with age standardized death rates (ASDR). However, such measures count the death of a 25-year-old as statistically equivalent to the death of an 84 year old, which seems inappropriate. In contrast, The YLL measure we use puts more weight on early deaths to produce an interpretable statistic.

To partition years of life lost into its cause-specific components, we follow a methodology laid out by Andersen, Canudas-Romo, and Keiding (2013) first constructing cause-specific analogues to the empirical mortality rate, probability of death between ages x and x+5, and share of the life-table population dying in age bin x:

$$ {}_5M_x^i = \frac{{}_5deaths_x^i}{{}_5pop_x} $$

$$ {}_5q_x^i = 5 \times \frac{{}_5M_x^i}{1 + (1-a) \times {}_5M_x} $$

$$ {}_5d_x^i = {}_5d_x \times \frac{{}_5q_x^i}{{}_5q_x} $$

Finally, for each five-year age group and each cause *i*, we calculate the sum of life-table deaths in younger age groups that can be attributed to the cause:

$$ {}_xf_{25}^i = \frac{\sum_{s=25}^{x-5} {}_5d_s^i}{\sum_{s=25}^{x-5} {}_5d_s}. $$

In each five-year age bin, years of life are lost by two distinct groups: those who died at earlier ages and thus lose all five years in the age bin, and those who will die in that age bin. The latter group will lose some fraction 1–*a* of the five years in the age bin in each they die. To find cause-specific years of life lost, we simply add up the years lost to people who died of cause *i* in earlier age bins and the years lost to people who die from that same cause in the current age bin:

$$ {}_5\Upsilon_x^i = 5 \times (1 - l_x) \times {}_xf_{25}^i + 5 \times (1-a) \times {}_5d_x^i $$

Total years of life lost in any age range is found by adding up the ${}_5\Upsilon_x^i$ terms corresponding to the age range. The terms can also be summed across *i* to find total years of life lost, or years of life lost due to more aggregated causes of death.

Decomposition of years of life lost by cause

Because years of life lost are additive, they can be used to measure gaps in life expectancy and decompose changes over time into cause-specific factors. As a result, we can directly compare years of life lost between low- and high-education groups:

$$\Gamma_y = {}_{60}^{y}\Upsilon_{25}^{L} - {}_{60}^{y}\Upsilon_{25}^{H}$$

The change in the gap over time will be:

$$\Delta = \Gamma_{2015} - \Gamma_{1990}$$

Finally, the share of the change in the educational gap in years of life lost that is attributable to cause i can be calculated as:

$$\pi^i = \frac{\Gamma_{2015}^{i} - \Gamma_{1990}^{i}}{\Delta}$$

Variability of life table estimates

While the life table estimates we produce (e.g., Γ_y^i) are calculations based on administrative data, two sources of variation must be accounted for when making inferences about comparisons of estimates between two populations or over time. First, population values that enter as denominators for age/sex/race/cause specific death rates are estimated from surveys. In 1990, the decennial census public use microdata sample (PUMS) contains the required data elements based on a 5% sample of the population, and in 2015, the American Community Survey collected data on 1% of the U.S. population. In addition, both surveys included complex sampling design. Second, while counts of deaths are derived from the universe of deaths in the U.S., it is common practice in demographic analysis to treat the realized number of deaths in any given year as the outcome of a stochastic process (Chiang 1984).

We used bootstrap methods to estimate standard errors of our estimates of years of life lost to specific causes of death. For estimates of population counts in 1990 and 2015 we generated 1000 random replicates for each age/sex/race combination assuming population estimates are distributed normally with a mean equal to the administrative total and a standard error calculated using Census bureau formulas. We assume death counts for each age/sex/race/cause combination follow a Poisson distribution with means (and hence, variances) equal to the value reported in administrative data. From the 1000 replicates, we calculated the standard deviation of each life table estimate.

As shown for illustrative causes in Table A3, the size of the confidence intervals implied by these calculations is quite small. The change in gap (difference in difference) estimates discussed in the text are statistically different from zero at a p value 5%.

Table A3 Estimates of change in education gap in years of life lost to selected causes between ages 25 and 84, 1990–2015

	Change in gap	std. error	95% C.I.		Change in gap	std. error	95% C.I.
Non-Hispanic White Men				Non-Hispanic Black Men			
Despair	.94	.03	(.88, 1.00)	Despair	.06	.06	(–.06, .18)
CVD	.61	.04	(.54, .68)	CVD	1.33	.13	(1.07, 1.59)
Cancer	.79	.03	(.72, .86)	Cancer	1.08	.11	(.87, 1.29)
Total	3.26	.07	(3.12, 3.4)	Total	2.98	.23	(2.53, 3.43)
Non-Hispanic White Women				Non-Hispanic Black Women			
Despair	.68	.02	(.64, .72)	Despair	.00	.04	(–.07, .07)
CVD	.38	.03	(.32, .44)	CVD	.68	.11	(.47, .89)
Cancer	1.01	.03	(.95, 1.07)	Cancer	.96	.09	(.79, 1.13)
Total	3.54	.06	(3.42, 3.66)	Total	2.24	.18	(1.88, 2.60)

Causes of death

We estimated the contribution to growing educational gaps in mortality of a wide range of underlying causes, using diagnostic categories of the 9th and 10th Revisions of the International Classification of Diseases, matching using standard concordances (Anderson et al. 2001; Butler 2007; Centers for Disease Control and Prevention 2017, n.d.). Table A5 lists the causes analyzed, and their corresponding ICD-9 and ICD-10 codes.

Competing risks

The cumulative incidence approach we use to account for years lost to a particular cause is similar in spirit to the classic cause elimination life table method (Chiang 1960) but unlike that approach the cumulative incidence approach has the advantage that years lost to each cause add up to the total years lost, as it is invariant to the order in which causes are eliminated in the calculation (Andersen et al. 2013). Cause elimination methods are typically used to study the life expectancy impact of a single cause, because the sum of the impact of multiple causes does not equal the total years of life lost.

To calculate the effect on person years lived of eliminating a cause *i*, this method generates an alternative series of survival probabilities at each age interval *x* to *x+5*, as

$$ {}_5p_x^{(-i)} = {}_5p_x^{\left(1-\frac{{}_5D_x^i}{{}_5D_x}\right)} $$

Then

$$ l_x^{(-i)} = {}_5p_{x-5}^{(-i)} l_{x-5}^{(-i)} $$

gives the fraction surviving to age x in the absence of cause *i* and

$$ {}_5L_x^{(-i)} = \frac{1}{2} * 5 * (l_x^{(-i)} + l_{x+5}^{(-i)}) $$

gives the person years lived in the interval *x* to *x+5* in the absence of cause *i*. The difference between this value and the all cause version of person years lived $\left({}_5L_x \right)$ gives an estimate of the years of life lost to cause *i*. As shown in Table A4, the results for selected causes using this method are qualitatively similar to those using the cumulative incidence measure shown in Table 2. While the assumption of independence of competing risks is required for cause elimination (Andersen et al. 2013), neither approach is immune to the general problem of competing risks in interpreting findings. As outlined carefully by Manton and Stallard (Manton and Stallard 1984) to interpret the cause elimination approach as predictive of a counterfactual in which a cause is removed, one must assume that no other cause-specific death rates change in the absence of the eliminated cause, which is unlikely to be true. Similarly, since the cumulative incidence approach we use attributes all life years lost to the first cause that kills, it must assume, also unlikely, that other causes would not be operative for individuals felled by that first cause. If two causes are closely related, say alcohol related deaths and other drug related deaths, depending on how it were achieved it is possible to imagine the elimination of one cause could result in either an increase or a decrease in the death rate for the other. Thus, both methods are likely to overstate the number of years lost to any single cause of death.

Table A4 Change in years of life lost (cause-elimination method) between ages 25 and 84, by race, sex, education, for selected causes of death, 1990–2015

	Low Ed			High Ed			*Change*	*Share of*
	1990	2015	Change	1990	2015	Change	*in gap*	*total*
NonHispanic White Men								
Despair	.81	1.95	1.14	.47	1.05	.58	*.56*	*17%*
CVD	4.01	2.72	–1.30	3.39	1.83	–1.56	*.27*	*8%*
Cancer	2.70	2.43	–.26	2.53	1.84	–.69	*.42*	*13%*
Total	14.36	14.91	.55	11.07	8.34	–2.73	*3.28*	
NonHispanic Black Men								
Despair	.87	.84	–.03	.58	.58	.01	*–.03*	*–1%*
CVD	4.26	3.59	–.67	4.23	3.05	–1.17	*.50*	*17%*
Cancer	2.97	2.38	–.59	3.08	2.11	–.97	*.38*	*13%*
Total	19.09	16.43	–2.66	17.69	12.09	–5.60	*2.95*	
NonHispanic White Women								
Despair	.24	1.02	.78	.17	.47	.30	*.48*	*14%*
CVD	2.65	1.85	–.79	1.97	1.05	–.92	*.12*	*3%*
Cancer	2.35	2.31	–.04	2.39	1.70	–.70	*.66*	*19%*
Total	8.66	10.91	2.25	7.00	5.71	–1.29	*3.54*	
NonHispanic Black Women								
Despair	.40	.42	.02	.20	.25	.04	*–.02*	*–1%*
CVD	3.76	2.62	–1.14	3.59	2.07	–1.52	*.38*	*17%*
Cancer	2.15	2.19	.03	2.61	2.02	–.59	*.62*	*28%*
Total	11.47	10.77	–.71	11.18	8.23	–2.96	*2.25*	

Table A5: Causes of death, by ICD-9 and ICD-10 code

Cause	ICD-9 (1990)	ICD-10 (2015)
Accidents/undetermined	Other accidental cause (E800-E848,E860-E869,E880-E929); Other undetermined causes (E980.6-E989)	Other accidental cause (V01-X39, X45-X59, Y85-Y86); Other undetermined intent causes (Y15-Y34, Y87.2, Y89.9)
Homicide	Homicide (E960-E969)	Assault(Homicide) (*U01-*U02, X85-Y09, Y87.1)
HIV	HIV/AIDS (*042-*044)	HIV/AIDS (B20-B24)
Diabetes	Diabetes mellitus (250)	Diabetes Mellitus (E10-E14)
Cardiovascular disease	Major cardiovascular disease (390-448)	Major cardiovascular disease (I00-I78)
Alcoholic liver	Chronic liver disease and cirrhosis (571.0-571.3)	Alcoholic liver disease (K70)
Other liver	Chronic hepatitis and biliary cirrhosis, cirrhosis and other unspecified chronic liver disease without mention of alcohol (571.4-571.9)	Other chronic liver disease (K73-K74)
Lower respiratory	Chronic obstructive pulmonary diseases and allied conditions (490-496)	Chronic lower respiratory disease (J40-J47)
Kidney disease	Nephritis, nephrotic syndrome, and nephrosis (580-589)	Kidney disease (N00-N07, N17-N19, N25-N27)
Other infectious disease	Other infectious disease(001-041, 046-139)	Infectious and parasitic diseases (A00-B19,B25-B99)
Mental/behavioral	Mental disorders (290,292-299,300-302,305-319)	Mental/behavioral disorder (F01-F09, F20-F99)
Lung cancer	Malignant neoplasms of trachea, bronchus, and lung (162)	Malignant neoplasms of trachea, bronchus, and lung (C33-C34)
Other Cancer	Neoplasms (140-161, 163-239)	Neoplasms (C00-C32, C37-D48)
Opioids and other drugs	Poisoning by opiates and related narcotics (965.00-965.02, 965.09); Heroin (E850.0); Methadone (E850.1); Other opiates and related narcotics (E850.2) Suicide by drugs, medicaments, and biologicals (E950.0-E950.5); Accidental poisoning by drugs, medicaments, and biologicals (E850.3-E858); Injury undetermined from poisoning by drugs, medicaments, and biologicals (E980.0-E980.5); Alcohol psychoses (291); Alcohol dependence syndrome (303); Drug dependence (304)	Poisoning by drugs with T-codes for heroin and opioids (X40-X44, X60-X64, X85, Y10-Y14 with T-codes T40.1-T40.4) Poisoning by drugs with T-codes not indicating heroin or opioids (X40-X44, X60-X64, Y10-Y14 without T-codes T40.1-T40.4); Mental and behavioral disorders due to alcohol (F10); Mental and behavioral disorders due to other psychoactive substance use (F11-F19)
Non-drug suicide	Other suicides (E950.6-E950.9, E951-E959)	Other intentional self harm (*U03, X65-X84, Y87.0)

Other internal causes	Disease of blood and blood-forming organs (280-289); Endocrine and metabolic diseases and immunity disorders (240-249,251-279); Diseases of the nervous system and sense organs (320-389); Diseases of veins and lymphatics, and other diseases of circulatory system (451-459); Diseases of the respiratory system(460-487,500-519); Diseases of the digestive system (520-562, 564-570, 572-579); Skin and musculoskeletal system and connective tissue (680-739); Genitourinary system, excluding kidney disease (590-629); Pregnancy/childbirth (630-676); Perinatal origin (760-779); Congenital origin (740-759); Symptoms and signs (780-799)	Blood diseases (D50-D89); Other Endocrine/Metabolic disease (E00-E07,E15-E88); Nervous System & Sense Organs (G00-H93); Other circulatory (I80-I99); Other respiratory system (J00-J39,J60-J98, U04); Digestive system (K00-K66, K71-K72, K75-K92); Skin & Musculoskeletal (L00-M99); Genitourinary system (N08-N16,N20-N24,N28-N98); Pregnancy/childbirth (O00-O99); Perinatal origin (P00-P96); Congenital origin (Q00-Q99); Symptoms/signs (R00-R99)
All other deaths	Misadventures during medical care, abnormal reactions, and late complications (E870-879); Drugs, medicaments, and biological substances causing adverse effects in therapeutic use (E930-949); Legal execution (E978); Other legal intervention and late effects of injury due to legal intervention (E970-977); Injury resulting from operations of war (E990-E999)	Legal intervention (Y35, Y89.0); Operations of war and their sequelae (Y36, Y89.1); Complications of medical and surgical care (Y40-Y84, Y88)

ADDITIONAL ANALYSES/TABLES/FIGURES

Effects of trends in behavioral risk factors: Smoking and Obesity

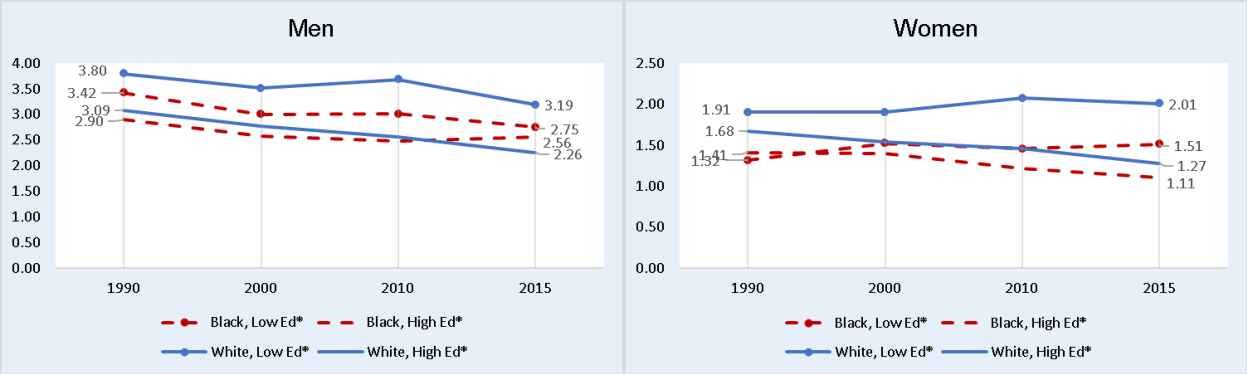
To study the contributions of changing smoking or obesity on mortality we first quantified the prevalence of current and past cigarette smoking and of the standard WHO classifications of body mass index (BMI) by age, sex, race and educational rank using the National Health Interview Surveys (NHIS) from 1990, 2000, 2010, and 2015 obtained from IPUMS (Blewett et al. 2016). Using results on relative mortality risk from the literature, we then calculated fraction of deaths by age for each group attributable to smoking and obesity. Smoking related mortality risks were derived from the estimates of Rogers and colleagues (Rogers et al. 2005) who quantified the excess mortality risk of current and former smokers relative to those who had never smoked. Excess mortality attributable to obesity was derived using estimates from a meta-analysis of 141 studies (Flegal et al. 2013). These estimates found a positive association between mortality risk and class 2 and 3 obesity (BMI>=35) relative to normal weight (18.5<=BMI<25). For each age, we calculated the share of total deaths attributable to smoking or obesity, and using the same methodology discussed above, we calculated the YLL to each in each year studied. In Figure A1, we show estimates of YLL due to smoking. Between 1990 and 2015 YLL due to smoking increased by .2 years for low-education white and black women. All other groups saw YLL due to smoking fall over this period. YLL due to smoking declined among white women in the top 3 quartiles of education by .4 years and by .3 years among black women in the top 3 quartiles. Among men low education whites and blacks both experienced declines of .6 years; white men in the top 3 quartiles saw a decline of approximately .8 years and black men in the top 3 quartiles saw a decline of approximately .3 years.

In Figure A2, we show estimates that YLL due to obesity increased for all groups. The combined rates of severe and morbid obesity (BMI>=35) increased for every group as defined by sex, race, age, and educational rank, with some groups (white men 75–84) experiencing ten-fold increases in the rate. Applying Flegal's (2013) estimate of the hazard ratio (1.34) of elevated mortality risk for this group relative to those in the normal range is applied over the entire period, we estimate that YLL due to obesity increased by .4 years among black women in the lowest educational quartile, by .5 years among white women in the lowest quartile, by approximately .3 years among black women in the top 3 quartiles of education and by .2 years among more highly educated white women. Men exhibited similar patterns as women.

In Table A6 we show the effect of these trends on the educational gaps in life expectancy. In every case but one, changing behavioral risk factors widened the gap between the lowest education quartile and the top three quartiles. For women, changes due to smoking had a larger effect on the education gap than changes due to obesity. Among white women, for example, the gap in smoking related YLL grew from .23 years in 1990 to .74 years in 2015, while the gap in obesity related YLL grew from .08 to .38 years. For white men, smoking was also a bigger contributor, but the contribution of obesity to the LE gap increased more, from .05 years to .32 years. Analyzing different data, earlier researchers came to broadly similar conclusions with respect to smoking among whites (Ho and Fenelon 2015); however, they did not study blacks. We found for black men, trends in smoking related mortality actually narrowed the education gap from .52 years in 1990 to .19 years in 2015 while trends in obesity related mortality widened the gap slightly.

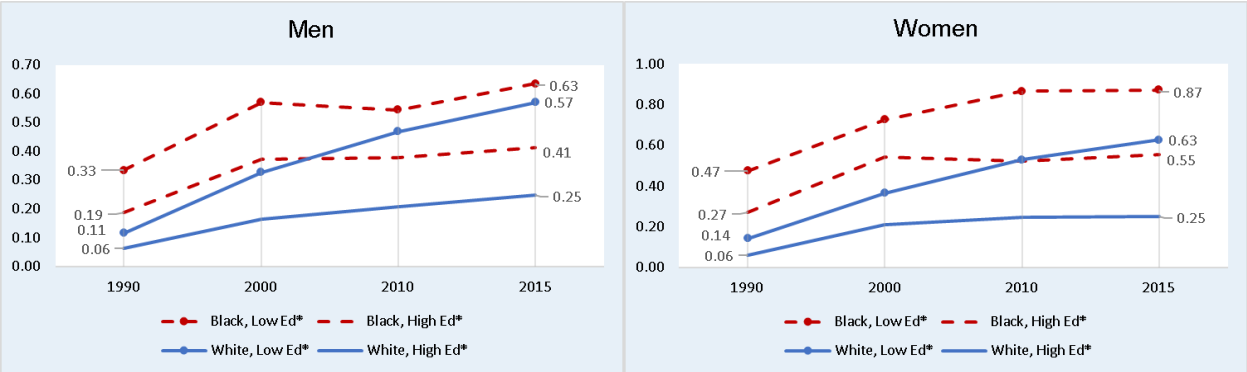
While the total effects of smoking and obesity contributed to the increased inequity in LE for most demographic groups studied, the contribution was modest. For example, among white women for whom the contributions of these behaviors were the largest, changes in obesity rates accounted for .5 years and changes in smoking rates accounted for .3 years, out of the 3.94 year increase in educational inequity in YLL. For white men .2 years were lost to obesity, and .3 years to smoking out of an increase of 3.62 YLL; for black women, .5 to obesity and .1 to smoking out of a total increase of 2.51 YLL; and for black men, .1 year was lost to smoking rate changes out of total increase of 3.09 years, while changes in obesity lowered YLL by –.3 years.

Figure A1: Years of Life Lost to Smoking, age 25–84 by race, sex, and education, 1990–2015



Source: Population data sources: U.S. Census Bureau, Centers for Disease Control and Prevention
*Low Ed refers to bottom 25 percent of education distribution, and High Ed refers to the top 75% of the education distribution within cohorts defined by sex, race/ethnicity, year of birth. Data on education distribution derived from U.S. decennial Census, 1940–2000 and American Community Survey, 2005–2015.

Figure A2: Years of Life Lost to Obesity, age 25–84 by race, sex, and education, 1990–2015



Source: Population data sources: U.S. Census Bureau, Centers for Disease Control and Prevention
*Low Ed refers to bottom 25 percent of education distribution, and High Ed refers to the top 75% of the education distribution within cohorts defined by sex, race/ethnicity, year of birth. Data on education distribution derived from U.S. decennial Census, 1940–2000 and American Community Survey, 2005–2015.

Table A6: Education gaps in YLL due to behavioral risk factors

	YLL due to obesity			YLL due to smoking		
	Low Education*	High Education*	Education Gap	Low Education*	High Education*	Education Gap
Black Women						
1990	.47	.27	–.20	1.32	1.41	.09
2015	.87	.55	–.32	1.51	1.11	–.40
White Women						
1990	.14	.06	–.08	1.91	1.68	–.23
2015	.63	.25	–.38	2.01	1.27	–.74
Black Men						
1990	.33	.19	–.15	3.42	2.90	–.52
2015	.63	.41	–.22	2.75	2.56	–.19
White Men						
1990	.11	.06	–.05	3.80	3.09	–.71
2015	.57	.25	–.32	3.19	2.26	–.93

Source: Population data sources: U.S. Census Bureau, Centers for Disease Control and Prevention
*Low Education refers to bottom 25 percent of education distribution, and High Education refers to the top 75% of the education distribution within cohorts defined by sex, race/ethnicity, year of birth. Data on education distribution derived from U.S. decennial Census, 1940–2000 and American Community Survey, 2005–2015

REFERENCES

Andersen, Per Kragh, Vladimir Canudas-Romo, and Niels Keiding. 2013. "Cause-Specific Measures of Life Years Lost." *Demographic Research* 29:1127–52.

Anderson, Robert N., Arialdi M. Miniño, Donna L. Hoyert, and Harry M. Rosenberg. 2001. "Comparability of Cause of Death between ICD-9 and ICD-10: Preliminary Estimates." *National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 49(2):1–32.

Blewett, Lynn A., Julia A. Rivera Drew, Miriam L. King, and Kari Williams. 2016. *IPUMS Health Surveys: National Health Interview Survey, Version 6.2*. Minneapolis, MN: University of Minnesota.

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.

Butler, Rhonda R. 2007. "ICD-10 General Equivalence Mappings: Bridging the Translation Gap from ICD-9." *Journal of AHIMA* 78(9):84–86.

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 of the United States of America* 112(49):15078–83.

Case, Anne and Angus Deaton. 2017. *Mortality and Morbidity in the 21st Century*. Brookings Institution.

Centers for Disease Control and Prevention. 2017. "ICD - ICD-10-CM - International Classification of Diseases, Tenth Revision, Clinical Modification." Retrieved July 29, 2017 (https://www.cdc.gov/nchs/icd/icd10cm.htm).

Centers for Disease Control and Prevention. n.d. "ICD - ICD-9-CM - International Classification of Diseases, Ninth Revision, Clinical Modification." Retrieved July 29, 2017 (https://www.cdc.gov/nchs/icd/icd9cm.htm).

Chiang, Chin Long. 1960. "A Stochastic Study of the Life Table and Its Applications: Ii. Sample Variance of the Observed Expectation of Life and Other Biometric Functions." *Human Biology* 32(3):221–38.

Chiang, Chin Long. 1984. *The Life Table and Its Applications*. Malabar, Fla: R.E. Krieger Pub. Co.

Flegal, Katherine M., Brian K. Kit, Heather Orpana, and Barry I. Graubard. 2013. "Association of All-Cause Mortality With Overweight and Obesity Using Standard Body Mass Index Categories: A Systematic Review and Meta-Analysis." *JAMA* 309(1):71.

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

Ho, Jessica Y. 2017. "The Contribution of Drug Overdose to Educational Gradients in Life Expectancy in the United States, 1992–2011." *Demography* 54(3):1175–1202.

Ho, Jessica Y. and Andrew Fenelon. 2015. "The Contribution of Smoking to Educational Gradients in U.S. Life Expectancy." *Journal of Health and Social Behavior* 56(3):307–22.

Manton, Kenneth G. and Eric Stallard. 1984. *Recent Trends in Mortality Analysis*. Academic Press.

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

National Bureau of Economic Research. n.d. "Mortality Data -- Vital Statistics NCHS' Multiple Cause of Death Data, 1959-2015." Retrieved July 31, 2017 (http://www.nber.org/data/vital-statistics-mortality-data-multiple-cause-of-death.html).

National Center for Health Statistics, Office of Analysis and Epidemiology. 2009. *The National Health Interview Survey (1986-2004) Linked Mortality Files, Mortality Follow-up through 2006: Matching Methodology*. Centers for Disease Control and Prevention.

National Center for Health Statistics, Office of Analysis and Epidemiology. 2018. *The Linkage of National Center for Health Statistics Survey Data to the National Death Index — 2015 Linked Mortality File (LMF): Methodology Overview and Analytic Considerations.* Centers for Disease Control and Prevention.

Olshansky, S. Jay, Toni Antonucci, Lisa Berkman, Robert H. Binstock, Axel Boersch-Supan, John T. Cacioppo, Bruce A. Carnes, Laura L. Carstensen, Linda P. Fried, and Dana P. Goldman. 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.

Rogers, Richard G., Robert A. Hummer, Patrick M. Krueger, and Fred C. Pampel. 2005. "Mortality Attributable to Cigarette Smoking in the United States." *Population and Development Review* 31(2):259–92.

Rostron, Brian L., John L. Boies, and Elizabeth Arias. 2010. "Education Reporting and Classification on Death Certificates in the United States." *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research* (151):1–21.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. "Integrated Public Use Microdata Series: Version 7.0."

Sasson, Isaac. 2016. "Diverging Trends in Cause-Specific Mortality and Life Years Lost by Educational Attainment: Evidence from United States Vital Statistics Data, 1990-2010" edited by J. Patra. *PLOS ONE* 11(10):e0163412.

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.

Sorlie, P. D. and N. J. Johnson. 1996. "Validity of Education Information on the Death Certificate." *Epidemiology (Cambridge, Mass.)* 7(4):437–39.