Increasing mortality of white Americans, a systematic deviation from Gompertz law, and a trend break in cohort health

Nicholas Reynolds*

November 16, 2018

Abstract

I suggest that recent increases in the mortality rate of white Americans are caused by a decline in the health of cohorts born since the middle of the last century, relative to the trend for earlier-born cohorts. In support of this theory, I present evidence of systematic deviations of the mortality rate of white Americans from its usual log linear relationship with age. In each year since 1985, log mortality rates of white men and women exhibit a sharp trend break for men born after 1947 and for women born after 1950 — leading to higher mortality than would be predicted by the log linear Gompertz curve. This trend break can explain the timing of increases in the mortality rate of white men and women aged 45 to 54, and the staggered timing of mortality increases across other age groups. To understand the cause of the recent increase in the white mortality rate, we must understand the sources of this cohort-specific decline in health.

Keywords: mortality | cohorts | white Americans

JEL Classifications: I12, I14, J11

*Nicholas_Reynolds@brown.edu, Economics Department and Population Studies and Training Center, Brown University. Providence, RI 02912. I thank Anna Aizer, Kenneth Chay, Andrew Foster, and Emily Oster for comments and encouragement throughout the project. I also benefited from comments from Angus Deaton, Stefan Hut, Stelios Michalopoulos, and David Weil and from participants at the Applied Microeconomics Lunch and Health Economics Breakfast at Brown University. I appreciate fellowship support from an NIH Training Grant (T32 HD007338) and the Fellowship on the Economics of an Aging Labor Force from the National Bureau of Economic Research, supported by the Sloan Foundation. And am grateful to the Population Studies and Training Center at Brown University, which receives funding from the NIH (P2C HD041020) for general support.

1 Introduction

Case and Deaton (2015) document a disturbing trend of increasing mortality among 45 to 54 year old non-Hispanic white Americans since 1999. This mortality increase could be due to some factor that changed between 1999 and 2012, such as the increasing availability of prescription opioids and heroin. Alternatively, the underlying health of white Americans who were 45 to 54 in the two periods may differ. White Americans born between 1958 and 1968 may be less healthy on average than those born between 1945 and 1955.

Cross-cohort differences in health would leave a distinct imprint on mortality rates in other years as well. I document such a sharp cohort-specific pattern. The mortality rate of white men and women has deviated, in a way that is systematically related to cohort, from it's usual log-linear relationship with age, the so-called Gompertz curve (Gompertz, 1825). In each year since the 1980s those born after the 1947 cohort for men and 1951 cohort for women have higher log mortality rates than the Gompertz curve would predict. Put another way, in each year of data the log mortality rates of white Americans born between 1930 and 1965 are better approximated by a piece-wise linear, trend break model — with the trend break occurring at or near the 1951 cohort for women and the 1947 cohort for men.

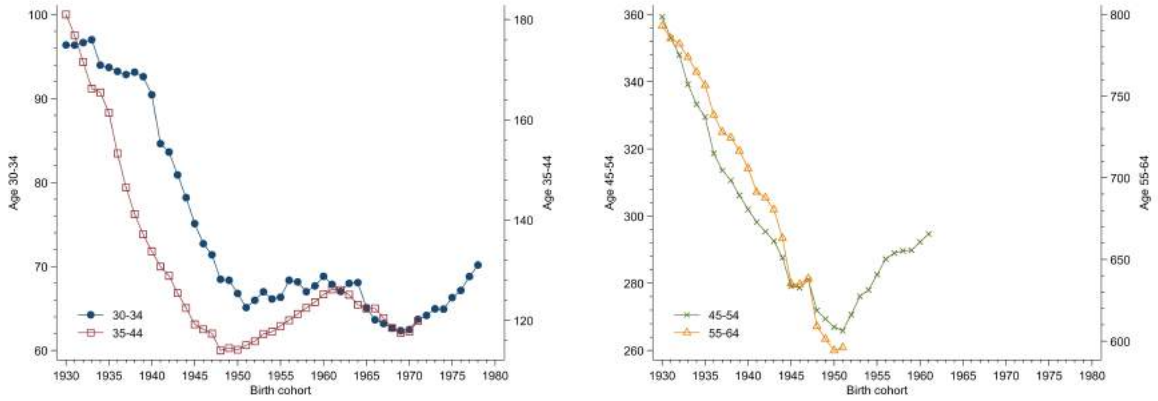
The fact that the cohort specific trend break is evident across years, starting in the 1980s when affected cohorts were in their 30s and persisting to older ages and recent years, suggests it reflects persistent differences in cohort health. White Americans born in 1960 appear to be less healthy in some fundamental sense than those born in 1950, and have been so since at least their mid-30s. Therefore, contemporaneous factors in the 1990s, such as increasing availability of prescription opioids and heroin, likely contributed to increased mortality but did not initiate it.

In this paper, I do not take a stand on a particular underlying cause of these health differences, but instead document a sharp cohort-based pattern that any potential explanation must match. Case and Deaton (2017) posit a preliminary theory — "cumulative disadvantage" — in which worsening opportunities at labor market entry for whites with low levels of education triggers progressively worse outcomes, culminating in an increased likelihood of untimely death. This explanation is cohort-based — suggesting that individuals who entered the labor market in a particular year should have elevated mortality across years. Other potential explanations focusing on early-life factors such as maternal smoking or pollution exposure, would also generate a distinct cohort-specific pattern.

Figures 1 and 2 present preliminary, suggestive evidence that cohort-specific factors

may be an important cause of recent mortality increases of white Americans. Figure 1 shows the mortality rate of white women by year of birth, separately for different age groups. The mortality rates of 30-34 , 35-44, 45-54, and 55-64 year olds all declined steadily between the 1930 and 1950 cohorts, but then near the 1950 cohort began to suddenly flatten or increase. [1]

Figure 1: Mortality rate of white women — deaths per 100,000

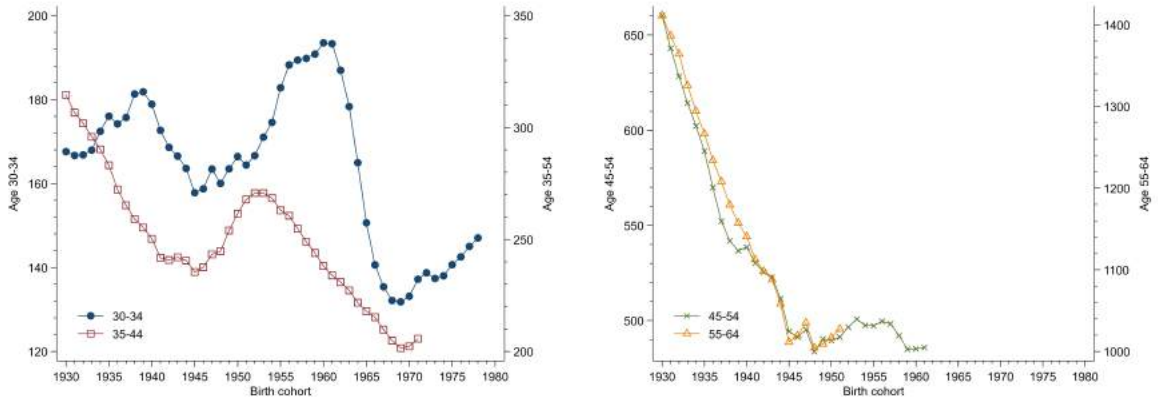


Note: Each series shows the crude death rate of white women, experienced by different birth cohorts in a given age group, in deaths per 100,000 people per year.

Figure 2 presents similar data for white men. The cohort pattern is less clear for men in their 30s and early 40s — potentially due to the large and unusual impact of the AIDS epidemic on these age groups that varied substantially by year. However, the mortality rates of 45-54 and 55-64 year old men both show consistent declines between the 1930 and 1947 cohort, and then flatten suddenly.

[1] Because Hispanic origin is not reported consistently in death data until 1997, I focus on the mortality rate of all whites regardless of Hispanic origin. If the phenomenon I document is truly concentrated only among non-Hispanic whites then my results will be attenuated estimates of this true effect.

Figure 2: Mortality rate white men — deaths per 100,000



Note: Each series shows the crude death rate of white women, experienced by different birth cohorts in a given age group, in deaths per 100,000 people per year.

In this paper, I present novel evidence of the importance of cohort effects in recent mortality trends of white Americans. My approach can be interpreted as a novel solution to the so-called "age-period-cohort" problem. If one imposes log-linearity in the impact of age on mortality in any given year, then deviations from this linearity can be interpreted as being driven by cohort effects. This approach therefore imposes a parametric assumption on the impact of age on mortality in each year, but is flexible with regards to year effects and even allows the log-linear age-effects to vary by year. It also yields an intuitive overidentification test — if the non-linearites are truly driven by cohort-specific factors then the trend break should occur at the same cohort in each year.

I first present graphical evidence of this cohort-specific trend break. For each year of data I remove a linear trend in age from log mortality. I then plot these detrended log mortality rates by cohort, for different sets of years. The shape of the detrended log mortality rates by cohort is similar across years — for women declining between the 1930 and 1951 cohort and then increasing; and for men declining between the 1930 and 1947 cohort and then increasing.

I then provide more formal evidence of the cohort-specific trend break using the framework of Hansen (2000) for estimation and testing of structural breaks of unknown location. For each year, I estimate a separate trend break model, and allow the data to determine at which cohort a trend break occurs and the size of that break. Across years, I estimate that the trend break in the log mortality rate of white men occurs near the 1947 cohort. For white women the trend breaks are persistently estimated to be near the 1950 cohort.

In the next section, I estimate a single structural break model across all years of data. This model imposes that the cohort-specific break occurs at the same cohort in all years, but allows the size of the break to vary. I again estimate that the trend break in log mortality occurs at the 1947 cohort for men and the 1951 cohort for women. The location of this estimated cohort break is robust to including increasingly flexible age-by-year controls, up to even allowing a separate cubic in age for each year.

The smallest estimate across specifications suggests a cross-year average size of the trend break in log mortality of .025 for white men and .026 for women. These model estimates suggest for example that the 1957 cohort of white men has had on average an approximately 25 percent higher mortality rates than they would have absent the trend break. Similarly, it implies that the 1960 cohort of white women had on average approximately 26 percent higher mortality rates than what would be predicted by the trend for preceding cohorts.

I next use the model to assess whether this cohort-specific pattern can explain the timing of recent increases, and stagnating improvements, in mortality by age. Case and Deaton (2015) focus on the mortality rate of non-Hispanics whites aged 45-54, and document a sudden increase starting in 1999. Gelman and Auerbach (2016) suggest that the changing age-structure of this group impacted mortality rates, and that the mortality rate of non-Hispanic white women increased steadily between 1999 and 2012, while that of men suddenly stagnated.

However, the timing of trend breaks in the age-adjusted mortality rates of this group at other ages differs. The age-adjusted mortality rate of white women aged 35-44 suddenly began to increase in 1991 and continued to do so until the early 2000s; while that of those aged 55-64 was declining until 2010 before suddenly beginning to increase. The timing of these mortality increases appear amenable to a cohort-based explanation — with "unhealthy" post-1951 cohorts driving mortality increases first at young ages and subsequently at older ages, as they age into different age bins.

I use the previously estimated model to assess more formally whether the timing of mortality increases by age can be explained by the cohort-specific trend break. I use the estimated model to simulate mortality rates by year for age bins, and compare the simulated rates to observed mortality rates. For white women the simulated rates closely track the true series, and in particular can match the staggered timing of increases by age bin —- suggesting that cross-cohort differences in health are a plausible underlying cause of the recently documented mortality increases for white women.

The year-over-year pattern of the mortality rate of white men by age-bin show less

prima facie evidence of cohort effects. However, the cohort-specific trend break model also fits these patterns relatively well for men. In particular, the model predicts the timing of stagnation in mortality improvements at ages 45-54 in 1997, and at age 55-64 for near 2010. For 35-44 year olds the AIDS epidemic appears to have left a noticeable imprint on the year-over-year trend in mortality rates.

In addition to Case and Deaton (2015), my paper contributes to a larger literature in demography documenting patterns in the all-cause mortality rate of white Americans. Yang (2008) and Masters et al. (2014) emphasize the importance of a cohort-perspective and estimate age-period-cohort models of mortality whose results suggest that mortality declines since 1960 are largely driven by cohort factors. I present evidence that recent increases in mortality also exhibit a clear cohort pattern. I also present a new approach to documenting this cohort pattern, which does not rely on the standard additively-separable age-period-cohort model.

A large literature also examines widening gaps in mortality by education level, see for example Meara et al. (2008) and Olshansky et al. (2012). Masters et al. (2012) argues that widening differences in cohort effects can explain this widening gap in all-cause mortality and mortality risk from heart disease and lung cancer. My study differs from this literature, and follows Case and Deaton (2015), by focusing on an increase in the mortality rate of the aggregate population of white men and women of all education levels.

2 Data

To calculate mortality I use the Multiple Cause of Death File from the Center for Disease Control and intercensal population estimates from the Census Bureau and the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute. Using these sources I calculate mortality and mid-year population by age, race, and sex. I then calculate "crude death rates," as reported for example by the CDC.

Hispanic-origin was not reported on death certificates in all states until 1997, and is therefore not consistently recorded in the Multiple Cause of Death File until that year. I therefore primarily report mortality rates for all whites, including Hispanics and non-Hispanics. In some supplementary analysis I report post-1997 mortality rates for non-Hispanic and Hispanic whites separately.

3 Graphical evidence

In this section, I present graphical evidence of break in the cross-cohort mortality trend, evident across years. I restrict my sample to log mortality rates for 30-84 year old white men and women, born between 1930 and 1970, for the years 1968 to 2014. I conduct all analysis separately for men and women.

For each year of data I remove a linear trend in age, by estimating the following model:

$$ln(mort_{apc}) = \beta_p a + \mu_p + \epsilon_{apc} \tag{1}$$

I then form "detrended log-mortality" estimates as the residuals from this regression.

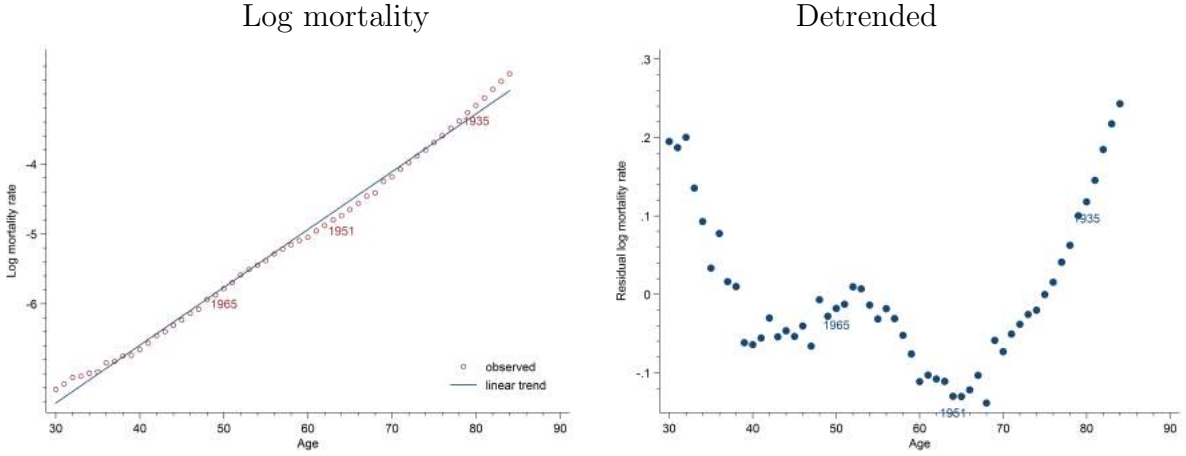
Figure 3 outlines this procedure graphically using 2015 data as an example. The procedure is shown for the log mortality rate of white women in the top panel and for white men in the bottom panel. The dots in the left graphs show raw log-mortality data by age, and the line shows the predicted log mortality from the log-linear model estimated from equation 1. The right panel shows the residuals, ie. the difference between the raw data and the model shown in the left panel.

The graphs show first that log mortality is remarkably linear in the ages examined — the dots in the left graphs are remarkably close to the linear trend. It shows second that, at least in 2015, there are systematic deviations from this linear relationship. White women aged approximately 65, and therefore born in approximately 1951, have particularly low log mortality than predicted by the linear trend. Proceeding in either direction from this minimum to younger or older cohorts, there is a trend in the residual mortality — each subsequent cohort has higher mortality in relation to the linear trend prediction. A similar pattern is evident for white men with the 1947 cohort having the minimum residual.

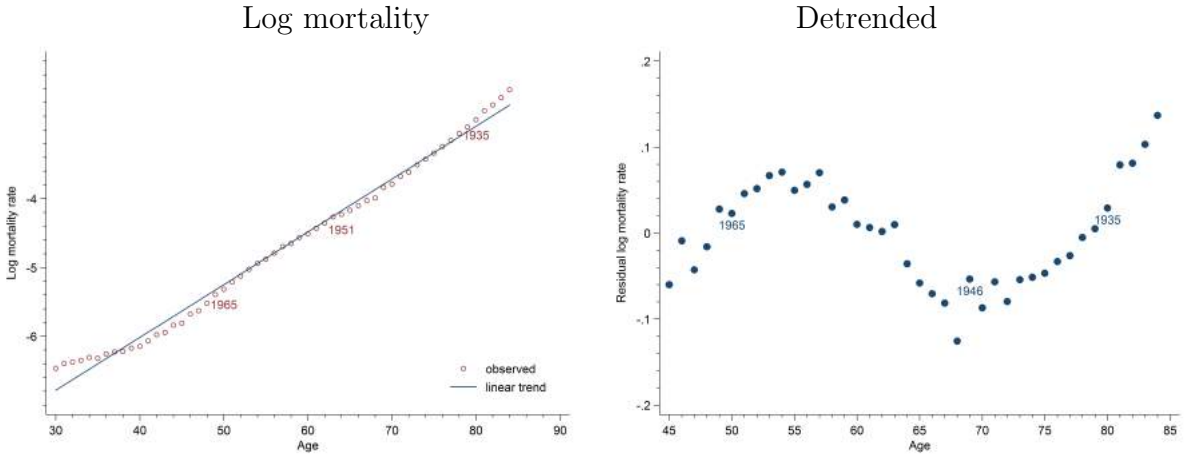
These patterns suggest that the 2015 log mortality rates for white men and women over 45 would be better fit by a piecewise linear, trend break model, with a trend break near the 1951 cohort form women and the 1946 cohort for men.

Figure 3: Graphical depiction of detrending of log mortality rates

White women, 2015



White men, 2015

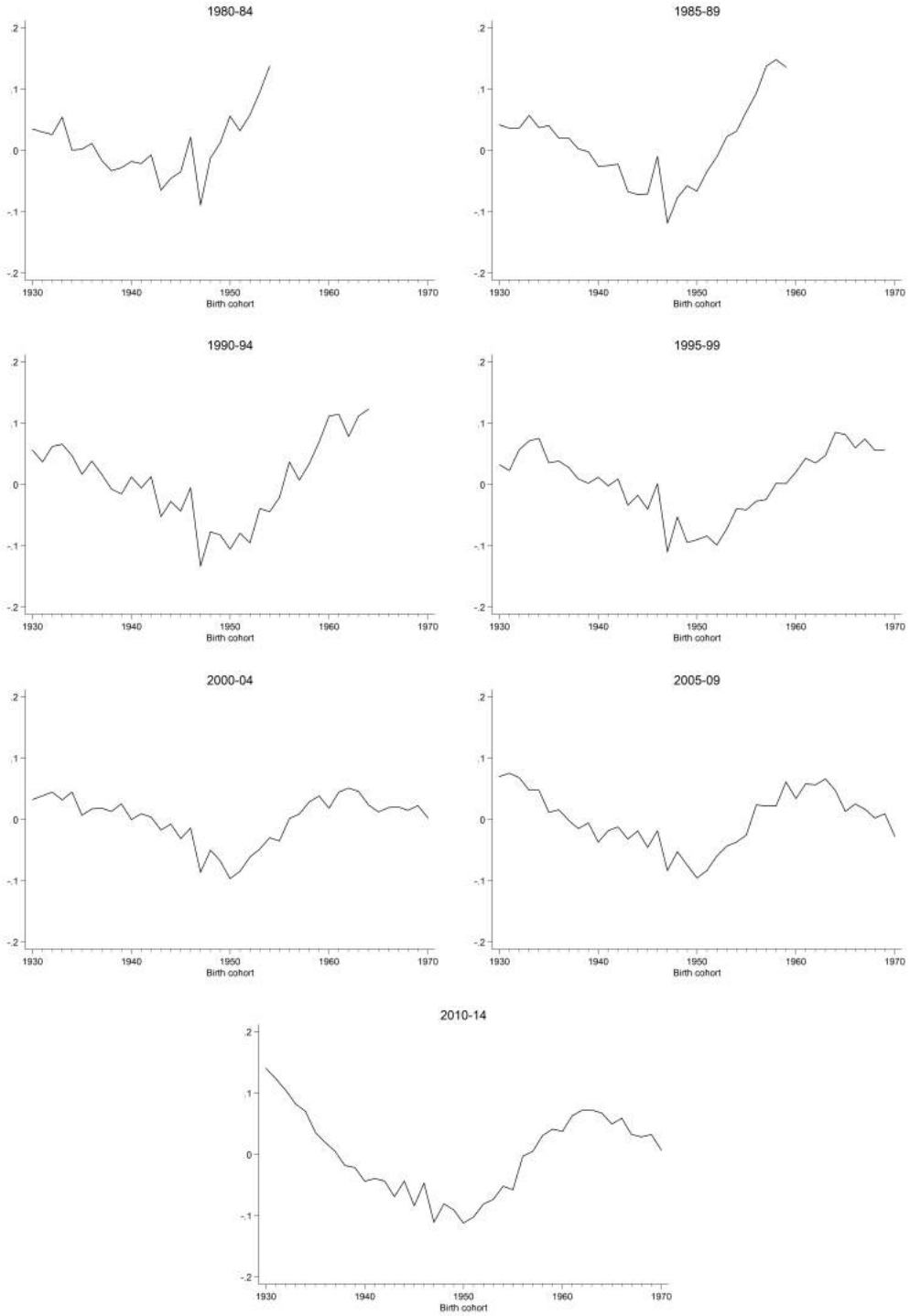


Note: These figures show the log mortality rate of white men after the removal of a linear trend in age. Each plot shows the average across the listed 5-years of the residuals from a regression based on equation 1.

I then take the average in 5 year bins, for each cohort, of the residuals from this regression. This will provide evidence that the deviations shown above for 2015, have the same pattern by cohort in other years — starting in the 1980s. If the normal impact of aging on mortality is linear in each year, then a break in this detrended series — at the same cohort for each year — can be interpreted as evidence of a trend break in underlying cohort health.

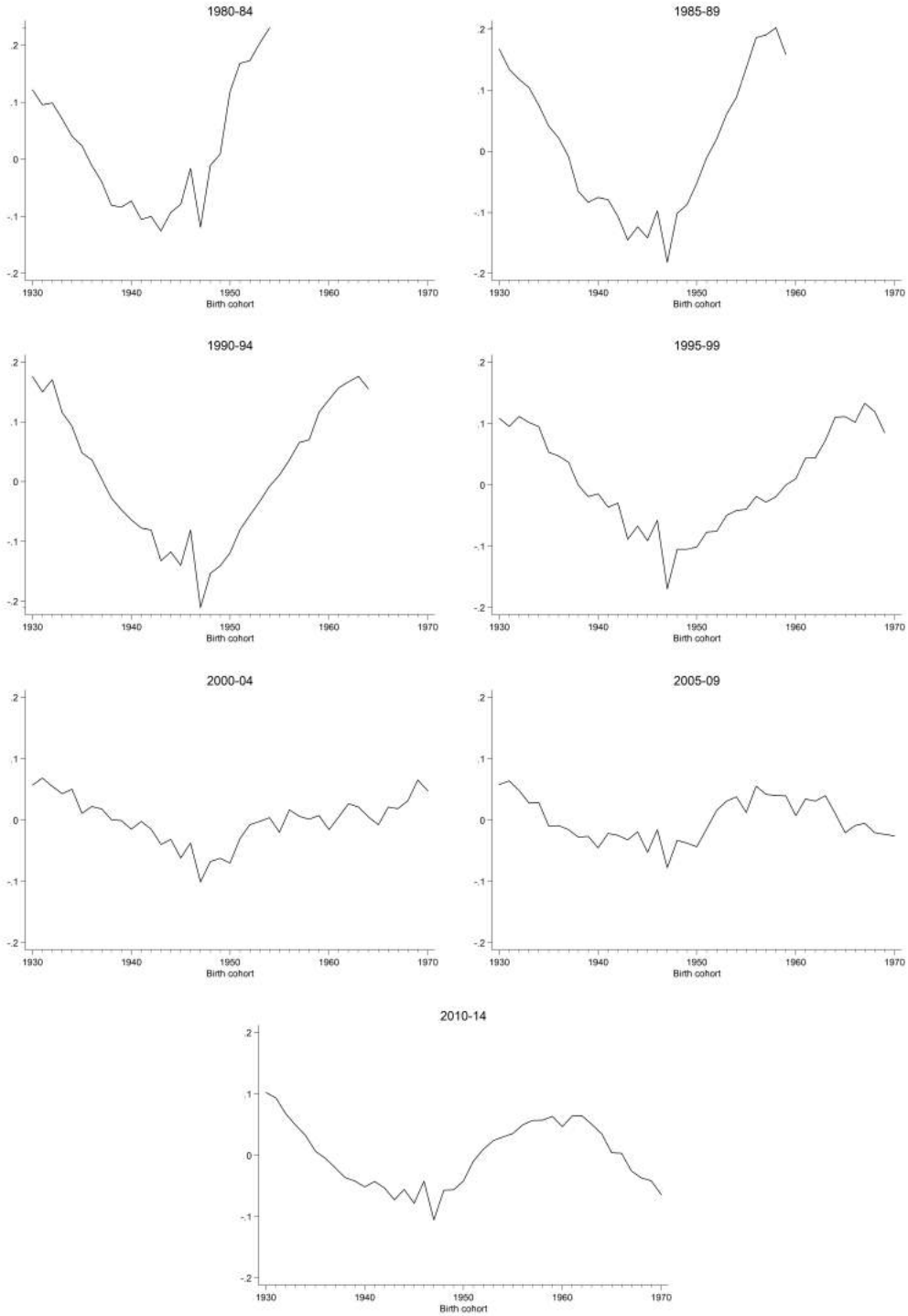
Figure 4 and 5, show these detrended log mortality rates for white women and men respectively. For each 5-year bin between 1980 and 2014, there is striking visual evidence of a trend break. For women the break appears consistently near the 1950 cohort. For men it occurs near the 1947 cohort. In later years, the detrended mortality rate declines again for early 1960s cohorts — but is still far above the pre-1950 trend.

Figure 4: Detrended log mortality rates, white women



Note: These figures show the log mortality rate of white men after the removal of a linear trend in age. Each plot shows the average across the listed 5-years of the residuals from a regression based on equation 1.

Figure 5: Detrended log mortality rates, white men



Note: These figures show the log mortality rate of white men after the removal of a linear trend in age. Each plot shows the average across the listed 5-years of the residuals from a regression based on equation 1.

4 Structural break tests by year

Next, I use the structural break estimation and testing framework of Hansen (2000) to provide formal evidence of a break in the cross-cohort mortality trend.

Consider the following model of log-mortality:

$$ln(mort_{apc}) = \beta_a^p a + \beta_c^p c + \delta^p \cdot (\gamma^p - c) \cdot 1_{c \geq \gamma^p} + \mu^p + \epsilon_{apc} \tag{2}$$

where a denotes age, p denotes period (eg. year), c denote cohort; and $ln(mort_{apc})$ denotes the log-mortality rate of individuals age a, in period p, and from cohort c.

Because age and cohort are perfectly collinear in each year of data the linear trends in age and cohort, β_a^p and β_a^p, are not separately identified. However, the following transformed model is identified:

$$ln(mort_{apc}) = \tilde{\beta}^p a + \delta^p \cdot (\gamma^p - c) \cdot 1_{c \geq \gamma^p} + \mu^p + \epsilon_{apc} \tag{3}$$

where $\tilde{\beta}^p = \beta_a^p - \beta_c^p$. And the location, γ^p, and size δ^p, are identified.

I estimate the model separately for each year, by concentrated, weighted least squares, following the methodology in Hansen (2000). Algorithmically, this amounts to looping through different assumed values of the trend break location γ^p, and selecting the location with the lowest sum of squared, weighted residuals.

Following Hansen (2000) I invert the following likelihood ratio statistic to form 99 percent confidence intervals for γ^p:

$$LR(\gamma^p) = n \frac{S(\gamma^p) - S(\hat{\gamma^p})}{S(\hat{\gamma^p})} \tag{4}$$

where n denotes the number of observations, $S(\gamma^p)$ is the weighted sum of squared residuals from estimating equation 3 with the trend break location fixed at a given γ^p, and $S(\hat{\gamma^p})$ is the sum of squared residuals with the *estimated* break location $\hat{\gamma^p}$. I constuct the 99 percent confidence interval as those values of γ^p such that $LR(\gamma^p) <= 10.35$, the critical value given in Hansen (2000). While I allow for heteroskedasticity in inference on other parameters, this test requires homoskedasticity.

Hansen (2000) also suggests that inference on δ^p is unaffected by treating γ^p as unknown. I therefore form confidence intervals for δ^p using the standard formula for weighted least squares.

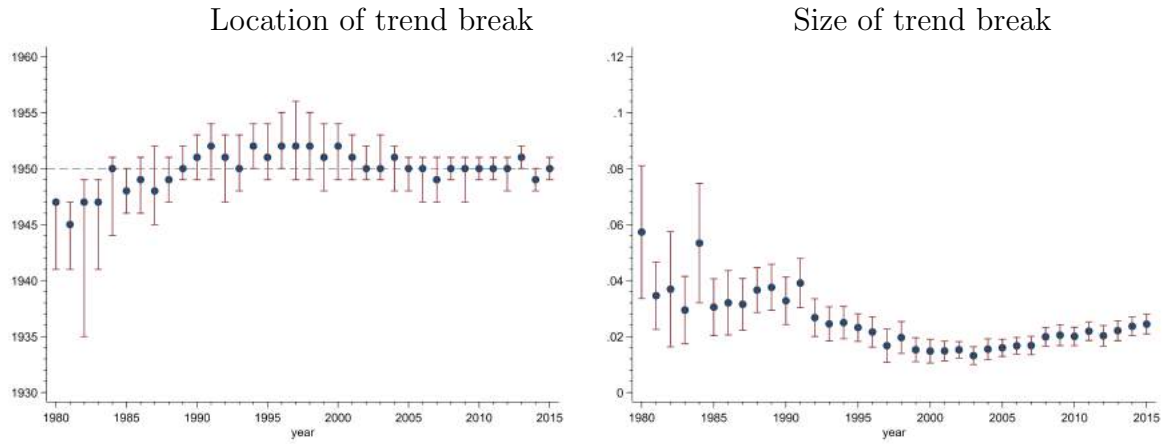
I implement the above described approach, restricting the sample to the years 1980-

2015, ages 30-75, and cohorts 1930 to 1965. For each year, I restrict the location of the cohort-break γ^p to not be one of the 3 youngest or oldest cohorts in the sample.

Figure 5 shows the results across years for the log mortality rate of white women. The left panel shows for each year of data the estimated location of the cohort specific trend break, $\hat{\gamma^p}$, as well as the 99 percent confidence intervals. For the first few years I examine, 1980-1984, the trend break is estimated to be located near the 1946 cohort. For all following years from 1985 to 2015, the estimated trend break is between 1948 and 1952, and the confidence interval includes 1950. The right panel shows for each year the size of the estimated trend break, $\hat{\delta^p}$. The size of the trend break in log mortality is initially near .04 and then declines to nearer .02 after 1990.

I also use the asymptotically valid bootstrap procedure suggested in Hansen (1999) to test the null hypothesis of no break, eg. $H_0 : \delta^p = 0$. For all models, the value of the F-type statistic for the true data is larger than that calculated in all of the 1000 bootstrap repetitions that I run — implying a p-value of less than .001 for the null of no break.

Figure 6: Trend break and size estimates, log mortality, white women



Note: This figures show the results of estimation of the trend break model in equation 6, with the log mortality rate of white women by age, year and cohort as the dependent variable. A separate model is estimated for each year by concentrated, weighted least squares, following the approach outlined in Hansen (2000). The sample includes ages 30-75, and cohorts 1930 to 1965. The left panel shows for each year of data the estimated location of the cohort specific trend break, $\hat{\gamma^p}$, as well as the 99 percent confidence interval calculated by inverting a likelihood ratio test statistic. The right panel shows for each year the size of the estimated trend break, $\hat{\delta^p}$.

Figure 6 shows analagous results for the log mortality rate of white men. The location of the trend break is slightly less stable for men than for women. In the first few years, 1980-1983, the trend break is estimated to be located near the 1943 or 1944 cohort. For

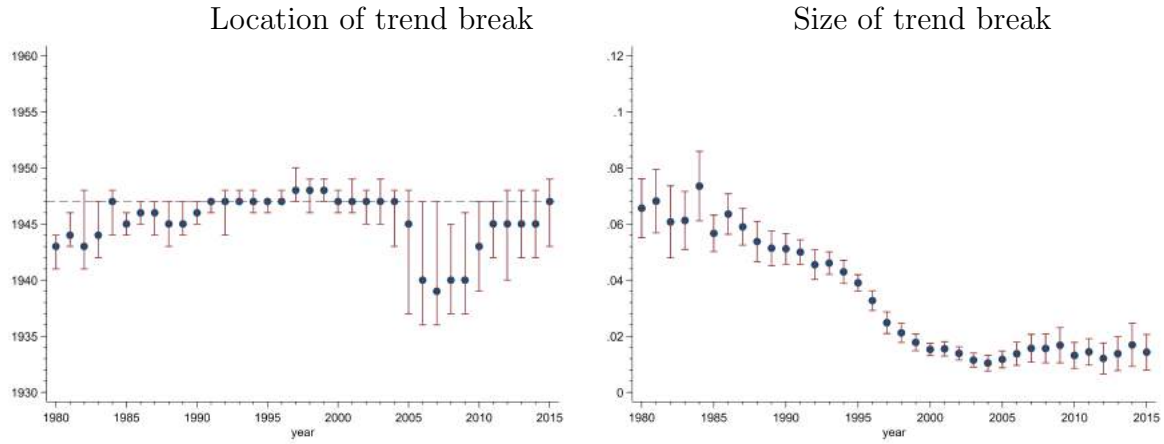
a middle period from 1984 to 2004, the estimated trend break is precisely estimated and consistently located between 1945 and 1947, and the confidence interval includes 1947.

After 2004, the location becomes less precisely estimated. Between 2004 and 2009 the point estimate drops to near 1940 and the confidence interval no longer includes 1947. For 2009 to 2015, the point estimate jumps to 1945 and the confidence interval again includes 1947.

The size of the trend break for log mortality is initially large for men than previously shown for women, it is above .06 for the first 7 years examined. In later years however it falls to near .-02, similar to that estimated for women.

For men in all years, the implied p-value from the bootstrap procedure is less than or equal to .001 for the null of no break.

Figure 7: Trend break and size estimates, log mortality, white men



Note: This figures show the results of estimation of the trend break model in equation 6, with the log mortality rate of white men by age, year and cohort as the dependent variable. A separate model is estimated for each year by concentrated, weighted least squares, following the approach outlined in Hansen (2000). The sample includes ages 30-75, and cohorts 1930 to 1965. The left panel shows for each year of data the estimated location of the cohort specific trend break, $\hat{\gamma^p}$, as well as the 99 percent confidence interval calculated by inverting a likelihood ratio test statistic. The right panel shows for each year the size of the estimated trend break, δ^p.

5 A single structural break model

The above analysis allowed the location of the trend break in log mortality to vary by year. The results appear to suggest that the trend break occurs in the same cohort across years. Therefore, in this section I estimate a single structural break model across years, which imposes that the cohort-specific break occurs at the same cohort in all years. Guided by the

above results, I allow the size of the break to vary across years. Estimation of this single model allows me to probe the robustness of the trend break results by including different specifications of a control function — which allows a separate polynomial in age for each year.

I again use the approach of Hansen (2000) to estimate the following model:

$$ln(mort_{apc}) = \delta^{p}_{1,c} \cdot c + \delta^{p}_{2,c} \cdot 1_{c \geq \gamma} \cdot (c - \gamma) + f(a, p) + \epsilon_{apc} \tag{5}$$

where $ln(mort_{apc})$ denotes the log mortaliy rate of the cell of age a, period p, and cohort c — for either white men or women. $\delta^{p}_{2,c}$ estimates the size of the break in each year p, and γ estimates the cohort at which a break occurs. I include increasingly flexible specifications of the "control function" $f(a, p)$. In most specifications the cohort trend $\delta_{1,c}$ is not separately identified from aspects of the control function, but main objects of interest, the size and location of the trend break are identified.

As above all models are estimated by concentrated, weighted least squares, following the approach outlined in Hansen (2000). The sample includes the years 1980-2015, ages 30-75, and cohorts born from 1930-1970.

Table 1 reports the results of estimating equation 4, with the log mortality rate of white women for single age-by-year bins as the dependent variable. Each column contains the results from a separate regression, with progressively more flexible specifications of the control function $f(a, p)$ from left to right.

Column 1 shows results from a model including a full-set of year fixed-effects, and a full set of age fixed-effects. This model is very flexible with respect to the shape of the age-profile in all years, but does not allow this shape to vary by year. All remaining columns include a full-set of year FEs, as well as progressively higher-order polynomials in age *interacted with year*. At the most extreme, column 4 allows a separate cubic in age for each year, so allows the impact of age on mortality to vary by year, albeit in a smooth parametric way.

The location of the trend break is consistently estimated to occur between the 1950 and 1951 cohort across all specifications. Additionally, the 99 percent confidence intervals, calculating by inverting the likelihood ratio statistic of Hansen (2000) are very tight, each including at most two cohorts. Graphs of these likelihood ratio tests for each specification are shown in Appendix Figure 1. The horizontal line denotes the 1 percent critical value of 10.59.

The average size of the estimated cohort break — the average value of $\delta_{2,c}$ across all years—changes depending on the specification of the control function. It ranges from .043,

Table 1: Log mortality of white women, shared cohort-specific trend break

	(1)	(2)	(3)	(4)
Average size of break	0.043	0.027	0.035	0.026
	(0.003)	(0.002)	(0.003)	(0.003)
Location of break	1950	1949	1950	1951
	[1950, 1950]	[1949, 1950]	[1950, 1950]	[1951, 1951]
P-value for existence of break	$< .001$	$< .001$	$< .001$	$< .001$
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	No	No	No
Linear-age-by-year	No	Yes	Yes	Yes
Quadratic-age-by-year	No	No	Yes	Yes
Cubic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on equation 4, with the log mortality rate of white men for single age-by-year bins as the dependent variable. All models are estimated by concentrated, weighted least squares, following the approach outlined in Hansen (2000). The sample includes the years 1980-2015, ages 30-75, and cohorts born from 1930-1970. The row titled "Average size of break" reports the average value of $\delta_{2,c}$ across all years, with the standard error in parentheses calculated by the delta method. The row titled "Location of break" reports the estimated cohort at which a trend break occurs, with a 99 % confidence interval in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence of break" reports p-value from a F-type test for the null hypothesis that no trend break occurs, based on 1000 bootstrap samples.

with a standard error of .003, in the model with age and year fixed-effects, to .026, with the same standard error, when I allow for cubic-age-by-year controls.

Even the smallest estimate of .026, suggests a substantial impact on mortality. It implies that the white women born in 1961 have had on average a 26 percent higher mortality rate then if their mortality experience matched the trend for pre-1950 cohorts.

For each model I follow the bootstrap procedure described in Hansen (2000) to test the null hypothesis that no trend break occurs, ie. that $\delta^p_{2,c}$ is equal to 0 for all p. For all models, the value of the F-type statistic for the true data is larger than all of the 1000 bootstrap repetitions — suggesting a p-value of less than .001 for the null of no break.

Table 2 reports analogous results with the log mortality rate of white men as the dependent variable. For white men, the estimated location of the trend break is even more consistent and precisely estimated. It is estimated to occur at the 1947 cohort across all specifications, and the 99 percent confidence intervals do not overlap any other cohorts. Graphs of these likelihood ratio tests for each specification are shown in Appendix Figure 2. In all the figures only the 1947 cohort is below this critical value.

As for women, the average size of the estimated cohort break for white men varies depending on the specification of the control function. The range of estimates are very similar to that for women: ranging from .047, with a standard error of .001, in the model with age and year fixed-effects, to .025, with the same standard error, when I allow for cubic-age-by-year controls.

Again the even the smallest estimate of .025, suggests a non-trivial impact of the trend break on mortality rates. It implies that the white men born in 1957 have had on average a 25 percent higher mortality rate then if their mortality experience matched the trend for pre-1947 cohorts.

For men as well, in all models the value of the F-type statistic for the true data is larger than all of the 1000 bootstrap repetitions — implying a p-value of less than .001 for the null of no break.

6 The cohort-specific trend break in log-mortality and year-over-year trends in mortality at different ages

I next use the previously estimated model to assess whether the cohort-specific pattern can explain the timing of recent increases, and stagnating improvements, in mortality by age. As described above, the timing by year of mortality trend breaks has not been uniform across

Table 2: Log mortality of white men, shared cohort-specific trend break

	(1)	(2)	(3)	(4)
Average size of break	0.047	0.037	0.034	0.025
	(0.001)	(0.001)	(0.001)	(0.002)
Location of break	1947	1947	1947	1947
	[1947, 1947]	[1947, 1947]	[1947, 1947]	[1947, 1947]
P-value for existence of break	$< .001$	$< .001$	$< .001$	$< .001$
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	No	No	No
Linear-age-by-year	No	Yes	Yes	Yes
Quadratic-age-by-year	No	No	Yes	Yes
Cubic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on equation 4, with the log mortality rate of white men for single age-by-year bins as the dependent variable. All models are estimated by concentrated, weighted least squares, following the approach outlined in Hansen (2000). The sample includes the years 1980-2015, ages 30-75, and cohorts born from 1930-1970. The row titled "Average size of break" reports the average value of $\delta_{2,c}$ across all years, with the standard error in parentheses calculated by the delta method. The row titled "Location of break" reports the estimated cohort at which a trend break occurs, with a 99 % confidence interval in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence of break" reports p-value from a F-type test for the null hypothesis that no trend break occurs, based on 1000 bootstrap samples.

age: while the age-adjusted mortality rate of non-Hispanics white women aged 45-54 began to increase in 1999 (Case and Deaton, 2015; Gelman and Auerbach, 2016), that of those aged 35-44 began to increase earlier in 1991, while that of 55-64 year olds continued to decline until a sudden break in 2010. Below, I use the shared cohort-specific trend break model estimated in the previous section to assess whether this cohort pattern can explain the staggered timing of mortality increases by age-bin. Intuitively, the question is whether "unhealthy" post-1951 cohorts have driven mortality increases first at young ages and subsequently at older ages, as they move through the age distribution.

To do so, I use the estimation results from the shared trend break model based on equation 5 and described in the previous section. I use the specification including a full set of year fixed-effects and a separate linear age effect for each year, reported in column 2 of Tables 1 and 2. I then use the estimated model to simulate mortality rates by year and age group. Specifically, for each age-year cell I predict log mortality, and then calculate predicted mortality as the natural exponential of predicted log mortality, for each single age-by-year pair. Finally, I calculate the simulated age-adjusted mortality for age-bins as the simple average across single ages.

Figure 5 shows the true age-adjusted mortality rates and those simulated from the above model, for white women by year, for 35-44, 45-54, 55-64, year olds respectively. The simulated series closely tracks the true mortality rates for all age bins. Notably, the simulated mortality rates match realized trend breaks in the mortality rate of 45-54 year olds in 1998 and of 55-64 year olds in 2010. The simulated series also nearly matches the increase in mortality of 35-44 year olds, predicting an increase starting after 1992 rather than the observed break in 1991.

Figure 8: True mortality rates and those predicted by a model with cohort-specific trend break, white women

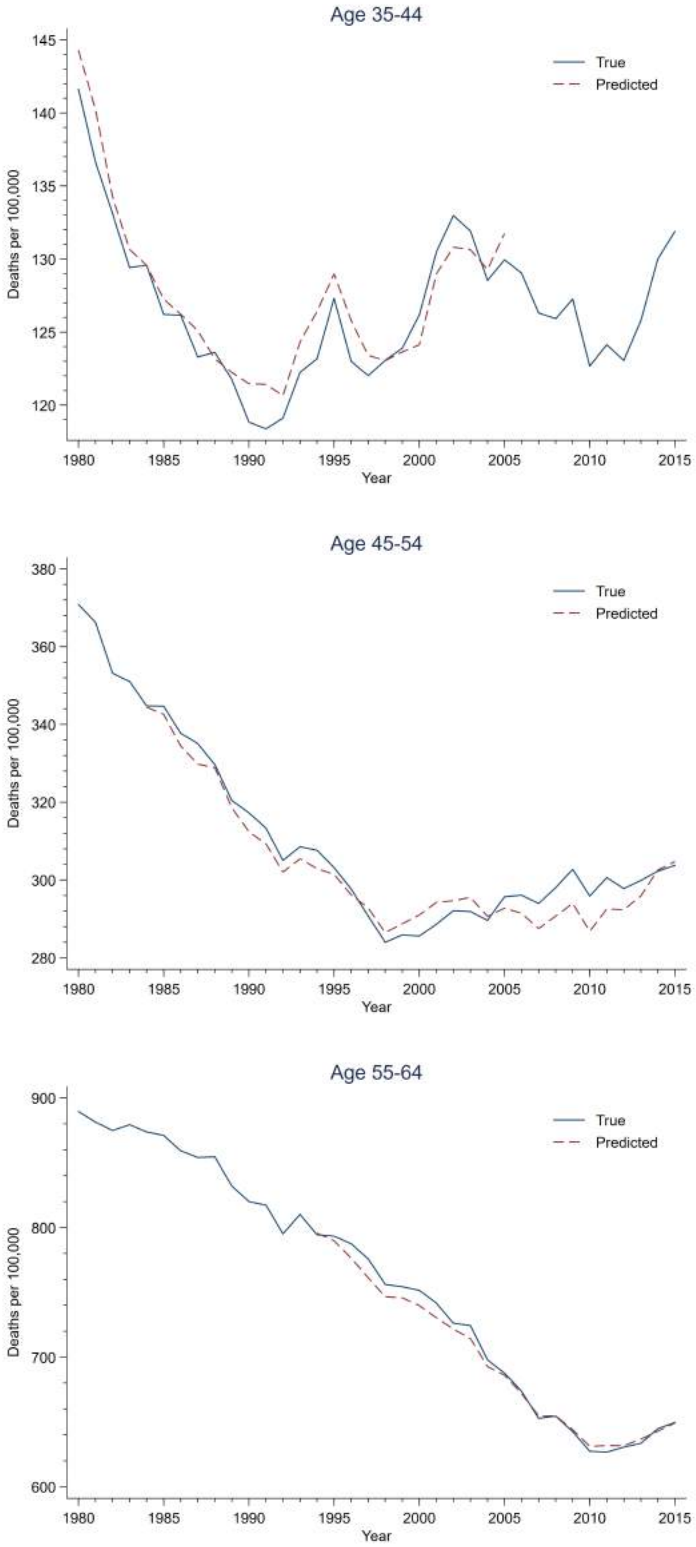
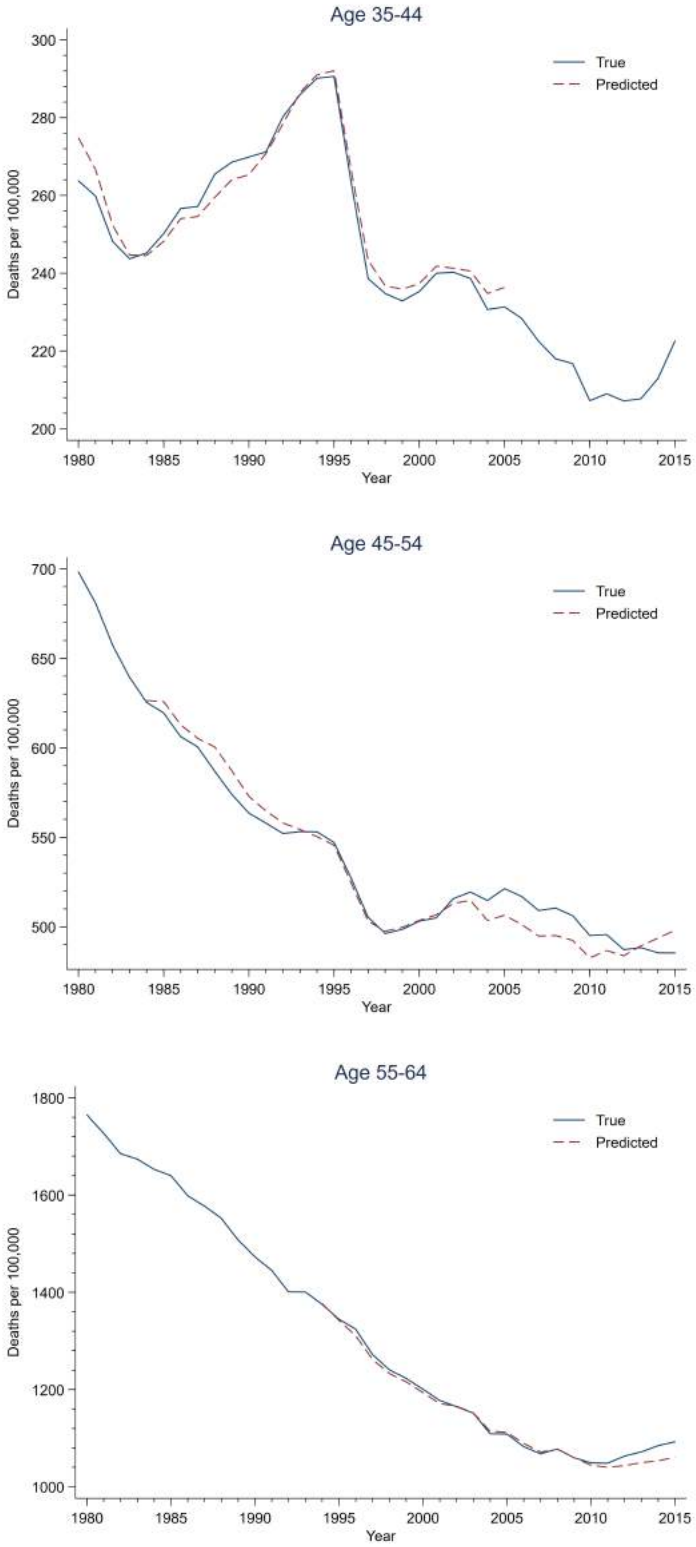


Figure 6 shows analogous results for white men. The mortality experience of white men aged 35-44 shows the clear imprint of the AIDS epidemic — increasing between the early 1980s and 1995 and then declining sharply. Never the less, the simulated series from the simple cohort-specific trend break model again closely tracks the true mortality rates for all age bins. Again for men, the simulated mortality rates matches the observed trend break in the mortality rate of 45-54 year olds in 1998. It also matches the timing of the break for 55-64 year olds in 2010 — though the size of the true increase is larger than the simulated one.

Figure 9: True mortality rates and those predicted by a model with cohort-specific trend break, white men



There is a relatively tight mapping between the mortality rates predicted from the cohort-specific trend break model — without any other non-linear age-by-year interactions — and the true year-over-year patterns in mortality by age. This suggests that cohort specific differences in health plausibly played a role in recent increases in mortality by age.

Conclusion

That increased mortality of white Americans appears cohort-specific suggests it's causes are as well: differing early life circumstances, or long-standing differences in health behaviors or labor market outcomes; rather than recent changes.

The apparently unhealthy cohorts are about to enter old age and their depressed health could increase health care spending, depress labor force participation, and impact the solvency of programs such as Medicare.

In ongoing research I am examining whether a similar cohort-specific pattern is evident in education, employment, and earnings. Preliminary results suggest that such a pattern does exists — cohorts after the middle of the century appear to suddenly have worse outcomes relative to the preceding trend.

References

Anne Case and Angus Deaton. Rising morbidity and mortality in midlife among white non-hispanic americans in the 21st century. *Proceedings of the National Academy of Sciences*, 112(49):15078–15083, 2015.

Andrew Gelman and Jonathan Auerbach. Age-aggregation bias in mortality trends. *Proceedings of the National Academy of Sciences*, 113(7):E816–E817, 2016.

Benjamin Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. *Philosophical transactions of the Royal Society of London*, 115:513–583, 1825.

Bruce E Hansen. Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics*, 93(2):345–368, 1999.

Bruce E Hansen. Sample splitting and threshold estimation. *Econometrica*, 68(3):575–603, 2000.

Ryan K Masters, Robert A Hummer, and Daniel A Powers. Educational differences in us adult mortality: A cohort perspective. *American Sociological Review*, 77(4):548–572, 2012.

Ryan K Masters, Robert A Hummer, Daniel A Powers, Audrey Beck, Shih-Fan Lin, and Brian Karl Finch. Long-term trends in adult mortality for U.S. blacks and whites: An examination of period-and cohort-based changes. *Demography*, 51(6):2047–2073, 2014.

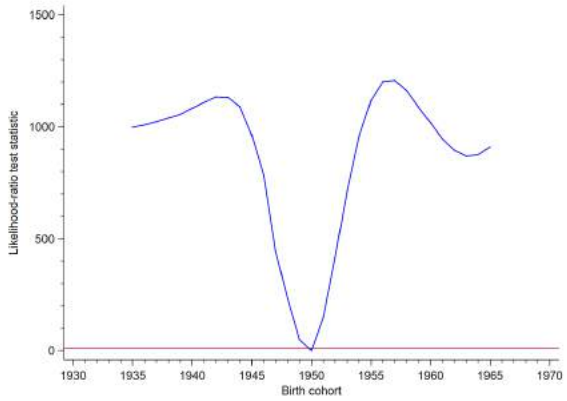
Ellen R Meara, Seth Richards, and David M Cutler. The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs*, 27(2):350–360, 2008.

S Jay Olshansky, Toni Antonucci, Lisa Berkman, Robert H Binstock, Axel Boersch-Supan, John T Cacioppo, Bruce A Carnes, Laura L Carstensen, Linda P Fried, Dana P Goldman, et al. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs*, 31(8):1803–1813, 2012.

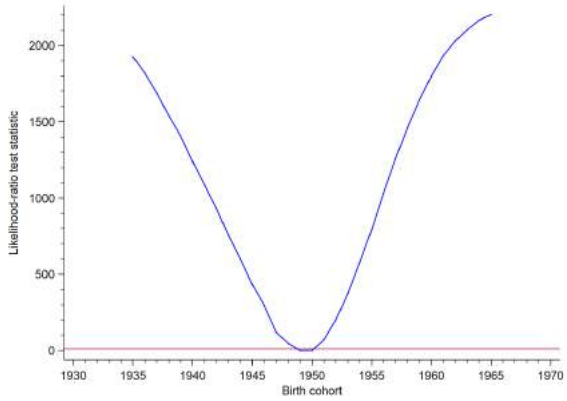
Yang Yang. Trends in U.S. adult chronic disease mortality, 1960–1999: Age, period, and cohort variations. *Demography*, 45(2):387–416, 2008.

Appendix Figure 1: Likelihood ratio test of location of shared cohort-break, by specification; white women

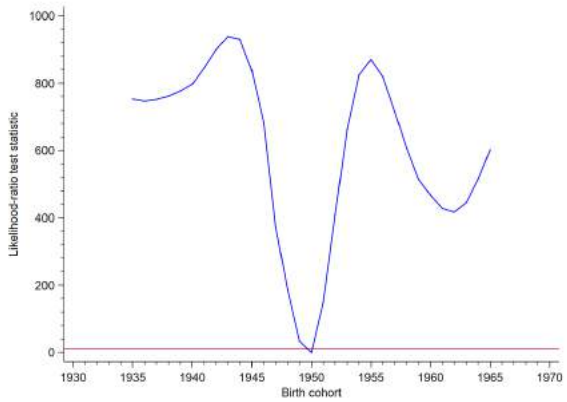
(1) Age FE, Year FE



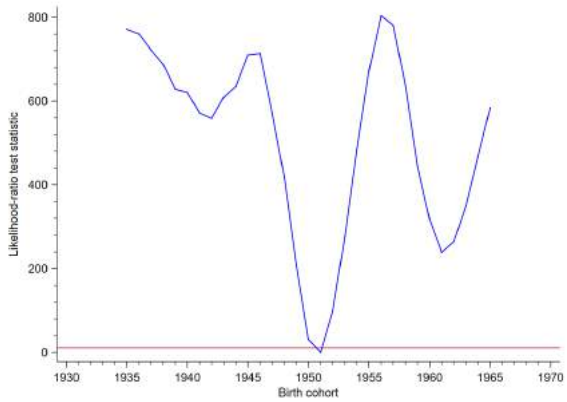
(2) Linear-age-by-year



(3) Quadratic-age-by-year

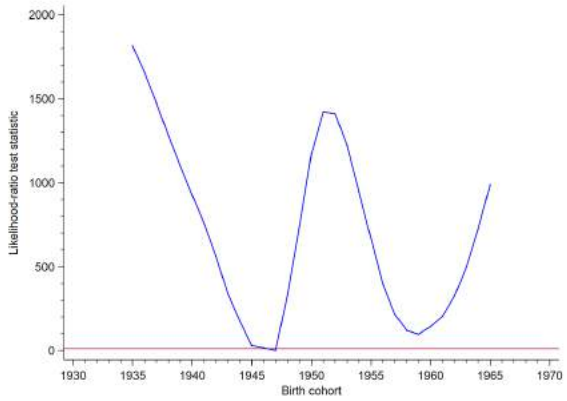


(4) Cubic-age-by-year

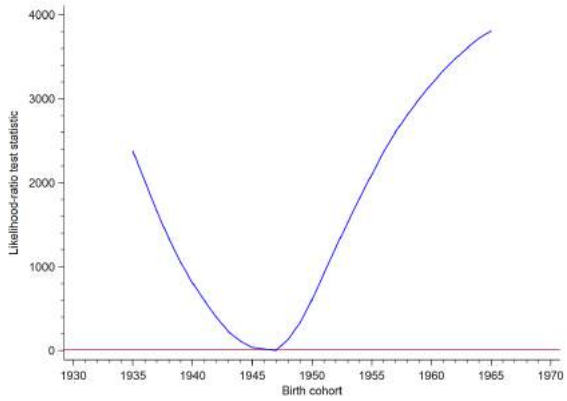


Appendix Figure 2: Likelihood ratio test of location of shared cohort-break, by specification; white men

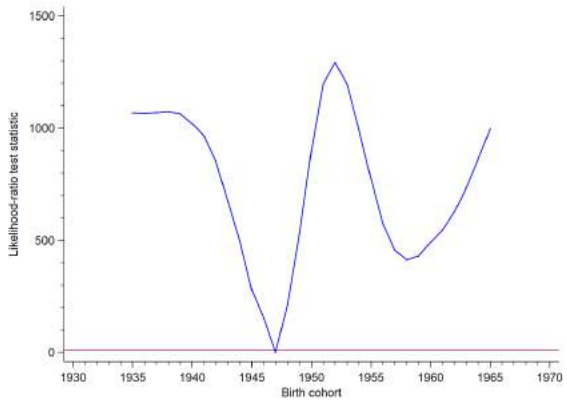
(1) Age FE, Year FE



(2) Linear-age-by-year



(3) Quadratic-age-by-year



(4) Cubic-age-by-year