The Broad Decline in Health and Human Capital of Americans Born after 1947

Nicholas Reynolds
University of Essex *

September 19, 2023

Abstract

I present evidence of a decline across cohorts in the health and human capital of Americans, beginning with those born after 1947 and continuing until those born in the mid-1960s. Age-adjusted educational attainment, wages, maternal health (proxied by the birth weight of infants), and mortality all exhibit trend breaks near the 1947 cohort, such that each outcome worsens for subsequent cohorts relative to the prior trend. Evidence of these breaks remain when I control for year and age effects or smooth age-by-year interactions, or use a nonparametric test. These cohorts also went through puberty later and scored lower on standardized tests, suggesting the roots of the decline date to at least adolescence.

The cohort decline is large enough to drive aggregate: i) educational declines in the 1960s, ii) increases in the low birth weight rate in the 1980s, iii) mortality increases since 1999, and to contribute substantially to iv) wage stagnation since the 1970s. My theory therefore partially unifies the disparate searches for the causes of each of these declines. The decline is remarkably widespread across geography and race, for those born in the United States.

*Nicholas.Reynolds@essex.ac.uk. I thank Kenneth Chay, Emily Oster and Anna Aizer for comments and encouragement throughout the project. I also benefited from comments from Andrew Foster, Ina Ganguli, Stefan Hut, Jeffrey Lin, Jesse Shapiro, Alison Swartz, Juan Pablo Uribe, Atheendar Venkataramani, and participants at seminars at Brown University, Haverford College, Lehigh University, the University of Essex, the Max Planck Institute for Demographic Research, and the NBER Summer Institute. Thanks to Ryan Kessler for sharing code and guidance on the Empirical Bayes estimation. I appreciate fellowship support from two National Institutes of Health training grants (T32 HD007338) and the Fellowship on the Economics of an Aging Labor Force from the NBER. I completed much of the research at the Population Studies and Training Center at Brown University, which receives funding from the National Institutes of Health (P2C HD041020) for general support.

1. Introduction

Health and human capital improved enormously across American cohorts born in the first half of the twentieth century. Robert Fogel and others have argued that the key driver of these improvements have been gains in health in early life.[1] In this paper, I present evidence that, in a number of dimensions, this progress in adult health and human capital hit a sudden stop and reversed trend for cohorts born after 1947. Instead, this paper suggests that the underlying health and human capital of Americans born in each year between 1947 and the mid-1960s has been worse than that of those born a year earlier. I present evidence that this decline likely originated in adolescence or earlier.

My cohort-based theory provides a partially unified explanation of puzzling societal declines occurring at distinct points over the last 60 years. As post-1947 cohorts have aged their depressed health and human capital have manifested as lower test scores, educational attainment, wages, and maternal health, and increased mortality rates — contributing to well-known aggregate declines in these outcomes. In the 1960s and 1970s, these cohorts drove the sharp declines in achievement tests, high school graduation rates, and college enrollment rates (National Commission on Excellence in Education, 1983; Card and Lemieux, 2001b). The depressed labor market ability of men in these cohorts then led them to have lower wages, contributing to wage stagnation since the 1970s , particularly for those without a college degree (Gould, 2014; Katz and Murphy, 1992; Autor et al., 2008). Depressed health as mothers led the women in these cohorts to give birth to less healthy infants, driving the upturn in the low birth weight rate in the 1980s (Centers for Disease Control, 1994; Currie and Gruber, 1996). Ultimately, the poor health of these cohorts has increased their likelihood of untimely death, contributing to recent mortality increases at midlife (Case and Deaton, 2015, 2017; National Academies of Sciences et al., 2021).

Figure 1 shows that the trend across cohorts in age-adjusted educational attainment, wages, maternal health (proxied by the birth weight of infants), and mortality all exhibit trend breaks near the 1947 cohort, such that each outcome declines for subsequent cohorts relative to the prior trend.[2] These simultaneous trend breaks, while striking, could in principle reflect differences in external factors which these cohorts were exposed to, rather than underlying differences in health and human capital. That is, cohorts born after 1947 may have been otherwise similar to earlier cohorts, but were merely unlucky to have experienced bad conditions throughout their lifetime.

I therefore provide evidence, under increasingly weak assumptions about the nature of exter-

[1] See for example Fogel (1986); Fogel and Costa (1997); Costa and Steckel (1997); Floud et al. (2011); Fogel (2012); Costa (2015).

[2] I examine maternal health by studying the birth weight of infants by the birth cohort *of their mother* using vital statistics microdata on more than 75 million births; wages and education , using CPS survey data; and the mortality rates of men and women, again using vital statistics data. Appendix Figure A1 shows similar patterns for more detailed educational categories.

nal age-by-year factors, that these patterns reflect a decline in the underlying health and human capital of cohorts born after 1947, relative to the prior trend. First, I estimate traditional age-period-cohort models, which assume cohort, age, and year factors are each additively separable. Second, I estimate models with a trend break of unknown location in cohort effects while allowing for a separate polynomial in age in each year, adapting methods from the structural break literature (Hansen, 1999, 2000). Third, I present a method for identification and estimation of a relative cohort decline in a general nonseparable framework in which underlying cohort factors and external age-by-year factors are both multidimensional and their interaction is unrestricted. I define a generalization of cohort effects as the counterfactual of what a cohort's outcomes would have been had they been exposed to a different set of external age-by-year factors than those they experienced. I show how to identify a second difference in "Local Average Cohort Effects" — the counterfactual average outcomes a cohort would have if they had experienced the external age-by-year factors of a cohort born one year earlier — under the assumption that age-by-year shocks are as-if-randomly-assigned between neighboring ages. This assumption allows for a permutation test of the null of no difference in underlying cohort factors.

All of these methods reveal strong evidence of a trend break at the 1947 or 1948 cohort in underlying health and human capital evident in each of the above outcomes. The remaining threat to validity would be nonsmooth changes in the impact of age across years, such as a large shock that impacts the health of individuals aged 30-and-under, but not of 31-year-olds, which repeatedly "hit" the same cohorts. Alternative explanations put forward in the related literature do not have this feature. For example, the effects of shifts in supply or demand will be smooth as long as individuals who are close in age are substitutable (Card and Lemieux, 2001a).

Simulations suggest that the cohort-based decline contributed substantially to the year-over-year declines described above. The structural-break-based estimates imply that absent this cross-cohort health and human capital decline: the rise in low birth weight births in the 1980s would not have occurred; there would have been very modest real wage growth between 1979 and 1993, rather than real declines; and midlife mortality of men and women would have continued to fall steadily at its pre-1999 rate. This suggests that previously independent searches for the causes of educational declines of the 1970s, increases in the low birth weight rate in the 1980s, increases in the midlife mortality of whites since 1999, and to a lesser extent wage stagnation over the last 50 years, can be at least partially unified. To find the cause of all of these declines will require understanding what went deeply wrong for cohorts of Americans born after 1947.

The above methods do not formally identify the long-run trend in cohort effects, and therefore cannot speak to whether cohort health and human capital declined in absolute terms or the trend of improvement just suddenly slowed. The absolute declines in educational attainment seem to point towards absolute declines in human capital. I show that under reasonable assumptions on

the long-run trend in period or age effects, cohort effects in the other outcomes also declined in absolute terms after 1947.[3]

I then present evidence that the health and human capital decline likely originated in adolescence or earlier. Most strikingly, males born in the 1960s appear to have had a later adolescent growth spurt than those born a decade earlier, and females born in the 1960s experienced menarche later. This represents a reversal of a long-run trend towards earlier puberty, which was thought to be driven by improvements in nutrition and health (Hauspie et al., 1997). Puberty is also generally later in low income countries and for people of lower SES (eg. Lleras-Muney et al., 2022). The long run "secular trend" in growth in childhood and adult height also suddenly stopped some time around the 1945 to 1955 cohorts, while continuing in other countries.[4] Cohorts born after 1947 also did worse on standardized tests as adolescents. The health decline is only evident for those born in the US, not for residents born elsewhere.

I provide a preliminary investigation of the underlying cause of the apparent early life health decline of these cohorts. First, I highlight that the health decline is widespread across demographic groups in the US; with the exception that the decline appears to be largest among residents of counties in the highest quartile of population density. I then provide evidence against a host of ex ante plausible causes. One promising candidate cause is lead pollution from motor vehicle exhaust. Motor vehicle use increased rapidly after 1945 and the gasoline used in motor vehicles included increasing quantities of lead additives. Fetal and early life lead exposure has been linked to poor health and cognitive development. High child blood lead content was ubiquitous across the US when these cohorts were children and highest in the densest urban areas.[5]

The evidence above suggests that something went deeply wrong early in life for cohorts of Americans born after 1947, and that as these cohorts aged their poor outcomes have contributed substantially to aggregate declines over the last 60 years. This does not rule out some role for period-based factors. What this paper points to strongly is that these recent crises have deep roots and that the first link in the chain of "cumulative deprivation," in the words of Case and Deaton, came early in life for these cohorts. It seems likely that the initial poor start for these cohorts was exacerbated by the poor demand for "unskilled" workers when they entered the labor market, and again by massively expanded supply of prescription and illegal opioids as they hit midlife — and future research should examine these interactions.

The remainder of the paper is organized as follows. Section 2 reviews relevant literature.

[3] I use approaches similar to those in Lagakos et al. (2018) and Fosse and Winship (2019).

[4] The stagnation in adult heights was noted by Komlos and Lauderdale (2007a,b). I extend their findings to include historical data on childhood height and additional years of data on adult heights.

[5] See the discussion in Reynolds (2023) and below. McFarland et al. (2022) estimate that the share of children with blood lead content above the 2015 threshold for "clinical concern" increased from 50 percent for the 1940-45 cohorts to 100 percent for the 1966-75 cohorts.

Section 3 describes the data and main outcomes. Section 4 presents a conceptual model to guide the analysis. Section 5 describes the method and results from the additively-separable age-period-cohort model. Section 6 describes the method and results from the models which allow for smooth age-by-year interactions. Section 7 presents the non-parametric method and results. Section 8 shows the sizeable impact of the cross-cohort declines on year-over-year changes in each outcome. Section 9 presents evidence that the decline likely originated in childhood. Section 10 present a preliminary investigation into the root cause of the cohort decline. Section 11 concludes.

2. Literature review

Some earlier research has discussed the possibility of cohort-based declines contributing to the aggregate patterns described above. Some papers also have presented evidence of cohort-specific declines in particular outcomes for Americans born after roughly the middle of the twentieth century.

My paper is unique in providing: a unified analysis which documents the shared timing of the cohort decline in multiple outcomes; a careful analysis showing that the break in trend is sharp, documenting its precise timing, and attempting to rule out external factors; evidence that the decline likely originated early in life; and an initial investigation into the cause of that early-life health decline.

Multiple outcomes

Few papers have noted that multiple outcomes have declined for post-1947 cohorts. An exception is Borella et al. (2020), who compare the outcomes for white Americans without a college degrees born in the 1940s to those born in the 1960s across a number of outcomes. They document declines in wages and projected life expectancy and increases in medical expenses for this group. There are a few important distinctions between their study and mine. First, I study average outcomes for all Americans in each cohort born between 1930 and 1970, while they focus on a selected subset. Second, their focus is on the effects of the declines in the above outcomes on labor supply, savings, and welfare, which they quantify using a life-cycle model. In contrast, my goal is to understand whether the declines in these (and other) outcomes are driven by "cohort effects" — long-standing differences in health and human capital — or external factors which have hit these cohorts at different points in their lifetime. I also have a particular focus on the sharp trend break in outcomes near the 1947 cohort.

Educational attainment

The declines in educational attainment for cohorts born after the late-1940s is well known but the cause is not. Heckman and LaFontaine (2010) estimate that the high school graduation rate peaked at around 80 percent in the late 1960s — approximately when the 1946 cohort was 18 — and has declined by 4-5 percentage points since. Card and Lemieux (2001b) show a sudden 12

percent fall in college entrance rates for men from 1968 to 1978 — approximately the 1947 cohort to the 1957 cohort — and a stagnation in prior improvements for women. They also show that the share of each cohort completing a Bachelor's degree suddenly began to decline around 1945 for men, and stagnated around 1950 for women. They present an extensive study of possible causes of the decline, concluding that for women it could be explained by low returns to education and cohort size, but that for men the decline represents a fundamental trend break with no observable explanation.[6] This decline in college completion rates is the key driver of increases in the college-high school wage gap in the "canonical model" of Katz and Murphy (1992) and subsequent studies (Card and Lemieux, 2001a; Autor et al., 2008). Acemoglu and Autor (2012) also note this stagnation in educational attainment, and presaging the arguments in this study, write that the sharpness of the change in trend by cohorts suggests it is unlikely to be caused by a sudden change in the school system, and that "other factors are thus likely to be at play." To my knowledge, my finding below of trend breaks at the 1947 cohort in the share of men and women earning an advanced degree is novel.

Wages and earnings

Initial research into wage stagnation beginning in the 1970s and the rapid decline in wages for the less educated in the 1980s considered but largely dismissed the role of a cross cohort decline in "earnings ability" (eg. Bound and Johnson, 1992; Katz et al., 1999). The results in this paper suggest it is perhaps time for a reconsideration. Interestingly, an earlier paper, Bishop (1989), estimated that the decline in cognitive ability, as measured by test scores, had a large effect on wage rates for these cohorts and would continue to contribute substantially to productivity decline over the coming decades. More recently, Guvenen et al. (2022) use panel data on earnings from tax returns to study the lifetime earnings of successive cohorts. They document a decline in men's median lifetime earnings of more than 10 percent between the 1942 and 1960 birth cohorts. They also note that these cohorts were already behind by the time they were 25, and therefore close the paper by stating that "the sources of the dramatic changes we have witnessed in the U.S. earnings distribution over the last 50 years may be found in the experiences of newer cohorts during their youth (and possibly earlier), and how those experiences differed from those of older cohorts." My paper builds on theirs by explicitly confronting the challenge of identifying whether these (and other) declines for these cohorts are driven by long-standing differences in health and human capital.

Maternal health

To my knowledge no prior papers have documented evidence of the cohort decline in maternal health shown in this paper. Both the idea of maternal endowments as contributing to birth outcomes

[6]Handy and Shester (2019) suggest that changes in the average parity or birth order by cohort can explain around one third of the decline in Bachelor's completion for white men born between 1946 and 1960.

of children (Rosenzweig and Schultz, 1983; Grossman and Joyce, 1990) and that of the life-course perspective on maternal health (Lu and Halfon, 2003) suggest the using birth outcomes of children can be useful to understand the underlying health of mothers. This strategy was previously used in Almond and Chay (2006) to study black-white differences in cohort health.[7]

Mortality

Case and Deaton (2015) documented a shocking increase in the all-cause mortality rate of non-Hispanic white Americans and noted a proximate cause was increases in so-called "deaths of despair" — drug and alcohol poisonings, suicide, and chronic liver diseases and cirrhosis. Subsequent research has examined mortality across racial groups and suggested that cohort differences in health and disadvantage which predated the 1990s are likely important as well (Case and Deaton, 2017; Lleras-Muney, 2017; Masters et al., 2017; Zang et al., 2018; Acosta et al., 2020; Reynolds, 2023; National Academies of Sciences et al., 2021).

Case and Deaton (2017) develop a preliminary theory of "cumulative disadvantage" which emphasizes worsening opportunities at labor market entry — particularly for whites with low levels of education — as a key driver of the mid-life mortality increases. They posit that these worse opportunities at labor market entry trigger various negative outcomes which build on each other, and culminate in an increased likelihood of untimely death. The results in this paper and my earlier work (Reynolds, 2023) suggest that these cohorts had already fallen behind before labor market entry. It seems likely that worsening economic conditions interacted with and potentially exacerbated the poor health and human capital of post-1947 cohorts.

Height

Adult height, often viewed as a proxy of childhood nutritional status (Tanner, 1990; Floud et al., 2011), also suddenly stopped increasing for white Americans born between approximately 1955 and 1974, while the height of Europeans continued to grow rapidly (Komlos and Lauderdale, 2007a,b; Komlos, 2010). I build on these findings by examining height and physical development of these cohorts as children. I also show evidence that the males in cohorts born in the 1960s grew more slowly in adolescence than those born around a decade earlier.

Other outcomes

Since I began work on this project a series of papers by Hui Zheng and coauthors have shown declines in other outcomes for approximately the same cohorts. Zheng (2021) finds declines in cognitive functioning in old age. Zheng and Echave (2021) find that a summary measure of biomarkers of inflammation, metabolic functions, lung function, and renal function appears to have worsened

[7] Also potentially relevant is Tilstra and Masters (2020), which finds that recent declines in birth weight were due to increases in induced labor and cesarean delivery at earlier gestational ages. These findings could be complementary to mine: either these cohorts are more likely to have ceasareans or both explanations are important and could more than explain recent declines in birth weight.

continuously from the Baby Boom cohort onwards.[8] Zheng et al. (2022) also find evidence of declines in similar measures and a chronic disease index, and attribute these declines to exposure to income inequality in childhood. Interpretation of the findings from these papers is complicated by the fact that all results come from a "mixed-model" in which age effects are modelled as fixed effects and cohort and period effects as random effects.[9] The analysis in these papers also does not identify the precise timing that decline begins.

Early-life health and human capital development

This paper presents evidence suggesting that the cohort decline likely originated in childhood or adolescence. This is plausible in light of the literature showing large effects of infant and childhood health on educational attainment, health, and labor market outcomes (see Currie and Almond (2011) for a review). Most economics papers is this area are focused on careful causal identification rather than aggregate implications of changes in childhood health. A notable exception is the earlier cited work by Robert Fogel and other economic historians, which focus primarily on earlier cohorts.[10]

3. Data and main outcomes

To document the cohort declines in health and human capital I use microdata from the Current Population Survey, and vital statistics data recording nearly the universe of births and deaths.

A. Maternal health

I use detailed microdata on 50 to 100 percent of the births in the US in each year between 1968 and 1995 to study patterns in infant health by mother's birth cohort. The data are known as the "natality microdata files" and are produced by the National Center for Health Statistics. They provide detailed information derived from birth certificates. They include a 50 percent sample of all births in 1968, and include progressively larger samples until 1985, after which they include the universe of births.

I consider the health of infants as a proxy for maternal health and also to provide evidence of an intergenerational effect of the apparent decline in cohort health. Under the usual assumption

[8] However, they do not find similar patterns in a summary measure of clinical risk factors for cardiovascular diseases, stroke, kidney disease, and type 2 diabetes.

[9] See Bell and Jones (2013); Luo and Hodges (2020) for critiques of these models. More broadly it would be useful to begin with age-adjusted patterns by cohort before turning to identification.

[10] Interestingly, they find evidence of a decline in height and other proxies of childhood and cohort health for cohorts born between the 1830 and 1890 cohorts (Steckel and Haurin, 1982; Komlos, 1987). This decline was then followed by rapid growth for the subsequent six decades (Costa and Steckel, 1997). Also related to my study are Almond and Chay (2006) and Chay et al. (2009, 2014) who find long-run effects of infant health gains for African Americans born in the South after the Civil Rights Act of 1966 led to gains in test scores, educational attainment, earnings, and maternal health (with an intergenerational effect on infants). They also suggest that these cohort improvements were large enough to drive changes in the aggregate black-white gaps in these outcomes, which had previously been attributed to period-based factors.

that a mother's health "endowment" enters the infant health production function (Rosenzweig and Schultz, 1983; Grossman and Joyce, 1990), a decline in the health of infants by *mother's birth cohort* may be viewed as evidence of a decline in maternal health for these cohorts. [11]

Mothers' year of birth is not recorded explicitly. I therefore calculate the approximate birth year of each mother as the infant birth year minus the mother's age. I restrict my analysis to births to mothers born between 1930 and 1970. I use the years 1968 to 1995 and births to mothers ages 18 to 40. This results in an sample of more than 75 million births.

As my main infant health outcome, I study birth weight. I study birth weight as a continuous measure, and also use the commonly defined measure of low birth weight percentage, the percentage of infants weighing less than 2500 grams. For all analysis I calculate these measures in cells by year, single age, and birth year using sampling weights; and then conduct regressions and other analysis on these cell means. In analysis of heterogeneity I calculate means in cells defined based on age-year-cohort crossed with additional variables such as race, education, or Census Region.

B. Wages and educational attainment

To document cross-cohort declines in wages I draw on a large, commonly-used data source: the Current Population Survey, Merged Outgoing Rotation Group (CPS-MORG). The CPS-MORG has the advantage of recording point-in-time hourly earnings, and its yearly frequency aids identification of cohort effects.

I use the CPS MORG data from 1979 to 1993. These files contain information on the hourly or usual weekly earnings of a third of the individuals in each monthly CPS. Combined with information on usual hours worked this allows for the construction of point-in-time hourly wages. These hourly wage measures have been used extensively in studies of wage inequality and trends, and have some notable advantages over other sources.[12] Pooling across all months in a year yields a sample three times the size of a monthly CPS — such as the March CPS. Further, Lemieux (2006) and Autor et al. (2005) suggest that the point-in-time nature of the earnings question reduces measurement error.[13]

I focus on the wages of prime-age men age 25 to 54. This avoids the need to address large confounding changes in women's labor market participation over this period (Goldin, 2006). Exact year of birth is again not recorded explicitly. I therefore calculate the approximate birth year as the survey year minus the respondent's age. I restrict my analysis to cohorts born between 1930 and 1965. I end the analysis in 1993 to avoid having to address issues related to the substantial

[11] A similar use of this data to study cohort health can be found in Almond and Chay (2006), who focus on the black-white gap in health by cohort.

[12] See for example DiNardo et al. (1996); Lemieux (2006); Autor et al. (2008); Acemoglu and Autor (2011); and Gould (2014).

[13] I use the CEPR Uniform Extracts (for Economic and Research", 2020) and use their earnings definition, which is designed to follow NBER's recommendation and similar to past research.

redesign of the CPS earnings questions in 1994.[14] These restrictions lead to a sample of 970,479 men with non-missing earnings used in the main analysis.

I address the problem of topcoding of earnings, with a changing threshold, by focusing on the median and other quantiles. I calculate the sample median and other quantiles separately for age-year-sex cells, using the provided survey weights. These estimated quantiles will be unaffected by the topcoding — as long as the earnings value of the particular quantile is below the top-coding value. For regressions and other model-based analysis, I follow Chamberlain (1994) and Chetverikov et al. (2016) and use a two step procedure: first estimating cell quantiles, and second estimating models on the cell quantiles.[15] In analysis of heterogeneity I use a similar procedure, calculating quantiles by cells defined based on age-sex-year crossed with additional variables such as race, education, or Census Region. My main outcome variable is therefore median real hourly wages by single age and year.[16]

To study educational attainment I also use the CPS-MORG data, 1979-2016, and include men and women aged 25-75, who were born from 1930 to 1965. I calculate approximate average years of schooling for each cohort based on the 16 schooling categories in the CPS, and calculate the share of each cohort who have achieved different levels of educational attainment.

C. Mortality

My main mortality analysis uses data from the Human Mortality Database (HMD) on number of deaths and population counts by year and age. These data are derived from official vital statistics and census estimates, and are adjusted for errors using a uniform method. I then define cohort as year minus age. I restrict my analysis to the years 1975-2019, ages 25 to 85, and cohorts born between 1930 and 1965.

To analyze mortality by Census Region I use data from the United States Mortality Database (USMDB). The USMDB uses similar procedures to the HMD to create subnational mortality estimates.

To calculate mortality separately by race 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 the number of deaths and the mid-year population by single age,

[14] After the redesign individuals with allocated earnings could not be identified for the next 21 months. See Autor et al. (2005) for a further description.

[15] For sufficiently large cells this approach has advantages over traditional LAD quantile regression. For example it is unbiased in the presence of left hand side measurement error, unlike traditional quantile regression, see Hausman et al. (2019). Also note, there are actually two topcoding values in the CPS MORG: one for hourly earnings for hourly workers, and one for weekly earnings for non-hourly workers. The hourly earnings threshold is rarely crossed. I include all cells for which less than 1 percent of a cell is topcoded, which in practice is mainly determined by the weekly earnings threshold.

[16] Earnings are adjusted for inflation using the CPI-U-RS, and reported in 2014 dollars.

sex, race cells. I then calculate crude death rates — number of deaths over mid-year population — within each cell.

4. Conceptual model

This section presents a conceptual model which generalizes the idea of "cohort effects" and suggests a link to counterfactuals and potential outcomes. This model will guide the identification and estimation in the remainder of the paper.

Assume that each of the main outcomes I consider are determined by the following model:

$$Y_{apc} = g(\theta_c, \epsilon_{a,p}) \tag{1}$$

where Y_{apc} denotes an outcome — such as log wages or the log mortality rate — for individuals who are age a, in the year or "period" p, and who are members of the cohort c, ie. they were born in year c. The outcome is a function of: i) θ_c, underlying, fixed features of the cohort and ii) ϵ_{ap} external factors in the year p which affect individuals of the given age, a. There is no restriction (yet) on the structural function $g(\cdot, \cdot)$ or on the dimensionality of θ_c or ϵ_{ap}.

Conceptually, θ_c represents the underlying, fixed differences between individuals born in different years which impact their outcomes. It reflects broadly the health, human capital, cognitive ability and "skills" of the cohort. θ_c can be multidimensional to represent both the potentially multidimensional nature of health and "skill" (eg. Heckman, 2007), and to represent differences in the distribution of skills within cohorts. I do not specify explicitly when or how these cohort differences develop, just that they originate before the age at which I begin measuring the outcome. There is a long history in social science as well as neuroscience and human biology of studying differences between cohorts and of the suggestion that they are likely to differ due to different experiences at "critical periods" in their life.[17] For example, the cohort differences could date to labor market entry, schooling age, infancy, or in utero. The model treats cohort effects as fixed from the age at which the outcomes began to be measured, abstracting from investment in skills after that age and scarring effects in adulthood.

In contrast, ϵ_{ap} represents the external factors such as technology or labor demand which will impact the outcome of a cohort who is age a in year p. The interaction between θ_c and ϵ_{ap} is unrestricted, so the model allows for example for a situation where the labor market for 30 year old workers would be a good match for the skills of one cohort but a poor match for those of another cohort.

In the context of this model, the concept of "cohort effects" can be generalized and linked to

[17] See for example Ryder (1965); Easterlin (1987); Fogel and Costa (1997), and Cunha et al. (2006) and the citations therein.

the idea of counterfactuals and potential outcomes.[18] For example, a natural way to summarize a decline in the health of cohorts born between 1947 and 1960 would be to ask: had both cohorts faced the set of external age-by-year factors actually experienced by the 1947 cohort how would their mortality rates had differed? This is fundamentally a counterfactual question and involves the comparison of the observed outcomes for the 1947 cohort to a set of potential outcomes for the 1960 cohort which will never be observed. In particular, the outcomes Y_{apc} will only be observed for cohorts, ages, and periods such that $c = p - a$. We can define the structural function $g(\cdot, \cdot)$ however for all pairs of θ_c and ϵ_{ap} — defining "potential outcomes" which a cohort would have had if they had been exposed to different external factors. Identifying these generalized cohort effects therefore is akin to the "Fundamental Problem of Causal Inference" (Holland, 1986).

5. Evidence of cohort decline from age-period-cohort models

I first present evidence of cohort declines in health and human capital under the assumption that the impact of age, year, and cohort are additively separable, such that:

$$Y_{apc} = \gamma_c + \phi_p + \alpha_a + \epsilon_{apc} \tag{2}$$

The main object of interest are the cohort effects, represented by the sequence of fixed effects. γ_c. While ϕ_p and α_a are full sets of fixed effects in year and age, respectively. ϵ_{apc} is an orthogonal error.

This is the additively separable age-period-cohort model, common in economics and demography.[19] Conceptually, the cohort effects reflect the impact of underlying, fixed differences between individuals born in different years. To identify the cohort effects, the above model assumes that the impact of external factors can be decomposed as additively separable year and age components. It is therefore unrestrictive with respect to the dynamics of year-over-year changes impacting each outcome, and with respect to the *shape* of age effects. However, it does not allow age-by-year interactions of any kind. Year-specific factors are assumed to impact all ages equally. Visually, year effects can be thought of as shifting the entire "age profile" of outcomes evenly. I will relax this assumption in later sections.

The additive separability in this model simplifies the counterfactual questions described above. The difference between the cohort effect for the 1947 cohort, γ_{1947}, and that of the 1960 cohort, γ_{1960}, reveals how the outcomes of individuals born in these two years would have differed — holding all external factors fixed. Large differences in cohort effects therefore imply large differences

[18] See for example Rubin (1974); Holland (1986); Heckman (2010).

[19] For a textbook treatment see Deaton (1997). Recent applications in economics include Aguiar and Hurst (2013) and Lagakos et al. (2018), and in demography include Yang (2008), Masters et al. (2014), Masters et al. (2017), and Zang et al. (2018).

in the latent health or human capital of different cohorts, not just that the cohorts were unlucky to experience poor labor market or health conditions.

Even this model is not identified due to the exact collinearity of age, period, and cohort (see eg. Hall, 1968; Deaton, 1997). Identification can be achieved by imposing one additional linear restriction, such as restricting two ages or two periods to have the same effect. I proceed in two directions. First, I estimate "detrended cohort effects", which will reflect the true cohort effect, minus some unknown long-run trend in cohort effects.[20] Graphical inspection of the sequence of detrended cohort effects will allow for the identification of *trend breaks* in cohort effects. Second, I will consider what additional assumptions on either the age or period effects would imply about the long trend in cohort effects, and therefore about the entire sequence of cohort effects. Importantly, this will allow me to consider under what assumptions the data suggest there were *absolute* declines in health and human capital for cohorts born after 1947.

For both approaches a few additional definitions are helpful. Denote the first cohort included in the model as 1 and the last as C, the first year as 1 and the last as P, and the first age included in the model as 1 and the last as A. Define linear "trends" in cohort, year, and age effects as follows:

$$\beta_c \equiv \tfrac{\gamma_C - \gamma_1}{C}\ ; \qquad \beta_p \equiv \tfrac{\phi_P - \phi_1}{P}\ ; \qquad \beta_a \equiv \tfrac{\theta_A - \theta_1}{A}$$

For the year effects, this trend is the average annual change in effects between the first and last year in the sample. Analogously, for cohort effects it is the average per-year-of-birth change in effects between the first and last cohort in the sample. For example, in most of my analysis it will represent how much health or human capital had improved or declined on average per year-of-birth from those born in 1930 to those born in 1965. For age effects it reflects the per-single-age difference between the effects for the youngest and oldest ages in the sample.

Then define "detrended" cohort, age, and year effects as:

$$\tilde{\gamma}_c = \begin{cases} 0 & \text{if } c = 1 \\ \gamma_c - \beta_c \cdot c & \text{if } c \in (2, C] \end{cases} \qquad \tilde{\theta}_a = \begin{cases} 0 & \text{if } a = 1 \\ \theta_a - \beta_a \cdot a & \text{if } a \in (2, A] \end{cases} \qquad \tilde{\phi}_p = \begin{cases} 0 & \text{if } p = 1 \\ \phi_p - \beta_p \cdot p & \text{if } p \in (2, P] \end{cases}$$

And finally define two composite linear trends:

$$\tilde{\beta}_{pc} = \beta_p + \beta_c\ ; \qquad \tilde{\beta}_{ac} = \beta_a - \beta_c$$

One can then rewrite the model in Equation 2 as:

$$Y_{apc} = \tilde{\gamma}_a + \tilde{\phi}_p + \tilde{\delta}_c + \tilde{\beta}_{ac} \cdot a + \tilde{\beta}_{pc} \cdot p + \epsilon_{apc} \tag{3}$$

This reparameterized model is identified — the design matrix is full rank. One can identify detrended age, period, and cohort effects. Additionally, one can identify two composite linear

[20] This is similar to approaches in Deaton (1997) and Chauvel (2011).

trends which represent i) the sum of the period and cohort trends, and ii) the age trend minus the cohort trend. By making additional assumptions on age or period trends one can identify or bound the cohort trend, and therefore the full set of cohort effects.

A. Detrended cohort effects

Figure 2 shows the results of estimating the detrended, additively separable age-period-cohort models of Equation 3 for six outcomes: men's median log wage, the low birth weight percentage by mother's birth cohort, the log mortality of men and women, and the average years of schooling of men and women. The estimated cohort effects for each of the different outcomes exhibit remarkably similar patterns. They each have an approximately piecewise linear shape with a large trend break located at or near the 1947 cohort, and declines for cohorts born after that year. These patterns are consistent with a large decline in cohort health and human capital, relative to trend. This decline began near the 1947 cohort and was broad enough to impact outcomes as disparate and far reaching as maternal health, wages, and mortality.

Panel A shows results for the median hourly wage of employed men. Again, the estimated cohort effects exhibit a clear piecewise linear shape, with a large trend break precisely at the 1947 cohort. The cohort effects increase from a normalized 0 in 1930 to .14 by 1947, before suddenly changing slope and declining for subsequent cohorts. Under the additive separability assumptions of this model, Under the additively separable assumption of the model, this pattern suggests a large break in the cross-cohort trend in labor market ability at the 1947 cohort, such that each cohort born after this year has declining ability relative to the trend for prior cohorts. The magnitude of the trend break is large. The results imply that the 1965 cohort had a median wage nearly 29 log points lower, ie. 33 percent lower, than they would have had the trend in labor market ability for the 1930 to 1947 cohorts continued.

Panel B shows analogous results for the low birth weight rate of infants by their mother's birth cohort. Recall that age and cohort in these models refer to the age and year of birth of the *mother*. Therefore the estimated cohort effects can be viewed as estimates of the maternal health of given cohorts using the health of their infants as a proxy, and also reflect an intergenerational effect of cohort health on infant health in the next generation. The estimated cohort effects again exhibit a piecewise linear shape: declining rapidly until the 1947 cohort, sharply changing slope after that cohort, and increasing nearly linearly until the 1965 cohort. The cohort effects decline from the normalized 0 in 1935 to a minimum of -.87 for the 1947 cohort, before reversing trend. The results imply a large break in the cross-cohort maternal health trend at the 1947 cohort, such that each cohort born after this year has declining health relative to the trend for prior cohorts. The size of the trend break suggests than the 1965 cohort would have had a low birth weight rate approximately 2.2 percentage points lower had the cohort health decline not occurred.

Panels C and D show similar results for the log mortality rate of men and women, using data

from the Human Mortality Database (HMD). The shape of the estimated cohort effects are not as sharply piecewise linear as those for the labor market and maternal health outcomes. However, they exhibit clear changes in trend near the late 1940s cohorts, consistent with elevated mortality and declining health for subsequent cohorts. For men's log mortality the cohort effects decline — not precisely linearly – from 0 to below -.1 by the 1947 cohort, then suddenly reverse trend and increase rapidly until reaching above .05 by the late 1950s cohorts. They then flatten and decline slightly for subsequent cohorts. This pattern suggests that men born in 1960 had mortality near .25 log points higher than they would have had health improvements continued at the same rate as for the 1930 to 1947 cohorts. For women's log mortality the cohort effects exhibit two smaller trend breaks at the 1947 and 1951 cohorts, but still show evidence of a decline in health after the late 1940s relative to the prior trend. The cohort effects decline from 0 to below -.11 by the 1947 cohort, they then change trend and are nearly flat until the 1951 cohort. They then change trend *again* after the 1951 cohort and increase nearly linearly to around 0 by the 1960 cohort. Overall, the size of the two trend breaks imply that the 1965 cohort has had nearly .25 log points higher than it would had the health improvements for the 1930 to 1947 cohorts continued at the same rate for later cohorts.

Panels E and F shows results for the average years of schooling of men and women, respectively. Again, the estimated cohort effects appear approximately piecewise linear in shape, with a large trend break precisely at the 1947 cohort. For men, the cohort effects increase from a normalized 0 in 1930 to .83 by 1947, before suddenly changing trend and declining for subsequent cohorts. For women the pattern is simlar though the differences in cohort effects are smaller in magnitude; and the 1947 trend break is smaller in magnitude but followed by a second trend break in the early 1950s (similar to the pattern for women's log mortality). Appendix Figure A2 shows results separately for the share of each cohort with a high school degree, bachelor's degree, and advanced degree. Consistent with the prior literature reviewed above, they all show evidence of a clear, large trend break at or near the 1947 cohort. These figures also suggest that the differing shape for women is driven by the patterns at bachelor's and advanced degree levels; the shape of cohort effects for high school degree attainment of women is strikingly piecewise linear and nearly identical to that for men.

B. Possible cohort effects under additional assumptions

The above approach does not identify the long-run trend in cohort effects, and therefore cannot speak to whether cohort health and human capital declined in absolute terms. In this section for the non-educational attainment outcomes I show what different additional assumptions about age or year effects would imply about the sequence of cohort effects, allowing me to assess under different assumptions whether health and human capital declined in absolute terms.

Restrictions on period trend

For low birth weight and log mortality I proceed by considering the implications for cohort effects of different restrictions on the period trend. Lagakos et al. (2018) call this the "Deaton-Hall" approach in reference to Hall (1968) and Deaton (1997). By restricting the trend in period effects in different ways, one can pin down the trend in cohort effects based on the composite period/cohort trend defined above.[21]

I consider three benchmark cases: i) assuming that the period trend, β_p, is equal to 0, ii) assuming that the cohort trend, β_c, is equal to 0, and an intermediate case iii) assuming that β_p and β_c are equal.[22] Of course many other cases are possible, but examining what these assumptions imply about the sequence of cohort effects can help to make the mechanics of the model and its underidentification clear.

Additionally, it is arguably natural to think of case i) as an upper bound on the period trend, and therefore a lower bound on the cohort trend. For example, for low birth weight assuming $\beta_p \leq 0$ is assuming that average growth in period effects between 1968 and 1995 is non-positive, in other words that over this long period there was at a minimum not a worsening of external, non-cohort factors which impact the low birth weight rate. For mortality analogously, it would be assuming that between 1975 and 2019 the external environment impacting mortality rates did not worsen. Assuming $\beta_p \geq 0$, implies that $\beta_c \leq \tilde{\beta}_{pc}$ — placing a lower bound on the trend in cohort effects. This corresponds to the cohort trend in case i) above. Therefore we can view the sequence of cohort effects in case i) as a "best case" scenario for cohort effects, in the sense that they represent the most that cohort effects impacting low birth weight or mortality could have declined between the 1930 and 1965 cohort (or any two cohorts for that matter). Under this assumption the true sequence of cohort effects must lie above the case i) sequence, with the case ii) and iii) sequences representing two other possible sequences above the lower bound.[23]

Panel B shows cohort effects under the different scenarios for the low birth weight percentage. The bottom series shows cohort effects under case i), assuming that β_p, is equal to 0. Even under this assumption, which could be viewed as the "best case" for the long run trend in cohort effects, the cohort effects increase between the 1947 and the 1960 cohort by more than .25. This would imply that maternal health declined across these cohorts enough to increase the low birth weight percentage for these mother's by more than .25 percentile points. The series based on scenarios ii) and iii) imply much *worse* declines in maternal health between the 1947 and 1965 cohorts.

Panels C and D show analogous results for men's and women's log mortality. Again, the bottom series shows cohort effects under case i), assuming that β_p, is equal to 0. For these outcomes, the "best case" for the long run trend in cohort effects implies a small set of cohorts across which

[21]This also closely follows the bounding logic in **?**, who provide intuitive graphical intuition for the approach.

[22]Note that case ii) will yield the "detrended" cohort effects presented above.

[23]Note that case ii) is not necessarily an upper bound on the cohort trend. It is possible that the long-run cohort trend in positive, ie. that cohort health has declined over the full sample.

cohort effects were worsening in absolute terms. For men they would imply that the impact of cohort health on log mortality was worsening between the 1947 and 1952 cohort, before beginning to improve again. For women, they would imply that the impact of cohort health on log mortality was worsening in absolute terms only between the 1951 and 1957 cohorts. Of course these conclusions depend essentially on the assumption put on β_p, and the series based on cases ii) and iii) imply much larger increases in cohort effects after the 1947 cohort, and a larger range of cohorts across which cohort health was worsening.

No growth in wage age effects after 50

It seems hard to argue that external factors impacting wages have definitely not improved over time, ie. that the trend in period effects for wages is less than 0. For wages, I therefore make assumptions on the age effects. In particular, I adopt what Lagakos et al. (2018) call the Heckman-Lochner-Taber approach in reference to Heckman et al. (1998).[24] Lagakos et al. (2018) emphasize that the three main theoretical mechanisms for life cycle wage growth — human capital investment, search, and learning — all share the prediction of little or no growth in wages in the last few years of a workers career. For example, the incentive to invest in human capital is low when a worker does not expect to work much longer. This implies that the age effects at late ages will be zero, or slightly negative if human capital depreciates.

I implement this approach by first estimating a restricted version of the age-period-cohort model in Equation 2. I restrict the age effects from ages 50 to 54 to be equal to 0.[25]. This represents a first benchmark case i) of no growth in age effects between ages 50 and 54, with no depreciation of human capital. I then consider two additional cases in which human capital depreciation (or other factors) lead age effects between age 50 to 54 to i) decline by .005 per single-year-of-age, and ii) decline by .01 per single-year-of-age. Because the model is estimated on log wages these correspond to human capital depreciation rates of .5 $\%$ and 1 $\%$ respectively (no depreciation and 1 $\%$ are benchmark cases considered in Lagakos et al. (2018)).

The cohort effects implied by these three scenarios are shown in Panel A of Figure 3. All three scenarios show declines in cohort effects between the 1947 and 1965 cohorts — implying that median labor market ability declined across these cohorts. The magnitude of the decline differs substantially across the different assumptions. The bottom series shows cohort effects under the assumption that age effects are constant between ages 50 and 54, ie. that human capital does not depreciate. This series would imply that labor market skill declined between the 1947 and 1965 cohorts enough to reduce log wages by approximately .16. The middle series shows cohort effects

[24]They also reference the following papers as employing variants of the same idea: McKenzie (2006); Huggett et al. (2011); Bowlus and Robinson (2012); Schulhofer-Wohl (2018).

[25]Note, I do not omit the age 25 fixed effect in this specification. Also, note that this is a stronger assumption than is necessary to identify the sequence of cohort effects. It is chosen to correspond approximately to the benchmark case in Lagakos et al. (2018) of no experience effects in the last 5 years of a workers career.

under the assumption that age effects decline .005 per year for ages 50-54, ie. that human capital depreciates by .5 $\%$ per year. This assumption would imply instead that labor market skill declined between the 1947 and 1965 cohorts enough to reduce log wages by around .11. Finally, the top series shows cohort effects under the assumption that age effects decline .01 per year for ages 50-54, ie. that human capital depreciates by 1 $\%$ per year. This assumption of course does not change the size of the trend break in cohort effects, but because it implies a larger long run trend in period effects, it implies a much smaller absolute decline in cohort effects between the 1947 and 1965 cohorts, of around .02.

C. Summary

The above results from the age-period-cohort models are striking. However, they could in principle be biased by external factors which disproportionately impact individuals of particular ages in particular years. That is, cohorts born after 1947 may have been otherwise similar to earlier cohorts, but were just unlucky to have experienced bad conditions throughout their lifetime. Because the same pattern holds for very different outcomes, measured at different points in individuals lifetimes this "bad luck" would reflect a surprising coincidence. Each cohort born after 1947 would have to have been coincidentally exposed to poor obstetric conditions at the particular ages which they gave birth, to labor market conditions that were particularly bad for workers of their age, and to diseases and other mortality risk factors which specifically increased mortality for individuals of their age — which did not impact the outcomes of individuals of other ages in the same years. While such a coincidence seems surprising, below I implement two novel methodological approaches which allow for such external age-by-year interactions, and can still identify a relative decline in cohort health and human capital.

6. Evidence of cohort decline allowing for smooth age-by-year interactions

In this section I present evidence of a decline in cohort health and human capital using a novel methodology, which allows me to identify a trend break in cohort effects while allowing for smooth changes in external, age-specific factors across years. This approach yields similarly strong evidence, for all outcomes, of large changes in the trend of cohort effects at or near the 1947 cohort.

Consider the following model, which restricts the general model from Equation 1:

$$Y_{apc} = \underbrace{\beta \cdot c}_{\text{long-run trend in cohort effects}} + \underbrace{\mathbb{1}_{c \geq \lambda} \cdot \delta \cdot (c - \lambda)}_{\text{trend break in cohort effects}} + \underbrace{f^p(a)}_{\text{year-specific impact of age}} + \epsilon_{apc} \tag{4}$$

where as above Y_{apc} denotes an outcome — such as log wage — for individuals who are age a, in the year or "period" p, and who are members of the cohort c.

The first two terms on the right-hand-side specify the cohort effects as piecewise linear with a single, *unknown* trend break or knot. As above, β represents a long-run trend in cohort effects. The second term on the right-hand side now introduces a trend break in cohort effects at some unknown location λ. δ represents the size of this trend break, and λ estimates the cohort at which it occurs. This specification of the shape of cohort effects is similar to that seen visually in the detrended cohort effects estimated above and shown in Figure 2. I assume that the changing external age-by-year factors, $f^p(a)$, take the form of a polynomial of known order in each year. ϵ_{apc} is an orthogonal error.

The model is identified because the discontinuous change in the slope of cohort effects is orthogonal to the smooth, polynomial age-by-year interactions. The location and size of the trend break are still identified with the introduction of separate polynomials in age *in each year*.

I estimate the model by least squares, following the structural break methodology in Hansen (1999, 2000). The location of the trend break of these piecewise linear cohort effects is treated as an unknown parameter to be estimated. Algorithmically, I loop through different assumed values of break location λ, and select the location with the lowest sum of squared residuals. My baseline specification includes age fixed effects, year fixed effects, and a separate quadratic-in-age in each year as controls. I also probe the robustness of the estimated cohort trend break to including increasingly higher order polynomials in age in each year, up to a separate quartic in age in each year.[26]

Intuitively, I allow external factors which impact individuals of different ages to change but restrict them to do so smoothly. Most alternative explanations I want to rule out would take such a smooth form. For example, changes in the supply or demand for workers of different experience levels will have a smooth impact across ages, as long as individuals of nearby ages are sufficiently close substitutes. Labor economists traditionally control for quadratic experience terms in Mincerian wage regressions (Mincer, 1974). If the biological aging process is smooth then changes in the disease environment — such as the HIV epidemic — may disproportionately increase the mortality of young adults, but they will not discontinuously increase mortality for those under age 30 and have no impact on 31-year-olds. Log mortality is generally found to be remarkably *linear* in age (Gompertz, 1825; Chetty et al., 2016), suggesting that allowing for higher-order polynomials is quite unrestrictive.[27]

[26]For educational attainment I only include age fixed effects as important external year factors and age-by-year factors do not seem likely to drive educational attainment after age 25. Following standard practice, I employ an ad-hoc restriction to prevent the location of the cohort break λ to be estimated to be one of the youngest or oldest cohorts in the sample. In particular, in each year I restrict the location of the break to not be one of the 5 youngest or oldest cohorts.

[27]In an earlier paper I use this remarkable log linearity, known as Gompertz law, to provide graphical and statistical

To test for the existence of a trend break in cohort effects, I test the null that $\delta = 0$. A standard t-test would be invalid in this setting, because the location of the trend break λ is not identified under the null. I therefore follow the bootstrap procedure described in Hansen (1996, 2000).

Results

A graphical depiction of the estimation and inference of the location of the cohort trend break, λ, is shown in Figure 4. I invert the likelihood ratio statistic in Hansen (2000) to form 99 percent confidence intervals for λ. Each panel shows results for a different outcome variable, and plots the likelihood-ratio test statistic for different assumed locations of the trend break in cohort effects. The cohort with the minimum value of the likelihood-ratio test statistics yields the point estimate of the location, $\hat{\lambda}$. The 99 percent confidence region is those cohorts falling below the 1 percent critical value shown with a dashed grey line. For each of the six outcomes the cohort break location is precisely estimated and near 1947.

Table 1 provides the full results of estimating the piecewise linear cohort effects based on Equation 4. The estimated location of the cohort break — and therefore the implied cohort after which the health and human capital decline begins — are centered at the 1947 and 1948 cohorts. For both of the maternal health outcomes, infant mean birth weight and share low birth weight, the trend break is estimated to occur at the 1948 cohort. For both of these outcomes, the 99 percent confidence interval includes *only* a single cohort. For the men's median log wage it is estimated to occur at the 1947 cohort, with a confidence region including only 2 cohorts — 1946 and 1947. Cohort effects in models of the log mortality of men and women are estimated to have a trend break at the 1947 and 1949 cohorts respectively. The estimated cohort break location for years of schooling are 1948 and 1950 respectively. The estimated break locations for mortality and education are very precisely estimated — with the 99 percent confidence interval including only a single cohort.

The estimated trend breaks in cohort effects for all outcomes are large in magnitude and precisely estimated. The estimated size of the trend break, δ for the median log wage is -.016. This implies that the median man in the 1960 cohort has a wage roughly 23 percent lower than they would have had the cohort effects followed the pre-1947 cohort trend. For mean birth weight the size estimate of -6.35 grams implies that the 1960 cohort has given birth on average to 76.2 grams lighter infants, than they would have if the pre-1948 cohort trend had continued. More strikingly, the share low birth weight estimated trend break size of .23 implies that the 1960 cohort had a low birth weight rate near 2.8 percentage points higher than they would had the earlier cohort trend continued. The low birth weight rate nationally in 1975 was only 7.4 percentage points. Therefore

evidence for the importance of a cohort health decline in the recent increases in white mortality, in particular (Reynolds, 2023).

this effect of the health decline by the 1960 cohort is on the order of 40 percent. For men's log mortality the estimated size of the trend break is .028 and for women's log mortality it is .029, implying the single-year mortality risk for the 1960 cohort was roughly *1.5 times* what it would have been had the pre-1947/1949 trend continued.

For all six outcomes, I fail to reject null hypothesis of no break in the cohort trend at a very low significance level. As outlined above, I follow the bootstrap procedure described in Hansen (2000) to test the null hypothesis that no break in the cohort trend occurs, ie. that δ is equal to 0 and that cohort effects are linear. 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 linear cohort effects.

Appendix Table A1 show the robustness of these estimates to different specifications of the age-by-year control function. I examine the robustness to including different age-by-year interactions, from allowing none, to including separate quadratic, cubic, or quartic polynomials in age in each year. The results are generally quite robust.[28]

7. Non-parametric evidence of cohort decline

This section outline a novel, nonparametric generalization of McKenzie (2006) which identifies second differences in "Local Average Cohort Effects" under the assumption that age-by-year shocks are "locally exchangeable." I define Local Average Cohort Effects as the counterfactual average outcomes a cohort would have if they had experienced the external age-by-year factors of a cohort born one year earlier or later. I then show that applying this methodology to data on wages, mean birth weight by mother's birth cohort, and men's and women's log mortality yields strong evidence of the non-parametric equivalent of a trend break in cohort effects at the 1947 cohort.

Take as a starting point the general model described in Section 4. Treat the cohort factors θ_c as fixed and the age-by-year factors, ϵ_{ap}, are random variables.

[28]Results across *all* specification for both the mean birth weight and low birth weight percentage reveal estimated break locations between 1947 and 1949. The sign of the trend break estimates all imply a decline in maternal health. The magnitude of the implied decline *increases* once a quadratic-in-age in each year is added, and is stable with the addition of higher-order polynomials. For the median log wage, 3 out of the 4 specifications yield an estimated break at the 1947 cohort. Only the specification including cubic age-polynomials yields a different estimate — of 1953 — and the confidence region also includes 1946 to 1947, with a p-value of just .01 for the null that the break occurs at 1947. If one imposes the break location location to be 1947, the estimated break size is again negative and of nearly identical magnitude to the other specifications. For men's log mortality, including no age-by-year controls yields an estimated health decline beginning a few cohorts earlier, but with the addition of quadratic or higher polynomials the estimated location stabilizes at the 1946 or 1947 cohort. The magnitudes also stabilize near .029. For women's log mortality a cohort health decline, beginning between 1947 and 1950, is detected in models up to including a cubic-in-age in each year. When a quartic is included the estimated trend break turns negative. This instability may be the result of the cohort effects being misspecified as piecewise linear. Also note that findings for white mortality in Reynolds (2023) using a different methodology reveal highly robust evidence of a decline in cohort health located near 1946 for white men and 1949 for white women.

Define the "difference in local cohort effects" as:

$$\Delta^{ap}\psi_c = g(\theta_c, \epsilon_{ap}) - g(\theta_{c-1}, \epsilon_{ap})$$

This represents the difference between the outcome observed for a cohort, c, at year p and age a, and the outcome that the cohort born a year earlier, $c-1$, would have experienced had they been age a in year p and faced the same external age-by-year factors. So while $g(\theta_c, \epsilon_{ap})$ is observed, $g(\theta_{c-1}, \epsilon_{ap})$ is an unobserved potential outcome.

Then define the "second difference in local cohort effects":

$$\Delta\Delta^{ap}\psi_c = \Delta^{a-1,p}\psi_{c+1} - \Delta^{ap}\psi_c$$

This can be thought of as a non-parametric generalization of the second-difference in cohort effects considered in McKenzie (2006). For example if the additively-separable age-period-cohort model of equation 2 were the true data generating process, then $\Delta\Delta^{ap}\psi_c$ would be equal to: $(\gamma_{c+1} - \gamma_c) - (\gamma_c - \gamma_{c-1})$, for all a, p.

Also define the expected average second difference in local cohort effects over a set of P years $\mathcal{P}$, where expectation is taken with respect to to the (unspecified) distribution of the age-by-year shocks:

$$\Psi_{c,\mathcal{P}} = E\left[\frac{1}{P}\sum_{p\in\mathcal{P};a=p-c}\Delta\Delta^{ap}\psi_c\right]$$

This definition takes seriously the idea that the "human capital" of a cohort may be multidimensional and therefore it's impact may differ depending on the external factors prevailing in a given year for those of a given age. It therefore takes as a metric for comparing "human capital" of different cohorts the outcomes that would be observed for them given the external factors prevailing in a particular year for those of a given age. Consider wages as the outcome. If $\Delta\Delta^{ap}\psi_c$ is negative it implies that given the labor market in year p for workers of age a: the counterfactual difference in wages between what cohort $c+1$ would have earned and what c did earn, is larger than the difference between what cohort c earned and the counterfactual wages of cohort $c-1$. If for a given c, one knew that for many years and ages that $\Delta\Delta^{ap}\psi_c$ is on average large and positive — we might conclude that there is a "trend break" or "kink" in human capital at cohort c in the sense that the difference in counterfactual wages is on average smaller between cohort c and cohort $c+1$ than it is between cohort c and cohort $c-1$.

For example, suppose (as was suggested by the results above) that the change in human capital between the 1947 and 1948 cohort is much smaller than change in human capital between 1946 and 1947 cohort then $\Delta\Delta^{ap}\psi_{1947}$ with wages as the outcome should be large and negative, because: $\Delta\Delta^{ap}\psi_{1947} = \Delta^{a-1,p}\psi_{1948} - \Delta^{ap}\psi_{1947} = [g(\theta_{1948}, \epsilon_{ap}) - g(\theta_{1947}, \epsilon_{ap})] - [g(\theta_{1947}, \epsilon_{ap}) - g(\theta_{1946}, \epsilon_{ap})]$

A. Unbiased estimator under exchangeabililty of age-by-year shocks

The key assumption for identification will be that the age-by-year shocks in neighboring years hitting an age a are "exchangeable" with the age-by-year shocks in the same years hitting age $a-1$. That is that, for all a, p:

$$F(\{\epsilon_{ap}, \epsilon_{a,p-1}\}, \{\epsilon_{a-1,p}, \epsilon_{a-1,p-1}\}, \boldsymbol{\epsilon}) = F(\{\epsilon_{a-1,p}, \epsilon_{a-1,p-1}\}, \{\epsilon_{ap}, \epsilon_{a,p-1}\}, \boldsymbol{\epsilon})$$

where $\boldsymbol{\epsilon}$ denotes the vector of all other $\epsilon_{a',p'}$ not listed, and $F()$ denotes the joint probability distribution of the entire sequence of shocks. The pairs of shocks are exchangeable in the sense that the joint distribution of all of the age-by-year shocks in neighboring ages and years are invariant to permuting the "label" of ages on the shocks.[29]

For example, it means that it is "as-if-randomly-assigned" whether the set of shocks in neighboring years hits 30-year-olds or 31-year-olds. This assumption is somewhat unusual but has a number of appealing properties. It allows for unrestricted dependence in the shocks over time hitting a given age a. It also allows for there to be non-random shocks to a larger groups of ages in a given year — for example the effect of the supply of workers of nearby ages on wages in Card and Lemieux (2001a). What is assumed random is the very-local difference in external factors impacting those one year apart in age.

Under this assumption, one can show easily that the second difference estimator suggested in McKenzie (2006) is an unbiased estimator of the average "second difference in local cohort effects." Begin by defining the first-difference in outcomes of those who are the same age, in neighboring years: $\Delta^p Y_{apc} = Y_{apc} - Y_{a,p-1,c-1}$. And define the second difference in outcomes, which takes the difference in the above between neighboring ages:

$$\Delta^a \Delta^p Y_{apc} = \Delta^p Y_{a-1,p,c+1} - \Delta^p Y_{apc} = (Y_{a-1,p,c+1} - Y_{a-1,p-1,c}) - (Y_{apc} - Y_{a,p-1,c-1})$$

The idea is that the second difference in outcomes, is equal to "second difference in local cohort effects" plus the difference in age-by-year shocks for that year:

$$\Delta^a \Delta^p Y_{apc} = \underbrace{[g(\theta_{c+1}, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p})] - [g(\theta_c, \epsilon_{a,p-1}) - g(\theta_{c-1}, \epsilon_{a,p-1})]}_{\text{cohort second diff}} + \underbrace{(g(\theta_c, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p-1}) - (g(\theta_c, \epsilon_{a,p}) - g(\theta_c, \epsilon_{a,p-1})}_{\text{diff in age-by-year shocks}}$$

I use as an estimator the average of these second-differenced outcomes across a set of P years $\mathcal{P}$. This will be equal to the average "second difference in local cohort effects" plus the average difference in age-by-year shocks:

[29] See for example Draper et al. (1993); Bernardo (1996) in statistics and Brock and Durlauf (2001) for an economic application.

$$\hat{\Psi}_c \equiv \frac{1}{P}\sum_{p\in\mathcal{P}} \Delta^a\Delta^p Y_{apc} = \underbrace{\frac{1}{P}\sum_{p\in\mathcal{P}} [(g(\theta_{c+1},\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p})) - (g(\theta_c,\epsilon_{a,p-1}) - g(\theta_{c-1},\epsilon_{a,p-1}))]}_{\text{avg. cohort second diff}}$$
$$+ \underbrace{\frac{1}{P}\sum_{p\in\mathcal{P}} [(g(\theta_c,\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p-1}) - (g(\theta_c,\epsilon_{a,p}) - g(\theta_c,\epsilon_{a,p-1}))]}_{\text{avg. diff in age-by-year shocks}}$$

Then under the assumption of exchangeability the average difference in age-by-year shocks goes to zero in expectation, and the estimator is equal in expectation to just the expected average "second difference in local cohort effects":[30]

$$E\left[\hat{\Psi}_c\right] = E\left[\frac{1}{P}\sum_{p\in\mathcal{P}} \Delta\Delta^{ap}\psi_c\right] = \Psi_c$$

B. Permutation-based inference

The exchangeability assumption allows for a simple permutation test in the spirit of Fischer's exact test.[31] I will test the so-called "sharp null" hypothesis that the second-difference in cohort effects is equal to zero in all years. Note that this is different than testing the null that the *average* second difference in cohort effects is zero.

$H_0 : \Delta\Delta^{ap}\psi_c = 0$ for all a, p s.t. $c = p - a$

Appendix A provides more details on the test and shows that it has correct size in expectation. I take the set of observed first differences for a cohort $\tilde{c}$ and the neighboring cohort $\tilde{c}-1$: $\{\Delta^p Y_{ap\tilde{c}} | p - a = \tilde{c}\}, \{\Delta^p Y_{a,p,\tilde{c}-1} | p - a = \tilde{c} - 1\}$. I make repeated draws where in each draw: for each pair of neighboring first differences I randomly either permute them or keep them as observed. I use as a test statistic the second difference estimator $\hat{\Psi}_c$. The key idea is that under the dual assumptions of exchangeability and the null hypothesis, the sequence of test statistics calculated from the observed data and each of the permuted versions is a sequence of exchangeable random variables. Therefore by comparing the test statistics from the observed data with those from the permuted versions one can gauge how unlikely it would be to see the observed second difference by chance, under the null.

I also construct confidence intervals by inverting the permutation test following the approach in Imbens and Rubin (2015) and Ganong and Jäger (2018). The confidence interval can be interpreted as the set of constant second-difference in local cohort effects which the test fails to reject.

[30] Recall that only the age-by-year shocks are random variables, so expectation is taken with respect to their distribution.

[31] See for example Ernst (2004), Imbens and Rubin (2015), and Young (2019).

In particular I find interval the B such that for all $b \in B$ would fail to reject null hypothesis that: $H_0 : \Delta\Delta^{ap}\psi_c = b$ for all a, p s.t. $c = p - a$.

C. Results

The results of applying the above strategy to estimate the second difference in local cohort effects, shown in Figure 5, reveal striking evidence evidence of the non-parametric equivalent of a trend break in cohort effects located at the 1947 cohort for wages, maternal health, and mortality of men and women. For each outcome, the figures on the left show the sequence of point estimates $\hat{\Psi}_c$ for the cohorts between 1937 and 1957. Panels A and B show that for both median log wage and mean birth weight (by mother's birth cohort) the estimated average second difference in local cohort effects for 1947 are negative and large in magnitude — around 2 or 3 times larger in magnitude than any other cohorts' estimates. For both outcomes the other cohorts' estimated average second differences in local cohort effects are near 0 — with the confidence intervals including zero for all but one other wage estimates and all except two other birth weight estimates. Panels C and D show results for men's and women's log mortality. For both outcomes, the estimated average second difference in local cohort effects for 1947 are positive and large in magnitude, again on the order of 2 times the magnitude of any other estimates. For these outcomes, however the other cohort's estimates are much less tightly centered around zero however — oscillating around zero with the confidence interval often not including zero.

The panels on the right graphically illustrate the permutation-based test of the sharp null that all of the second difference in local cohort effects for the 1947 cohort are 0. The dashed line shows the point estimate of $\hat{\Psi}_{1947}$ and the histogram shows the permutation distribution based on 10,000 simulations. The implied p-value of the "sharp null" hypothesis that the second-difference in cohort effects is equal to zero in all years for log wage is .0013, while those for the other three outcomes are .0002.

Interestingly, the evidence of a relative decline in health and human capital beginning suddenly after the 1947 cohort using the nonparametric approach is *even stronger* than that shown above using parametric approaches. This is true with regards to the precise timing: with the 1947 cohort having the largest estimated magnitude of average second differences in local cohort effects in *all* four outcomes. In contrast, the estimated break locations shown in Table 1 for different outcomes are all near 1947, but the exact cohort where the trend break is estimated to occur varies slightly between 1946 and 1950. This could be due to misspecification in the additively-separable cohort trend break model (as well as the traditional additively-separable age-period-cohort models). In particular, the imposition that the size of the trend break in cohort effects is of the same magnitude in all years/ages is completely relaxed in the nonparametric approach presented here.

The magnitudes of the average second differences in local cohort effects are also all much *larger* than the estimated magnitude of the trend breaks for the same outcomes from the additively-

separable cohort trend break model. This could again be due to misspecification in the additively-separable cohort trend break model. Although it is important to note that the nonparametric estimates are only informative about very local differences between the 1946, 1947, and 1948 cohorts; while the other models if correctly specified are informative about differences in cohort health and human capital more broadly for the cohorts born 1930 to 1965 (up to trend).

The evidence above comes from a quite unrestricted model with weak assumptions. The remaining threat to validity would be non-random changes in the impact of age across years between neighboring ages. For example, a large shock in a given year to the health of those age 30 and under, which did not impact those age 31 and over. Alternative external explanations, or more broadly external factors impacting each of the four outcomes, are generally thought to be smooth in age. For example, the effects of shifts in supply or demand on wages will be smooth as long as individuals who are close in age are substitutable (Card and Lemieux, 2001a). Similarly, the effect of "biological aging" on health is generally thought to be a smooth, continuous process.

Discrete policy cutoffs based on age could seem to be a threat, but note that to generate the above results they would have to "follow" the same cohort over time. For example, moving from age 30 in 1997, to age 31 in 1998, to age 32 in 1999 and so on. A large shock in a single year to one age and not the neighboring age would not yield a statistically significant estimate because the permutation-based-inference procedure would correctly reveal that such a pattern is not particularly "unlikely" under the null.[32]

8. Role of cohort decline in year-over-year declines

The estimated cohort health and human capital decline is large in magnitude. In this section, I present the results of a simple simulation which demonstrates the contribution of this cohort decline to secular year-over-year trends in the outcomes studied above. The results suggest that absent the cohort health decline: the increase in low birth weight births in the 1980s would not have occurred, there would have been modest real wage growth since the late-70s rather than real declines, and midlife mortality of men and women would have continued to decline steadily at it's pre-1999 rate.

I perform a simple simulation to examine the counterfactual in which there was no cohort health decline. For each outcome, I use the estimated location, $\hat{\lambda}$, and size, $\hat{\delta}$, of the cohort trend break from the models estimated above based on equation 4.[33] For each outcome I "remove" the

[32]Consider a simple example where in one year the second difference in outcomes for the 1947 cohort is equal to Z, an economically meaningful magnitude, but in all other years the second difference is zero. The point estimate would be Z, and the resulting permutation distribution would have half it's mass at Z and half it's mass at $-Z$ — implying the outcome is quite likely under the null.

[33]To be conservative, for the low birth weight percentage I use estimates from models with only age and year fixed effects, for which the trend break is of smaller magnitude. For the other outcomes, I use estimates from the baseline specification reported in Table 1.

cohort trend break, by subtracting it from each observation. That is, I create a transformed dataset, where each observation takes the following form: $\tilde{Y}_{apc} \equiv Y_{apc} - \mathbb{1}_{c \geq \hat{\lambda}} \cdot \hat{\delta} \cdot (c - \hat{\lambda})$. I then calculate summary measures of each outcome by year using both the raw data, Y_{apc}, and the transformed data, $\tilde{Y}_{apc}$, in which the estimated cohort trend break has been removed. Comparing the trends in these summary measures then reveals the contribution of the trend break in cohort effects to the year-over-year trend.

Figure 6 shows the results. Panel A shows trends in the percent of infants born at low birth weight. The raw series shows that while the low birth weight rate declined from above 8 percent in 1968 to 6.7 percent by 1985, it then reversed trend and increased to near 7.3 percent by 1995 (Centers for Disease Control, 1994). In contrast, the transformed series in which the cohort trend break has been removed exhibits no similar increase. The series implies that absent the cohort health decline, the low birth weight rate would have declined much more rapidly until 1985 — falling to 5.7 percent by that year. These improvements would have slowed after the mid-1980s — but the low birth weight rate would have continued to decline slowly until 1995 to just below 5.5 percent.

Panel B shows similar results for men's wages. In particular it shows the average across men age 25 to 54 of median wages by single age-bins.[34] This measure of wages declines in real terms from 1979 to 1993, from 3.14 to below 3.04. By this measure real wages declined from 23 to below 21 dollars an hour, in constant 2014 dollars. In contrast the transformed log wage series, in which the cohort trend break is removed, increases very slightly from 3.16 in 1979 to 3.17 in 1993.

Panel C and D show analogous results for the age-adjusted mortality rates of men and women age 45 to 54, the ages focused on in Case and Deaton (2015). The raw data show declines in men's and women's midlife mortality between 1975 and 1990. A "bump" is then evident during the AIDS epidemic, larger for men. Both series then exhibit a clear change in trend near 1999 and slowing improvements in the mortality rate after that year. For men the mortality rate even increases between 1999 and the mid-2000s. However, this change in trend is completely absent in the transformed series in which the cohort trend break in log mortality is removed. That is, the decline in cohort health can completely explain the stagnation — and for men slight increase — in mortality after 1999. Panels E and F show the mortality rates of older men and women, aged 55 to 64. Mortality for both sexes declines between 1975 and around 2010, before changing trend and increasing. Again the change in trend is absent in the simulated series in which the cohort trend break is removed — suggesting the mortality increases for this group are also plausibly driven by the cohort health decline.

[34] If log wages are assumed to take a symmetric distribution, such as if wages are assumed to be log normal, then this provides an estimate of age-adjusted mean log wages. Where the age-adjustment is made with a uniform population distribution.

To provide an even closer comparison with Case and Deaton (2015, 2017), Figure A4 shows the results of this simulation exercise for midlife white mortality rates. Below I show evidence that a trend break in cohort effects is also evident for white mortality analyzed separately, so I use the estimated trend break in log mortality of whites in this simulation. The results suggest the decline in cohort health can completely explain the increase in midlife white mortality after 1999. My cohort-based theory therefore provides an explanation for the previously surprising timing of the midlife mortality increases — beginning in 1999, during an economic boom. Midlife mortality changed trend in 1999 not because of factors specific to that year, but because "unhealthy" cohorts born after the late-1940s began to age into the 45 to 54 age group after that year.[35]

9. Evidence suggesting decline originated in childhood

This section presents three sets of results suggesting that the cohort decline likely originated in childhood or adolescence: i) growth and physical development in adolescents was negatively affected beginning with approximately the same cohorts, ii) scores on standardized tests taken as early as 16 declined for the same cohorts, ii) the decline is concentrated among the US-born population with no evidence of a similar cohort decline for the foreign-born population.

A. Growth and physical development

Previous research has found that increases in adult height, often viewed as a proxy of childhood nutritional status (Tanner, 1990; Floud et al., 2011), suddenly stopped for Americans born between approximately 1955 and 1974, while the height of Europeans continued to grow rapidly (Komlos and Lauderdale, 2007a,b; Komlos, 2010). I build on these findings for adult height and also examine height and physical development of these cohorts as children. More detail on the data and exact specifications for the results in this section are given in Appendix B.

Stagnating adult height

Figure 7 panels A and B show binned scatter plots of height by birth cohort for men and women.[36] Following Komlos and Lauderdale (2007b,a); Komlos (2010), I combine data on measured heights from multiple waves of the National Health and Nutrition Examination Survey (NHANES) and it's precursor conducted in various years since 1959.[37] The sample sizes are

[35]This argument is developed in more detail in my earlier paper which focused exclusively on the contribution of a cohort health decline to recent increases in the mortality rate of white Americans (Reynolds, 2023). Figures A3 and A4 shows that the differential timing of overall and white mortality increases in different age groups can also be explained by the cohort-based theory.

[36]The methodology follows Cattaneo et al. (2019). The number of bins is choosen by their data-driven IMSE-Optimal selection method, with the bins restricted to be equally spaced. The confidence intervals are constructed based on a cubic B-spline estimate of the relationship between birth cohort and height.

[37]As in these past articles I focus on adults age 23 to 47, ages at which height is approximately constant, individuals born in the US, and use survey weights adjusted for the combined surveys as suggested in Korn and Graubard (1999). I include 12 more years of data than Komlos (2010), bringing the data up to 2018.

much smaller than those for the main outcomes examined above, leading to imprecision, but some patterns emerge. The general pattern for men and women are similar: an increasing trend from the 1920s cohorts to approximately the 1950s cohorts is clearly evident — followed by a stagnation beginning at some point in the 1950s and nearly constant heights thereafter. Estimating simple trend break regressions and imposing a trend break in 1947 yields for men: an estimated pre-1947 trend of around .4 inches of growth per decade, a large trend break and a trend of only around .08 inches of growth per decade after the 1947 cohort. For women the same estimates suggest increase of around .3 inches per decade before the 1947 cohort, and then only .03 inches per decade thereafter (indistinguishable from zero at a 95 % level). Estimation of Hansen-style trend break models with an unknown break location, as above, are imprecise but confidence intervals for the break location include the 1947 cohort for both men and women.

Stagnating childhood height

Growth in childhood height also appears to have stagnated around the 1950 birth cohort, at least for white boys. Surprisingly this was already noted in the 1970s by reports of the National Center for Health Statistics (Hamill et al., 1977b,a).[38] Trends in childhood height need not exhibit the exact same pattern as adult height. Height at various ages in childhood depends on so-called growth tempo, while adult height depends on both growth tempo and the age at which growth stops (Tanner, 1990; Cole, 2003).

Panels C shows evidence of this stagnation in heights of white boys at age 6. The pre-1950 data come from various historical studies which were compiled in Meredith (1964). They do not come from a nationally representative sample, but also did not explicitly target high or low socioeconomic status children, for example. Each point shows the mean reported in Meredith (1964) for an interval of birth cohorts, plotted at the middle cohort of the interval. The post-1950 data show sample means for white non-Hispanic males from the combined NHES and NHANES data for 5-year birth cohort intervals, again plotted at the middle cohort. At age 6 there is clear visual evidence of an increasing trend in heights from the 1870 to approximately 1950 cohort, from 43 to more than 47 inches. However, this growth suddenly stagnates and for the next 60 cohorts there is little or no growth in heights. The estimated trend in height for the pre-1950 cohorts is .45 inches per decade, but slows to an estimated trend of only .09 inches per decade after 1950.[39]

Later adolescent growth spurt and delayed puberty

The NHES and NHANES samples are relatively small and the early surveys occured irregu-

[38]The report describing new child growth curves estimated based on data from NHES and NHANES surveys (Hamill et al., 1977a) reads: "In the analysis of these data, the marked diminution and near cessation of the trend to constantly increasing size of successive generations of American children is the most dramatic and significant finding relating to human biology and human growth in general."

[39]Meredith (1964) also compiled data for 10 year old boys. Heights at this age shows evidence of a qualitatitively similar trend break when combined with NHES and NHANES data.

larly and for selected ages. However, (by lucky coincidence) it is possible to observe heights from age 6 though adulthood of the cohort born between 1951-1957 and compare them to heights at the same ages for those born 1960-1967, for relatively large samples.[40]

Panel E of Figure 7 shows the estimated differences in height at three ages, between native-born males in the 1951-1957 and the 1960-1967 cohorts.[41] The estimated differences in height are small and statistically indistinguishable from zero both in childhood, at age 6 to 11, and in adulthood, at ages 23-40. However, I estimate a statistically significant difference in height of *nearly half an inch* in adolescence, at ages 12 to 17.

These results imply that the two cohort groups were nearly identical in height as children but the later born cohort had a much smaller or later growth spurt in adolescence — causing them to fall behind in height at ages 12 to 17. The later born cohort then caught back up in height in adulthood.

Panel F of Figure 7 shows estimates of the differential growth rates implied by these differences in heights. The point estimates imply that the later-born cohort grew .35 inches less between ages 6-11 and ages 12-17 than the earlier-born cohort. However, this reversed between ages 12-17 and adulthood, when the later-born cohort then grew .43 inches *more* than the earlier-born cohort. The differences in growth are statistically significant.

The results appear consistent with males in the later born cohorts having a delayed adolescent growth spurt but then growing longer into early adulthood, allowing them to "catch-up." It is common for individuals or populations who have a later growth spurt in adolescence to then continue growing for longer into their early twenties, at least partially catching up in adult height (Tanner, 1986; Lleras-Muney et al., 2022). This later growth also particularly occurs in the legs as opposed to the upper body — due to later ossification of the leg bones (Tanner, 1986; Lam et al., 2021). Consistent with this explanation, Appendix Figure A5 shows that the differences in height and growth shown in Figure 7 are nearly entirely driven by differences in leg length, with "trunk length" not statistically different between the two sets of cohorts.

Panels A and B and Appendix Figure A5 show that there is no evidence of a similar pattern of differential growth for females born in 1960-66 versus those born in 1951-1957. The point estimates for the differences in height at all three ages are small in magnitude and not statistically different from zero, although differences of .2 to .35 inches cannot be ruled out at standard levels.

[40]The second NHES surveys was carried out in 1963-1965 and only sampled 6-11 year olds, and the third was carried out in 1966-1970 and only sampled 12-17 year olds. Therefore in these two surveys one can observe the 1952 to 1957 cohorts heights at various ages of childhood for relatively large samples. NHANES 1 and 2 sampled ages 1 (or zero) to 74 and were carried out in 1971-74, 1976-80 and allow one to observe the 1960-1967 cohort at similar ages and again with relatively large sample sizes. The remaining surveys allow one to observe adult heights.

[41]The estimates for 6-11 and 12-17 year olds come from a regression which includes age-in-years dummies and a quadratic in age-in-months, to account for small differences in age distribution of the two cohorts. All regressions use sampling weights.

However, there is other evidence that the females in these cohorts experienced puberty later. Panel D reports binned scatter plots of the age at menarche, the age at which girls have their first menstrual period, based on retrospective reports from adults in the NHES and NHANES surveys. The estimates suggest that average age at menarche declined from nearly 13 to less than 12.4 between the 1920 and approximately 1947 cohort. Then suddenly this trend towards earlier menarche stopped and instead age at menarche increased until the mid-1960s cohorts to around 12.75, before declining again for subsequent cohorts.

In addition to having worse outcomes as adults, it appears that the post-1947 cohorts went through puberty later. More research is needed but this strongly suggests that something had already gone wrong for these cohorts by at least adolescence. Puberty is generally later for individuals of lower socioeconomic status and there has been a secular trend towards earlier puberty over the last 200 years in many high-income countries.[42] For example, Lleras-Muney et al. (2022) find that the timing of the adolescent growth spurt and age at menarche are both negatively associated with socioeconomic status in a large sample of low-income countries. While the exact biological mechanisms determining the age of puberty and the adolescent growth spurt are not fully understood, the trend over time and differences by income are thought to be driven by differences in nutritional status and disease exposure (e.g. Cole, 2003). Recent findings suggest that caloric deprivations in childhood activate particular receptors in the brain which delay the onset of puberty and the adolescent growth spurt, but also lead to a period of leg growth longer into early adulthood (Lam et al., 2021). Broadly, it seems likely that these cohorts had some kind of hormone or nutrient deficiency in adolescence — either due to a shock occurring during adolescence or an earlier shock which particularly affected biological processes related to puberty.

B. Test score decline

There was also a widely noted decline in standardized test scores beginning in the late 1960s (eg. National Commission on Excellence in Education, 1983; Koretz, 1987), which appears to closely match the timing by cohort of declines in other measures of human capital shown above. Most widely noted was a decline in Scholastic Aptitude Test (SAT) scores beginning in the early 1960s and lasting until the mid-1980s. Panel B of Figure 8 shows SAT scores by approximate birth cohort, measured as the school year that the test was taken minus 17. By this measure, the average score on the Verbal portion of the SAT was nearly flat at around 475 points between the 1940 and mid-1940s cohorts. Scores then began to drop precipitously for cohorts born after the late-1940s, declining from a peak for the 1946 cohort of 478 points to 426 points by the 1965 cohort. Similarly,

[42] See for example Lleras-Muney et al. (2022); Wyshak and Frisch (1982); Hauspie et al. (1997); Cole (2003) and the citations therein. There is evidence the trend has slowed or stopped in some countries. Long run changes in height-for-age are also generally largest in the teenage years for boys, consistent with a long-run trend towards an earlier growth spurt Meredith (1964); Cole (2003).

scores on the Math portion of the test were flat near 500 points for the 1940 to mid-1940s cohorts, and then also began to fall rapidly: from 502 for the 1946 cohort to 468 by the 1965 cohort.

A large literature studying this decline concluded that it could not be explained by changes in the composition of test-takers or changes in the difficulty of the test but failed to find a conclusive cause (eg. Koretz, 1987). A review commissioned by the College Board suggested that the declines in the 1960s could be explained to some extent by a changing composition of test takers but that there was little scope for such observable selection effects in explaining continued declines after 1970 (Price and Carpenter, 1978). The share of 17-year-olds taking the test also remained nearly flat over the 1970s, ruling out a simple single-index selection story in which lower ability individuals were pulled into the test-taking population. The declines were also evident at the top of the score distribution, with the share of test-takers achieving perfect scores and scoring above 700 points both also falling in the 1970s. Finally, similar declines in test scores are evident in other achievement tests besides the SAT, including nearly universal tests given in Iowa and Minnesota (Harnischfeger and Wiley, 1975).[43]

C. Only evident for native-born

In this section, I present evidence that the cohort decline is only evident for individuals born in the US, not for foreign-born residents. The lack of a cohort decline for foreign-born residents can be viewed as further evidence that the cause of the declines was an early life factor specific to the US; with the caveat that the foreign-born estimates could be driven by changing selection of immigrants. The null result for the foreign born also gives credibility that the finding of a cohort decline is not a mechanical result of model misspecification.

I first use the natality data to examine mean birth weight by mother's place of birth. Panel A of Figure 8 shows detrended cohort effects separately for foreign and native-born mothers, based on additively-separable age-period-cohort models.[44] As for the earlier figures, the models are based on Equation 2, and assume no age-by-period interactions. The shape of cohort effects for native-born have a clear piecewise linear pattern, with a large break at the 1947 cohort — consistent with a decline in maternal health after that year, relative to trend. In sharp contrast, the cohort effects for foreign-born are nearly linear and exhibit no trend break, consistent with no relative decline in maternal health for mothers born outside the US.

Next, Table 2 shows the results of estimating models with piecewise linear cohort effects separately for foreign and native-born mothers. For both groups, I examine the robustness to including different age-by-year interactions, from allowing none, to including separate quadratic, cubic, or

[43] In contrast scores on intelligence (IQ) tests meant to capture general reasoning ability famously continued increasing across these cohorts (Flynn, 1984).

[44] I drop the first two years of data, 1968 and 1969, because mother's place of birth is not recorded in those years. Using the remaining years I estimate models separately for mothers born within one of the 50 states, and for those born outside of them.

quartic polynomials in age in each year. For the native-born, all specifications yield precisely estimated trend breaks between the 1947 and 1949 cohorts implying large declines in maternal health. In contrast the results for foreign-born mothers vary widely across different specifications — and do not consistently show evidence of a decline in maternal health. For example, in my baseline specification with a quadratic-in-age the bootstrap-based F-type test fails to reject the null of no trend break. Further the sign of the break is unstable across specification — some models suggest improvement in health while others suggest a decline in health.

Estimating wages and mortality separately for the foreign and native-born is not possible in the data I use. I use decennial census data to examine cohort declines in occupational status for the foreign and native-born separately. Table A2 reveals a quite similar pattern in white-collar status to that shown above for maternal health.

10. Preliminary investigation into root cause

I provide a preliminary investigation of the underlying cause of the apparent early life health decline of these cohorts.

A. Link between educational declines and other declines

In general, the educational declines for the post-1947 cohorts occurred before their declines in other outcomes. A natural question is therefore whether the direct causal effect of education on wages and health can explain the other declines in health and wages documented above.

Two main findings detailed in Appendix C point against such an "education-only" explanation. First, the increase in log mortality is too large to be explained by the decline in wages, unless one assumes the causal effect of education on mortality is much larger than commonly found. Second, a large cohort decline is evident at or very near the 1947 cohort in conditional maternal health (as proxied by infant birth weight) across the maternal educational distribution, from mother's with less than high school to those with more than 5 years of college education. In principle this could be explained by complex changes in the selection of women by maternal health after the 1947 cohort — but a broad decline in health appears more likely.[45]

B. Decline widespread among native-born

The cohort decline appears to be remarkably widespread across demographic groups born in the US. There is only one other clear finding of systematic heterogeneity. While a post-1947 cohort decline in maternal health appears to have occurred for residents of counties of all population densities, the cohort decline in maternal health appears to be *larger* in the densest counties. These

[45]Interestingly, the cohort decline in wages appears to be concentrated among those without a Bachelor's degree, suggesting a possible connection to the large literature on the college high-school wage gap.

findings suggest that the cohort declines were likely caused by an early-life factor, widespread within but specific to the US — which was potentially worse in the densest urban areas.

By race

Table 3 examines potential heterogeneity by race for four of the main outcomes. I estimate the piecewise-linear cohort effect models based on Equation 4, separately for different racial groups.[46] For nearly all of the outcomes there is evidence of a large cohort decline beginning near the 1947 cohort. The one notable exception is the mortality of black men, which shows evidence of improvements for post-1952 cohorts. However, black women exhibit evidence of cohort declines in mortality similar to those for whites, and black men and women exhibit cohort declines in all the other outcomes. Further, the wage decline for black men is significantly *larger* than that for whites.

By Region

Table 4 examines analogous heterogeneity by region of residence, and shows that very similar declines in cohort health and human capital are estimated across each of the four Census Regions.

By state-of-birth and commuting zone

I turn to estimating the size of the cohort decline separately for each state-of-birth (SOB) and for each community zone (CZ). Because of the relatively large number of states and counties, I use a two-step empirical Bayes procedure.[47] I focus on maternal health because the natality microdata has by far the largest sample and is the only dataset that records state-of-birth.

I collapse the natality microdata by either SOB or CZ. I then estimate the piecewise-linear model of Equation 4, with λ fixed at 1948, separately for each SOB/CZ. I use low birth weight percentage as the dependent variable and include as controls age fixed effects, year fixed effects, and a separate quadratic-in-age in each year. I treat the estimated means and variances by SOB/CZ of the size of the cohort decline, δ as priors. In the second step, I use the Gaussian Empirical Bayes procedure suggested in Morris (1983) to form posterior estimates of: i) the size of the decline by SOB/CZ, ii) the mean and variance in the size of the decline across *all* SOBs/CZs. These estimates take into account sampling error and therefore "shrink" the noisy first-stage estimates towards the "grand mean" across all SOBs/CZs. Noisier estimated declines, for example those in small commuting zones or states, are shrunk more.

The results, shown in Figure 9, suggest that the cohort decline is remarkably widespread.

[46]Given the smaller sample sizes I report estimates from models including just age fixed effects and year fixed effects as controls. I also restrict the sample to exclude cohorts born after 1960, to avoid focusing on the known improvements for blacks born after the Civil Rights Act documented in Almond and Chay (2006), Chay, Guryan and Mazumder (2009, 2014).The CPS data includes individuals of Hispanic-origin as a separate category, while the vital statistics data does not consistently include Hispanic-origin. For the log wage results I therefore estimate models for 4 racial groups: non-Hispanic whites, non-Hispanic blacks, Hispanics, and all other races pooled. For health outcomes, I estimate models for only 3 categories: whites, blacks, and other races (which each include Hispanics).

[47]For economics applications of the Gaussian Empirical Bayes procedure see McClellan and Staiger (2000); Kane and Staiger (2001); Chandra et al. (2016) among others.

Panels A and B show a map and histogram of the posterior estimates of the size of the decline by mother's state of birth. The posterior estimates of the size of the cohort trend break for all 50 states are greater than 0 — consistent with a cohort health decline occurring for mother's born in every state. Forty states have estimated breaks between .1 and .3. The estimates suggest that the decline was smallest for women born on the West coast and neighboring states. There is also a cluster of states in the Southeast with particularly large breaks: Arkansas, Louisiana, Mississippi, and Alabama. The remaining heterogeneity does not show an obvious spatial pattern.

Panels C and D of Figure 9 show analogous results by mother's commuting zone of residence. The posterior estimates of the size of the cohort trend break is greater than 0, implying a decline in cohort health, in 98 percent of the commuting zones in my sample. Further, 72 percent of commuting zones have estimated breaks between .1 and .3. It is hard to recognize patterns in the map as the estimates are much more variable, but there appear to be clusters of particularly large breaks in the Deep South, as well as in the Rust Belt. Note that these results are based on the location where women in the cohorts of interest are living when they give birth. With selective migration this of course could differ systematically from the locations where the women themselves were born.

Heterogeneity by county characteristics

I combine the Natality microdata with historical county characteristics from the County Data Book to examine potential heterogeneity in the magnitude of the cohort decline across counties with different observable characteristics. I again focus on maternal health because of the large sample size of the natality microdata.

I collapse the underlying natality microdata to calculate the low birth weight percentage for mother's of age a, giving birth in year p, themselves born in year c, and living in county l when the birth occurs, which I denote as $Y_{apc,l}$. I then merge the natality data with county characteristics from the County Data Book.

My first approach is to estimate models of the following form by weighted least squares, with the number of births in each cell as weights:

$$Y_{apc,l} = \beta \cdot c + \mathbb{1}_{c \geq 1948} \cdot \delta \cdot (c - 1948) + f^p(a) + X_l \cdot [\beta_x \cdot c + \mathbb{1}_{c \geq 1948} \cdot \delta_x \cdot (c - 1948) + g^p(a)] + \epsilon_{apc} \tag{5}$$

X_l is a characteristic of county l taken from the County Data Book, such as population density or median income. For all such characteristics I first take logs and then normalize as a Z-score (using the actual weights of regression sample). Both age-by-year control functions $f^p(a)$ and $g^p(a)$ are again specified as year fixed effects, age fixed effects, and a separate quadratic in age in each year. The specification therefore allows the external impact of age and year to be heteroge-

neous with respect to the county characteristic. Note that I fix the cohort break to occur at 1948 (the estimated break location in the full sample). The coefficient of interest δ_x estimates how the magnitude of the trend break in cohort effects differs with a 1 standard deviation change in the characteristic X_l.

The results in Panel A of Figure 10 show a remarkable lack of heterogeneity in the size of the cohort decline across counties with differing characteristics. The topmost point in blue shows the original estimate of the main cohort break without any interactions. The remaining points show estimates of the interaction term $\hat{\delta}_x$, for the listed county characteristic, each from a separate regression.[48] Interactions with population density in 1960, the share of the population with a college degree in 1970, income per capita in 1970, the share living in an urban area (as defined by the Census Bureau) in 1970, and the employment shares in manufacturing, construction, agriculture, and mining in 1960 are *all* small in magnitude and not significantly different from zero.

I use a second approach to test for more complex (eg. non-linear) heterogeneity in the cohort decline across counties with different characteristics. I partition the counties in the data into quartiles of a given county characteristic.[49] Separately for each quartile I then estimate cohort trend break models with unknown break location as in Equation 4.

Panels B and C of Figure 10 show the results for partitions based on population density, income per capita, and manufacturing share. The estimated location of the cohort decline is remarkably stable, with all of the point estimates within one year of 1948, and all confidence intervals including 1948. The most striking finding is that the estimated size of the cohort decline is significantly larger for counties in the top quartile of population density. The estimated size of the cohort break is between .15 and .21 for the bottom three quartiles of population density, but it jumps to .34 for counties in the top quartile of popultion density.

This finding that the cohort decline appears to be largest in the densest counties could point towards exposure to some kind of pollution which is worse in the densest urban areas as the cause. Of course the caveat should be kept in mind that these estimates are based on mother's county of residence, rather than their own county of birth, and hence could be biased by selective migration.

C. Evidence against some plausible causes

In this section, I briefly present evidence that appears inconsistent with a few ex ante plausible hypotheses for the cause of the cohort decline in health and human capital.

Readily observable differences in the family background of these cohorts or where they grew up do not appear the cause of their decline in health and human capital. Appendix Figures A5, A6,

[48]The estimated main cohort break $\hat{\delta}$ varies slightly across each specification, because it represents the magnitude of the cohort break for counties at the mean value of X_l. In practice these variations are small, so for simplicity I only report $\hat{\delta}$ from the uninteracted regression.

[49]The quartiles are estimated using the regression sample and number of births as weights.

and A7 show patterns by birth cohort for a range of such characteristics from the General Social Survey. They do not show evidence of trend breaks near 1947 in directions likely to negatively impact health. The share of parent's with a high school education, as well as measure of father's occupational prestige all increase smoothly and approximately linearly from 1930 to 1970 cohorts. Cross-cohort increases in the share of each cohort whose mothers and fathers had a Bachelor's degree actually *accelerated* after the late 1940s cohorts. The patterns by cohort in mother's employment, growing up in traditional two-parent households, and parent's nativity also suggest they are unlikely causes of the health and human capital decline.[50]

A single-index selection-based explanation — in which technology developed after 1947 suddenly allowed more unhealthy infants to survive to adulthood — also appears unlikely. Appendix Figure A9 shows that rapid improvements in the infant mortality rate beginning in the 1930s suddenly *slowed* beginning in the late-1940s. Interestingly there is also a sudden relative increase in comparison to other high come countries. Similarly, Panel C of Figure A10 shows that increases in the the share of each cohort surviving to age 18 began to decelerate after the late-1940s.

The decline also does not appear likely to be caused by the population size of the relevant cohorts. There is a long history of studying the potential negative impact of the large size of the baby-boom cohort on their wages (Welch, 1979; Freeman, 1979). Easterlin (1987) described numerous ways in which being a member of a large cohort could have negative effects through "cohort crowding" (Bound and Turner, 2007). However, the trend in cohort size, shown in Panels A and B of Figure A10, is quite different than the sharp trend break pattern seen in health and human capital. Cohort size increased sharply in 1946 and 1947, but for these cohorts health and human capital was still improving. Cohort size then stagnated for a few cohorts before increasing gradually until the mid-1950s cohorts. It then declined rapidly after the 1961 cohort. To match the observed patterns in health and human capital, therefore a theory would have to posit a complex lagged effect of cohort size on health and human capital.

My prior paper (Reynolds, 2023) which focused only on mortality of white Americans, presented a number of results which are relevant here. First, I showed that growth in smoking rates by women of childbearing age, estimated from Holford et al. (2014), slowed after 1945 (reproduced in Figure A11). This makes maternal smoking appear an unlikely driver of the cohort health decline.[51]. Second, I showed trends in eight important air pollutants from the Community Emissions Data System (O'Rourke et al.) (reproduced in Figure A12). None of them show evidence of trend breaks leading to more rapid increases after the late-1940s — making them appear unlikely causes

[50]The share of individuals who lived with both their mother and father at 16 does appear to decline rapidly after the mid-1950s cohorts. Estimating structural break models of unknown location on this outcome however leads to an estimated trend break at 1954, with a 99 percent confidence interval for the trend break location that does not include 1947.

[51]There does not appear to be survey evidence on maternal smoking specifically for this period

of the cohort decline. Use of the pesticide DDT grew rapidly after it was made publicly available in 1945, however its use then began to decline rapidly after 1959 (US Environmental Protection Agency, 1975). Also, as described in Reynolds (2023) the evidence on the human health effects of DDT exposure is still mixed (Turusov et al., 2002; Beard et al., 2006; Eskenazi et al., 2009), in contrast to that for lead exposure (discussed below). Finally, while atmospheric nuclear testing also began in 1945 and continued until 1963; fallout was especially severe in certain geographic areas, although it covered the entire US (including particularly a few southwestern states downwind of the Nevada Test Site (Meyers, 2019)). The health decline is not particularly large in these states.

D. Fetal lead exposure from motor vehicle gasoline

A number of facts point towards fetal lead exposure from motor vehicle gasoline as a possibly important driver of the cohort health decline worthy of further study.[52]

Motor vehicle production and fuel use were both restricted during World War II and began to rapidly increase soon after. Figure A13 shows time series of i) motor vehicle registrations and ii) fuel usage by motor vehicles, for 1930 to 1970 from Historical Statistics of the United States (US Census Bureau, 1975). Both series exhibit slow growth between 1930 and 1941, and declines during World War II. They then exhibit sharp trend breaks near 1945 and very rapid grwoth until 1970.

Exposure to pollution from lead gasoline additives also likely increased substantially over this period. The lead content of automotive gasoline decreased during WWII because the lead additive was needed for military planes (Oudijk, 2010). After this period, the concentration of lead in gasoline fluctuated but was generally increasing. With the rapid increase in total fuel consumption, the total tons of lead added to gasoline consumed in the US increased very rapidly.[53] When lead additives were phased out of gasoline beginning in the 1970s blood lead levels of children fell rapidly.

Early life lead exposure is thought to have broad and lasting negative health effects, impacting for example the development of multiple organ systems, cognitive ability and emotional regulation, and cardiovascular disease.[54] Lead exposure has also been linked to delayed puberty (Selevan et al., 2003; Wu et al., 2003).

[52] In my earlier paper (Reynolds, 2023) I suggested fetal lead exposure as a possible driver of cohort specific increases in the mortality of white Americans in particular. The discussion below is adapted from that paper.

[53] See Shelton et al. (1982) for estimates of the lead content of regular and premium gasoline. The series generally cited by economists on total lead consumed in gasoline lead additives, United States Bureau of Mines (1941-1970), eg. in Reyes (2007); Curci and Masera (2018), actually measures the volume of lead consumed in domestic *production* of these additives — so will include production for export and does not include imports. A series shown in Mielke et al. (2010) based domestic consumption also shows rapid increases from around 25,000 to more than 200,000 metric tons between 1945 and 1970.

[54] See for example McMichael et al. (1986); Needleman (2004); McFarland et al. (2022); Hollingsworth et al. (2022); Aizer et al. (2018) on the health and cognitive effects of fetal and childhood lead exposure.

When the cohorts whose health declined were children or in utero high blood lead content was ubiquitous across the US. McFarland et al. (2022) estimate that the share of children with blood lead levels above the 2015 threshold for "clinical concern" was 50 percent for the 1940-45 cohorts and reached 100 percent by the 1966-75 cohorts — meaning every single child's blood lead content was above this threshold. Nationally representative, laboratory-based blood lead estimates are first available for 1976-1980. In this perioed, the average blood lead content of 1-5 year olds was more than 3 times the above threshold for all of the following demographic groups: whites and blacks, children from high and low income families, residents of large and small MSAs, and residents of all four Census Regions (Egan et al., 2021). While lead exposure was widespread it appears to have been highest in the densest urban areas (Egan et al., 2021). As shown above the cohort health decline appears to also have been worst in the densest areas.

The facts above could point towards fetal lead exposure in particular as a possible driver of the early life health and human capital decline. However, the timing appears off by one year: the timing of the trend break in automobiles — of 1945 — would lead to a trend break in fetal exposure to automobile-based lead pollution at approximately the 1946 cohort. There is evidence that lead is stored in bones and remobilization of this lead during pregnancy and lactation is a potentially important source of toxins for fetuses and infants (Silbergeld, 1991; Gomaa et al., 2002). The lag could be due to time for the lead to be stored in women's bones, though this is conjecture. The fact that lead is stored in bones and remobilized during pregnancy could explain why there is not a sharp improvement in outcomes following the phaseout of leaded gasoline. Note that if *childhood*, rather than fetal, lead exposure has particularly large effects it would generate a trend break at an earlier cohort. For example, if all ages under 10 represented a critical period for lead exposure we would see health declines beginning with around the 1935 cohort, who would have been around 10-years old in 1935. Much more research is needed to establish or falsify the importance of fetal lead exposure for the documented health decline.

11. Conclusion

In this paper I present evidence of a precisely timed and broad decline in health and human capital across cohorts of Americans beginning with those born in 1947 and continuing until at least those born in the mid-1960s. This decline appears to have originated in adolescence or earlier and to have played a key role in: test score and education declines in the 1960s and 1970s, increases in the low birth weight rate beginning in the mid-1980s, wage stagnation since the 1970s, and recent mortality increases, particularly of white Americans. The cohort decline has even had an intergenerational effect through its effect on the health of mothers and their infants.

Something appears to have gone deeply wrong early in life for cohorts of Americans born after 1947. Future research should continue the search for the decline's underlying cause. There is

surprisingly little variation across states or racial groups in the size of the decline. It will therefore be important to compile data which allows for examination across additional dimensions, such as county-of-birth, parental characteristics, birth order, or family size.

The documented health and human capital decline may also be evident in other outcomes not examined in this paper. Secular changes over time in outcomes such as labor force participation or medical spending may be the result, at least in part, of cohort differences in health and human capital rather than period-specific factors. Further, the decline in health and human capital may have interacted in important ways with period-specific shocks — such as technological change, trade shocks, and increased supply of opioids. A full accounting of the broad and lasting impact of the cohort decline will require additional research.

Table 1: Evidence of relative cohort decline — cohort trend break models with unknown break location

controlling for year FEs, age FEs, and separate quadratic-in-age in each year

	Trend break in cohort effects		
	Size	Location	Existence
	δ	λ	*p-value*
Intergenerational infant health			
Low birth weight (%)	0.241	1948	$< .001$
	(0.014)	[1948, 1948]	
Mean birth weight (g)	-6.35	1948	$< .001$
	(0.35)	$[1948, 1948]$	
Labor market			
Median log wage, men	-0.016	1947	$< .001$
	(0.001)	[1946, 1947]	
Log mortality			
Men	0.028	1947	$< .001$
	(.001)	[1947, 1947]	
Women	0.029	1949	$< .001$
	(.001)	[1949, 1949]	
Years of schooling			
only controlling for year FEs, age FEs			
Men	-0.106	1948	$< .001$
	(.001)	[1948, 1948]	
Women	-0.074	1950	$< .001$
	(.001)	[1950, 1950]	

Each row shows the results of estimation of a model based on Equation 4, with the listed outcome as a dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (1999, 2000). All models except those for years of schooling control for year fixed effects, age fixed effects, and a separate quadratic-in-age in each year. The years of schooling models only control for age FEs. The column titled "Size" reports the estimated size of the cohort break δ, with the standard error in parentheses. The column titled "Location " reports the cohort at which the trend break is estimated to occur, with a 99 % confidence interval in brackets calculated by inverting a likelihood ratio statistic. The column titled "Existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear. Intergenerational infant health results are based on vital statistics natality microdata, 1968-1995, mothers age 18-40 who were born between 1930 to 1970. Labor market results are based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1965. Log mortality results are based on data from the Human Mortality Database, and include the years 1975-2019, ages 25-85, cohorts born from 1930 to 1965. Years of schooling results are based on CPS-MORG data, 1979-2016, and includes men and women aged 25-75, who were born from 1930 to 1965.

Table 2: Cohort decline concentrated among those born in the US — piecewise linear cohort effect models of intergenerational infant birth weight

robustness to varying age-by-year control function

	(1)	(2)	(3)	(4)
Panel A: Mothers born in US				
Size	-5.209	-7.292	-5.879	-7.059
	(0.167)	(0.438)	(0.650)	(0.839)
Location	1949	1948	1947	1947
	[1949, 1949]	[1947, 1948]	[1947, 1947]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel B: Mothers born outside US				
Size	1.888	-1.669	-3.213	6.565
	(0.351)	(0.815)	(1.194)	(1.591)
Location	1943	1947	1947	1951
	[1941, 1947]	[1938, 1962]	[1938, 1962]	[1951, 1951]
P-value for existence	$< .001$	.139	.096	.041
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Quadratic-age-by-year	No	Yes	No	No
Cubic-age-by-year	No	No	Yes	No
Quartic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on Equation 4, with the listed outcome in single age-by-year bins as the dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). The row titled "Size" reports the size of the trend break of cohort effects, δ, with the standard error in parentheses. The row titled "Location" reports the estimated cohort at which a trend break occurs, with a 99 $\%$ confidence region in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear. Results are based on vital statistics natality microdata, 1970-1995, mothers age 18-40 who were born between 1930 to 1970.

Table 3: Evidence of cohort decline across racial groups — piecewise linear cohort effect models

controlling for age FEs and year FEs

		(1) White	(2) Black	(3) Hispanic	(4) Other races
Median log wage	*Size*	-0.0130	-0.0251	-0.0133	-0.0143
		(0.0005)	(0.0018)	(0.0023)	(0.0036)
	Location	1946	1948	1949	1949
		[1945, 1947]	[1947, 1949]	[1946, 1953]	[1936, 1953]
Low birth weight (%)	*Size*	0.10	0.08		0.07
		(0.005)	(0.01)		(0.03)
	Location	1949	1947		1950
		[1948, 1949]	[1945, 1949]		[1938, 1957]
Log mortality					
Men	*Size*	0.0235	-0.0294		.0103
		(0.0003)	(0.0023)		(0.0012)
	Location	1944	1956		1945
		[1944, 1945]	[1954, 1956]		[1943, 1947]
Women	*Size*	0.0218	0.0125		0.0131
		(0.0005)	(0.0007)		(0.0024)
	Location	1950	1945		1953
		[1950, 1950]	[1944, 1947]		[1948, 1956]

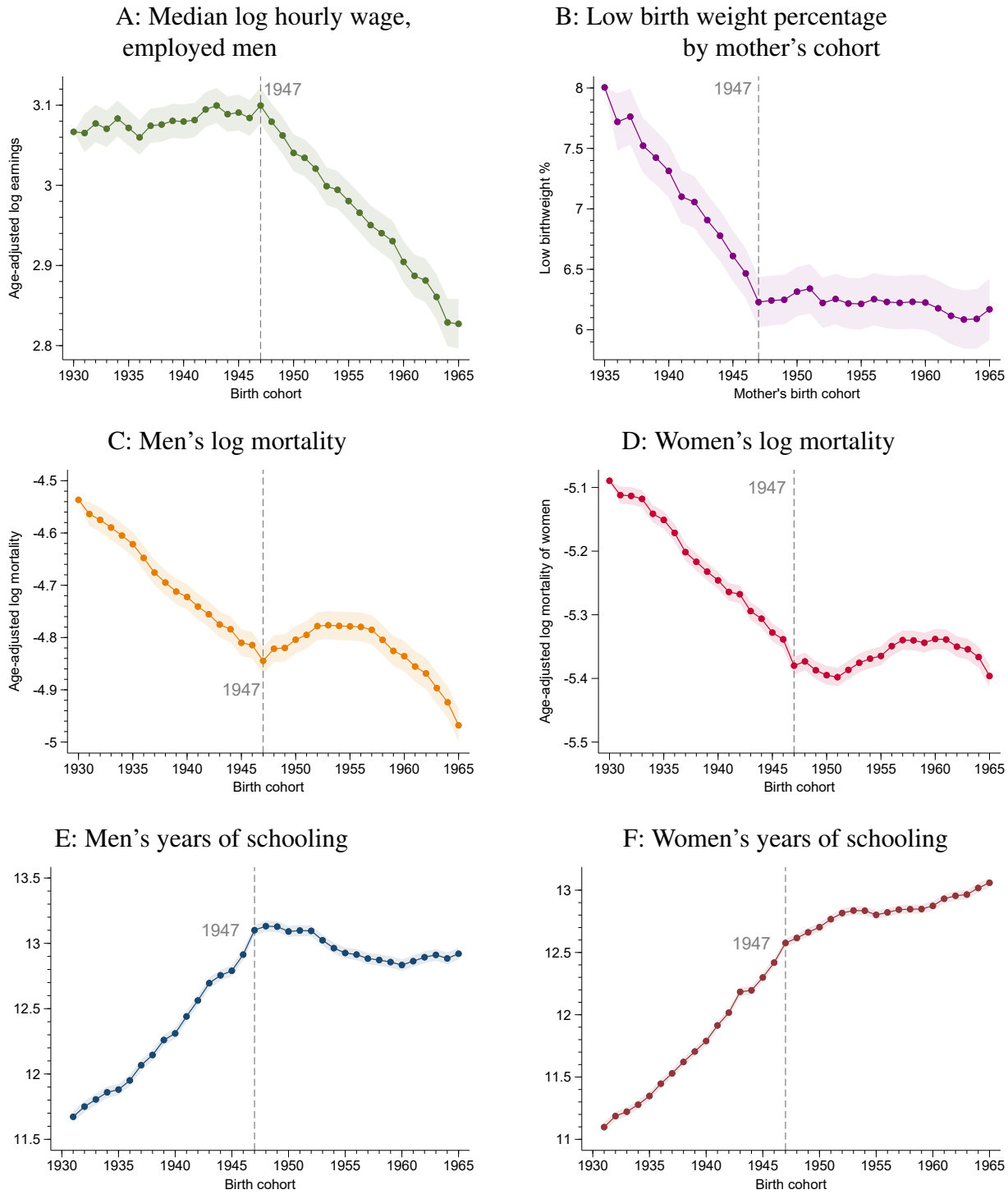
Each column shows the results of estimation of a model based on Equation 4, with the listed outcome as a dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). Intergenerational infant health results are based on vital statistics natality microdata, 1968-1995, mothers age 18-40 who were born between 1930 to 1960. Labor market results are based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1960. Log mortality results are based on data from the Human Mortality Database, and include the years 1975-2019, ages 25-85, cohorts born from 1930 to 1960. The column titled "Size" reports the estimated size of the cohort break δ, with the standard error in parentheses. The column titled "Location " reports the cohort at which the trend break is estimated to occur, with a 99 $\%$ confidence interval in brackets calculated by inverting a likelihood ratio statistic. The row titled "Existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear.

Table 4: Evidence of cohort decline across Census Regions — piecewise linear cohort effect models

		(1) Northeast	(2) Midwest	(3) South	(4) West
Median log wage	*Size*	-0.0170	-0.0194	-0.0139	-0.0149
		(0.0009)	(0.0013)	(0.0010)	(0.0012)
	Location	1947	1948	1946	1947
		[1946, 1947]	[1947, 1949]	[1945, 1947]	[1946, 1948]
Low birth weight (%)	*Size*	0.12	0.11	0.13	0.09
		(.01)	(.01)	(.01)	(.01)
	Location	1947	1949	1948	1948
		[1946, 1947]	[1948, 1950]	[1947, 1948]	[1946, 1949]
Log mortality					
Men	*Size*	0.0284	0.0287	0.0341	0.0286
		(0.0016)	(0.0014)	(0.0012)	(0.0015)
	Location	1946	1946	1946	1946
		[1946, 1946]	[1946, 1946]	[1946, 1946]	[1946, 1946]
Women	*Size*	0.0320	0.0288	0.0331	0.0269
		(0.0019)	(0.0019)	(.0016)	(0.0028)
	Location	1948	1950	1950	1949
		[1948, 1949]	[1950, 1950]	[1949, 1950]	[1949, 1949]

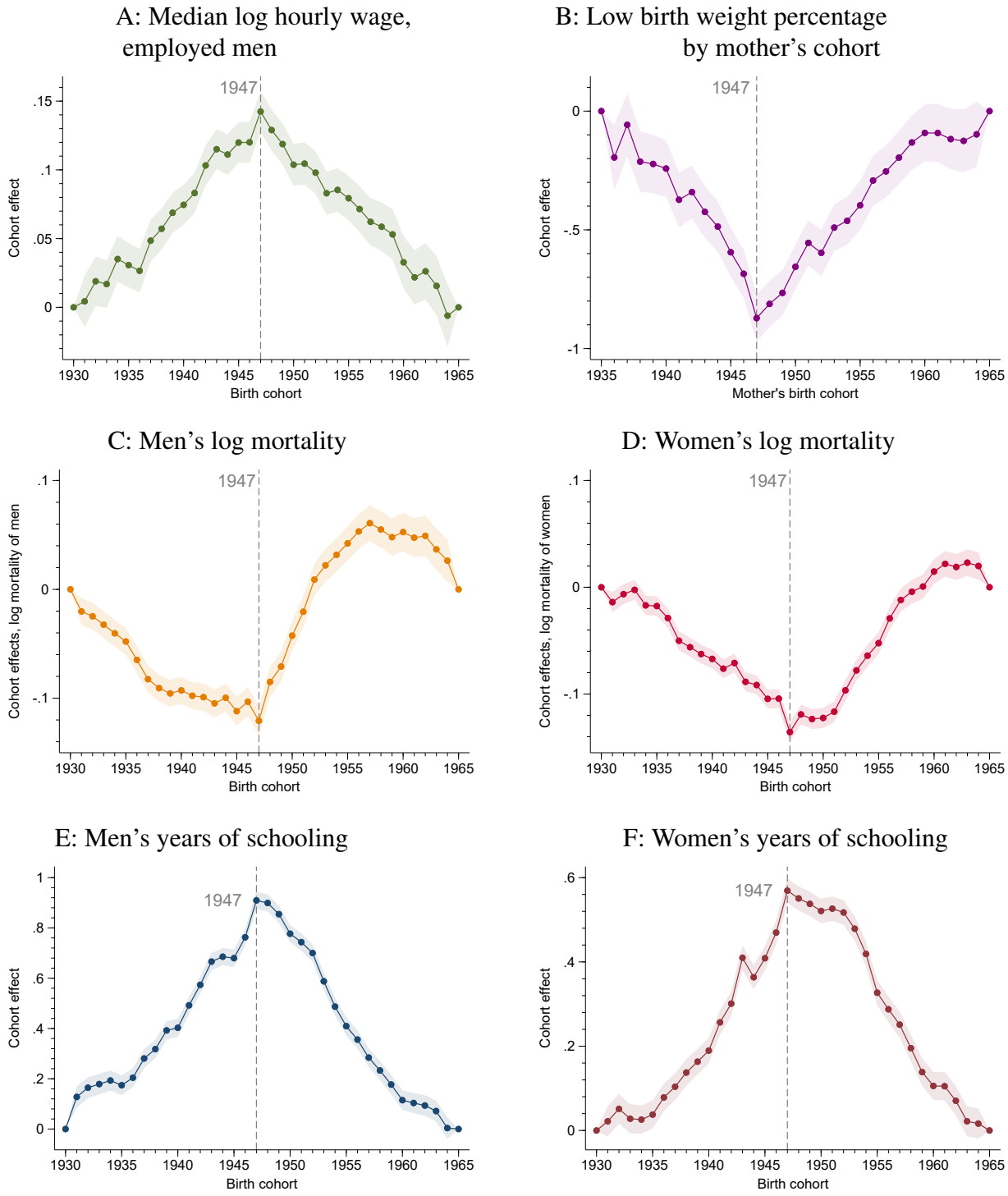
Each column shows the results of estimation of a model based on Equation 4, with the listed outcome as a dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). Wages and birth weight include controls for age fixed effects and year fixed effects — mortality also includes a separate quadratic-in-age in each year. The column titled "Size" reports the estimated size of the cohort break δ, with the standard error in parentheses. The column titled "Location " reports the cohort at which the trend break is estimated to occur, with a 99 $\%$ confidence interval in brackets calculated by inverting a likelihood ratio statistic. The row titled "Existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear. Intergenerational infant health results are based on vital statistics natality microdata, 1968-1995, mothers age 18-40 who were born between 1930 to 1970. Labor market results are based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1965. Log mortality results are based on data from the Human Mortality Database, and include the years 1975-2019, ages 25-85, cohorts born from 1930 to 1965.

Figure 1: Age-adjusted outcomes by birth cohort



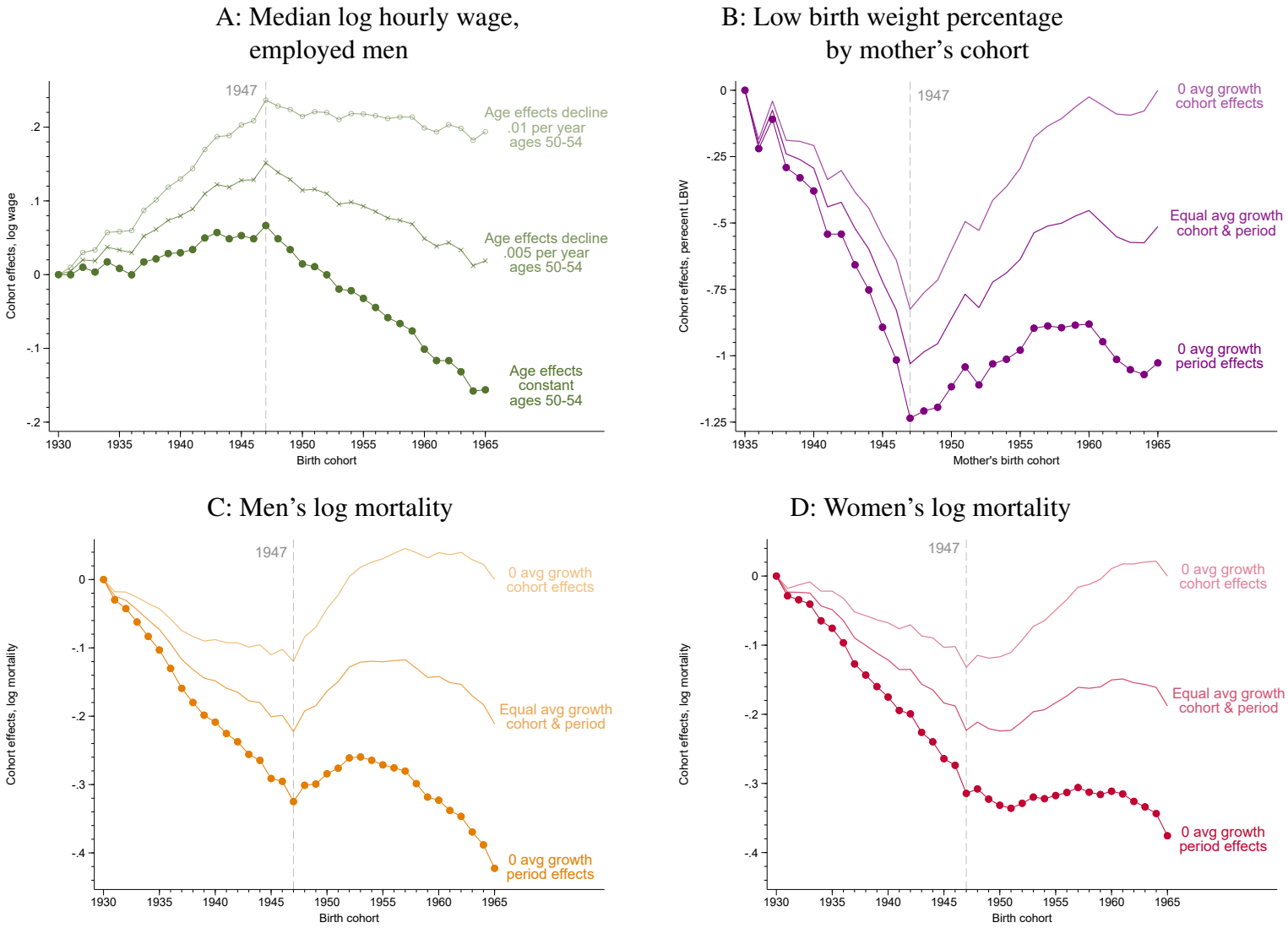
Each panel plots age-adjusted outcomes by birth cohort. Age adjustment is done by regressing the outcome on a full set of cohort fixed effects and age effects, and plotting predicted values. Panel A is based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1965. Panel B is based on vital statistics natality microdata, 1968-1990, mothers age 18-40, who were born from 1935-1965. Panel C and D are based on data from the Human Mortality Database, and include 1975-2019, ages 25-85, cohorts born from 1930 to 1965. Panels E and F are based on CPS-MORG data, 1979-2016, and includes men and women aged 25-75, who were born from 1930 to 1965.

Figure 2: Detrended cohort effects from additively separable age-period-cohort models



Each panel plots detrended cohort effects from estimation of age-period-cohort models based on Equation 3. Data sources and sample restrictions for each outcome are the same as in Figure 1.

Figure 3: Cohort effects under different assumptions, additively separable age-period-cohort models

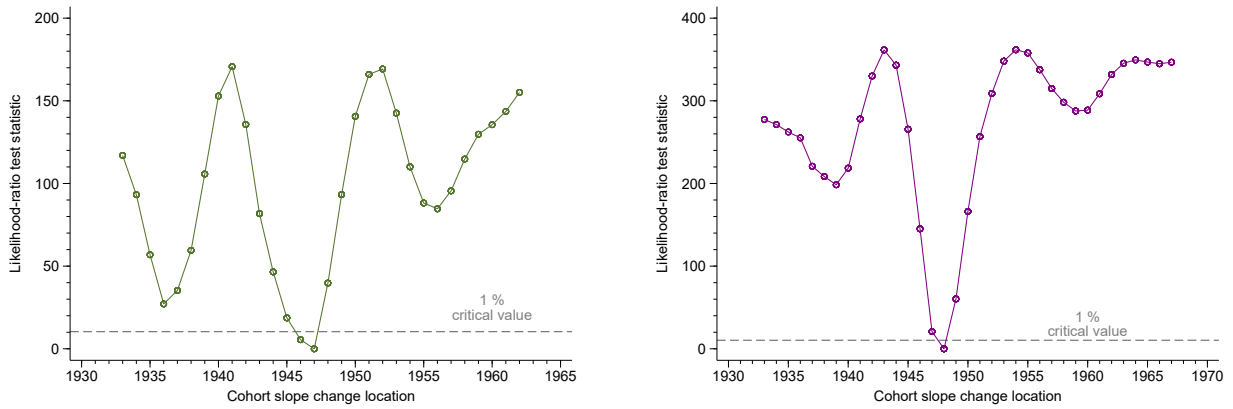


Each panel plots cohort effects from estimation of age-period-cohort models based on Equation 3. Each series shows cohort effects under different assumptions on the age effects, or the average growth in period/cohort effects. See Section 5b. Data sources and sample restrictions for each outcome are the same as in Figure 1.

Figure 4: Hansen-type estimation of cohort at which decline begins

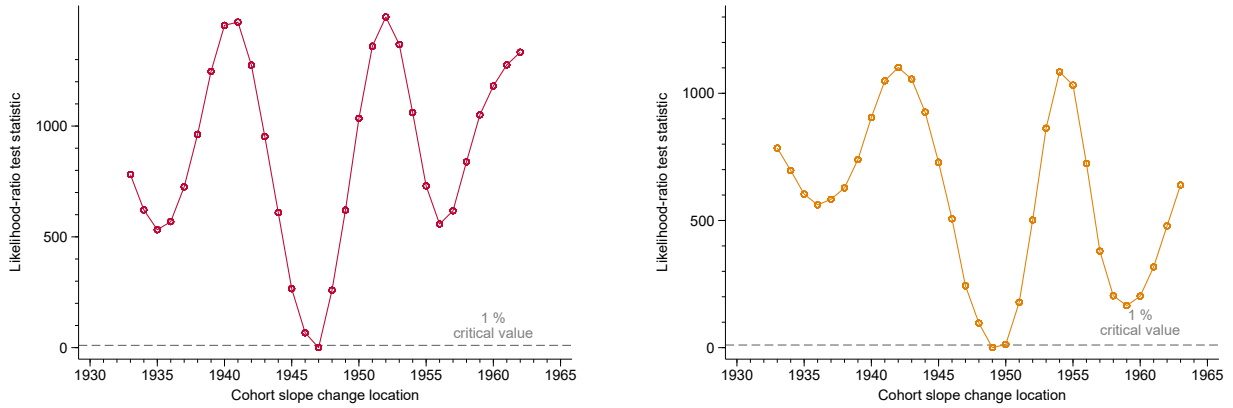
A: Median log hourly wage, B: Low birth weight percentage

controlling for age FEs, year FEs, and separate quadratic-in-age in each year



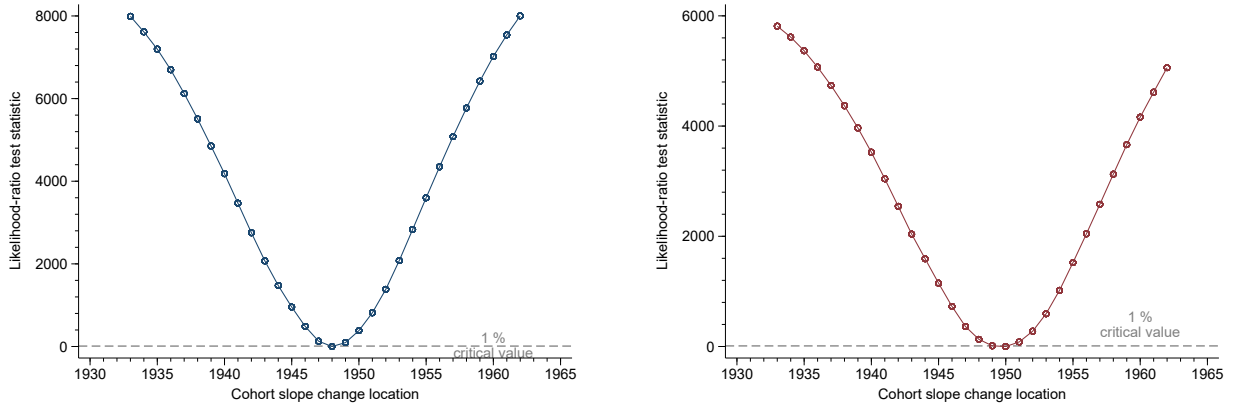
C: Men's log mortality D: Women's log mortality

controlling for age FEs, year FEs, and separate quadratic-in-age in each year



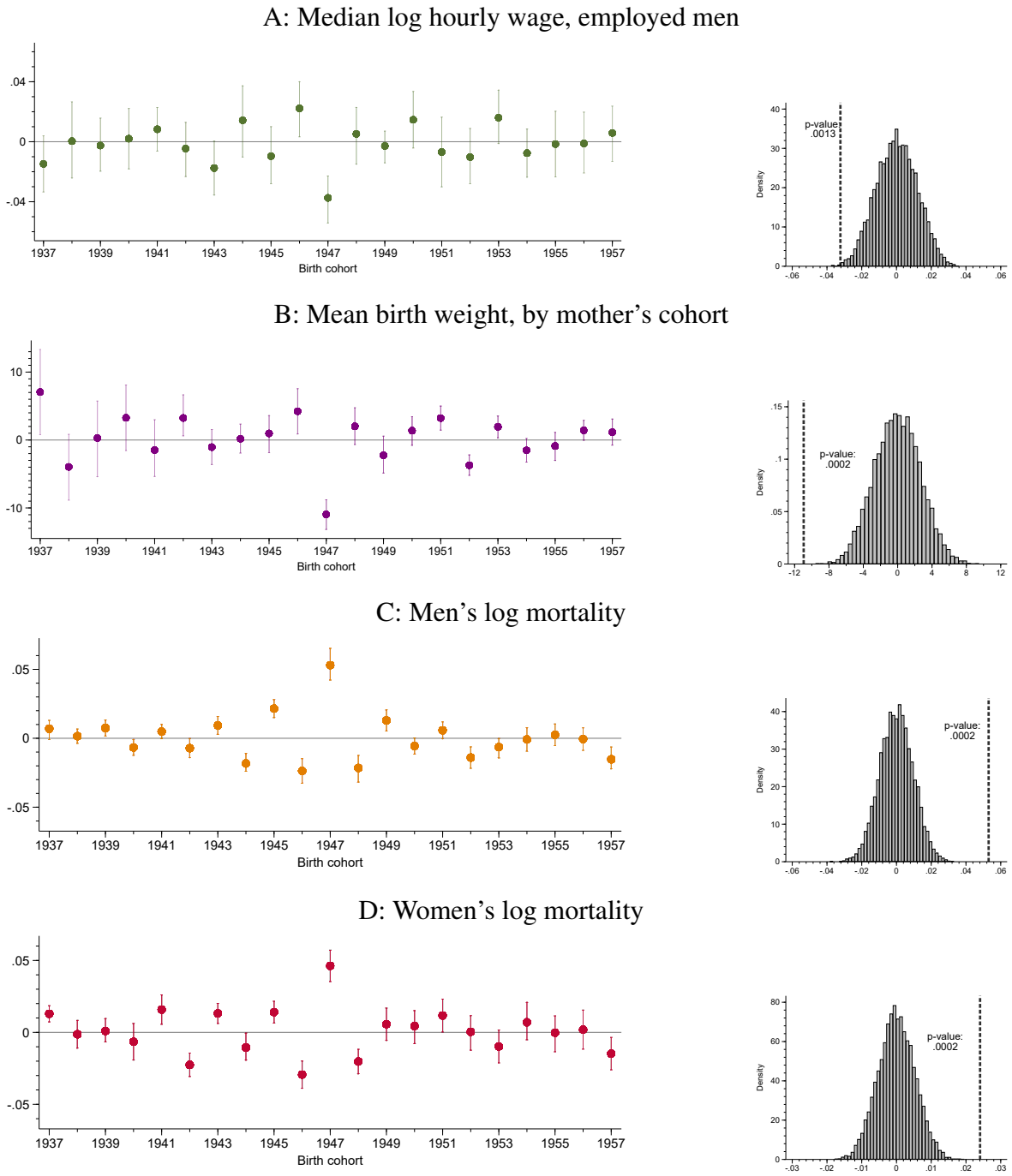
E: Men's years of schooling F: Women's years of schooling

controlling for age FEs



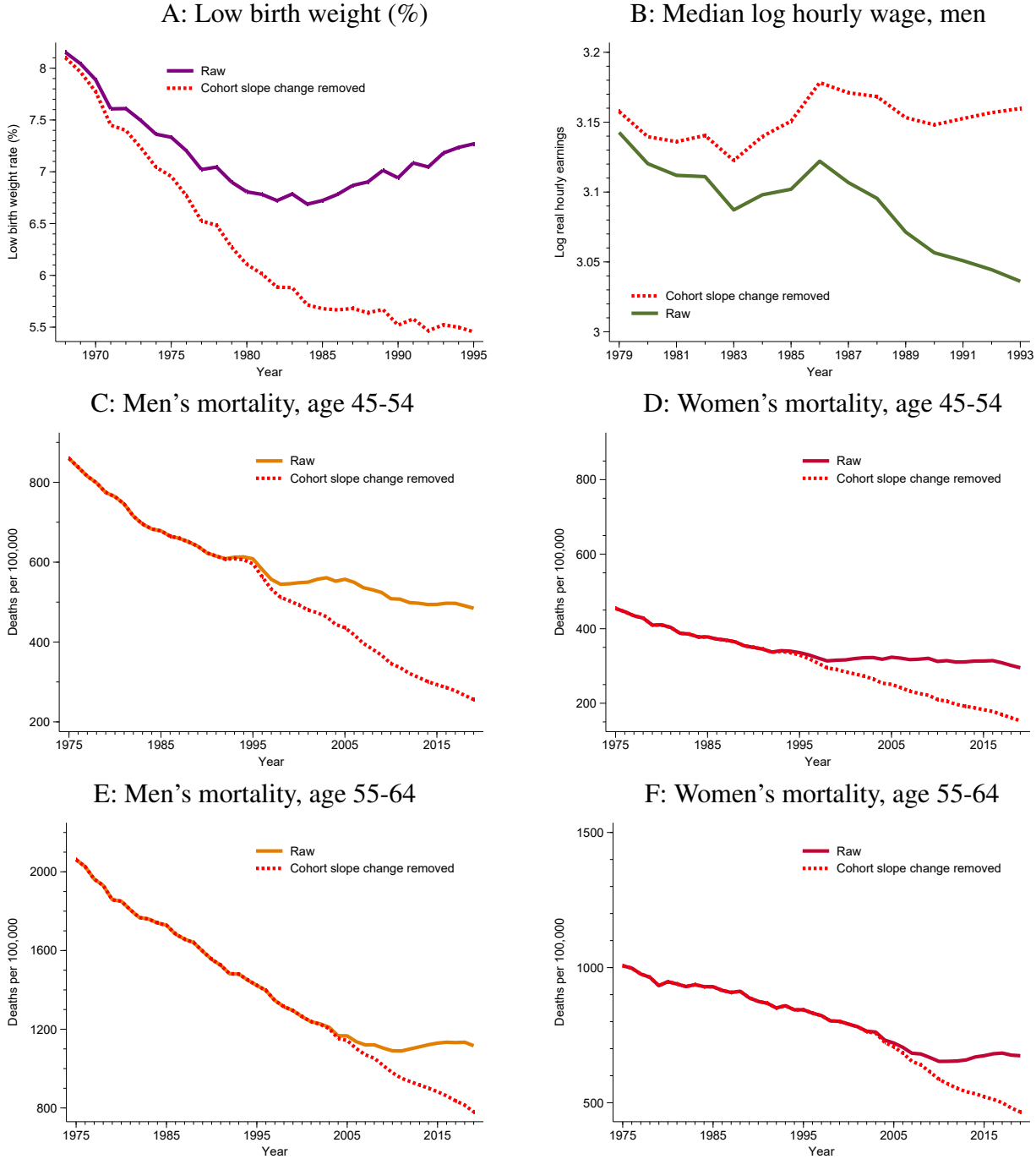
This figure shows estimation and inference of the location at which a cohort decline begins based on the model in Equation 4, with age fixed effects, year fixed effects, and separate quadratic-in-age in each year included as controls. Each plot shows the likelihood-ratio test statistic suggested in Hansen (1996), for a model with the listed outcomes as the dependent variable. The point estimate of this location is the cohort with the minimum value of the LR-test statistic. The 99 % confidence region is those cohorts with a LR-test statistic below the critical value shown with a dashed grey line. Data sources and sample restrictions for each outcome are the same as in Figure 1.

Figure 5: Nonparametric estimation of average second difference in local cohort effects



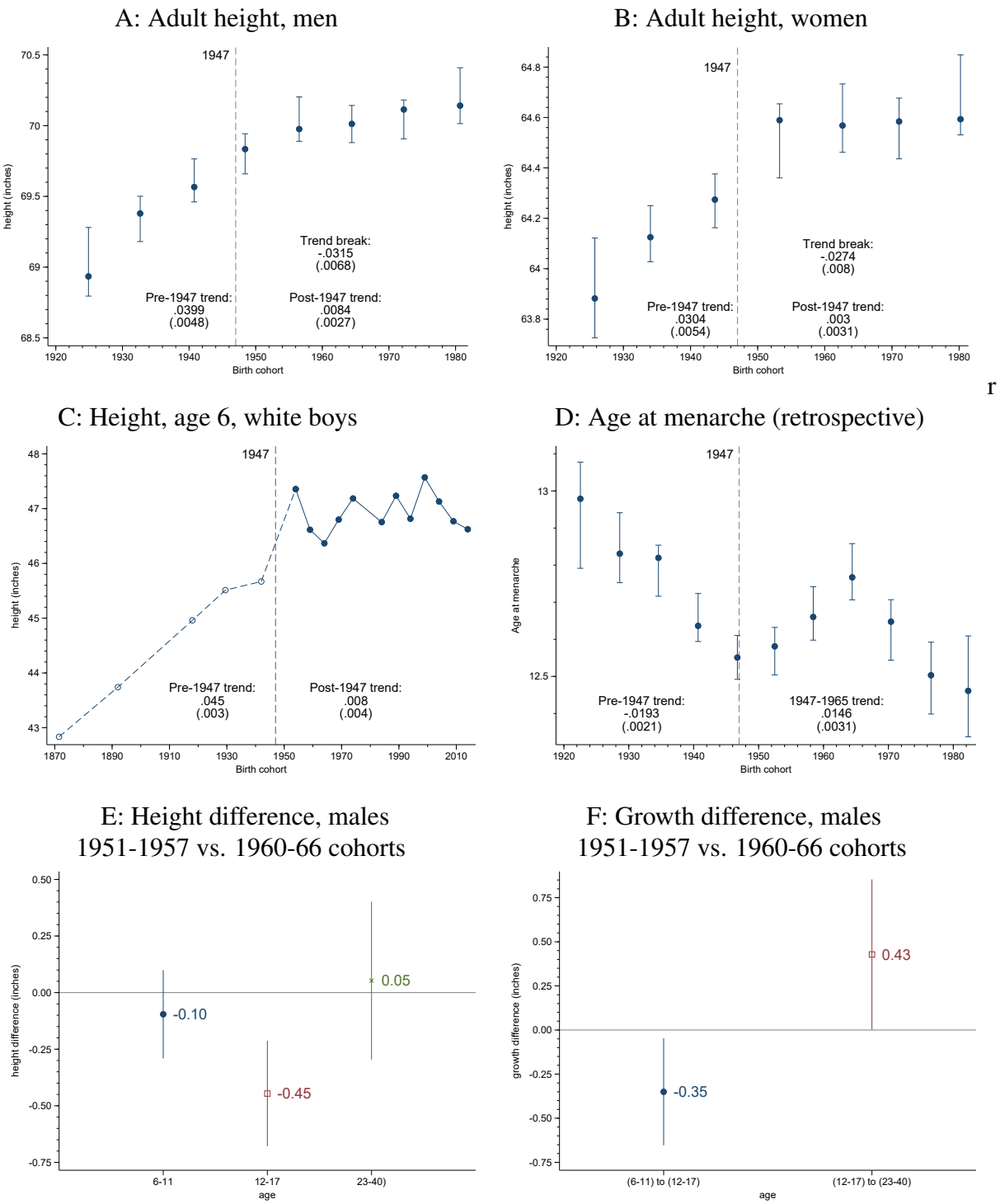
This figure shows the results of implementing the estimator of the average second difference and in local cohort effects, as well as the permutation-based inference, described in Section 7. For each outcome, the panels on the left show the sequence of point estimates $\hat{\Psi}_c$ for the cohorts between 1937 and 1957. They also show permutation-based, constant-effect, 95 % confidence intervals based on 1,000 simulations of the permutation distribution and a bisection algorithm. The confidence interval can be interpreted as the set of constant second-difference in local cohort effects which the test fails to reject. The panels on the right graphically illustrate the permutation-based test of the sharp null that all of the second difference in local cohort effects for the 1947 cohort are 0. The dashed line shows the point estimate of $\hat{\Psi}_{1947}$ and the histogram shows the permutation distribution based on 10,000 simulations. Data sources and sample restrictions for each outcome are the same as in Figure 1.

Figure 6: Impact of cohort decline on year-over-year declines



This figure shows simulated counterfactual year-over-year trends in 4 outcomes had the trend break in cohort effects not occurred, ie. if the pre-break cohort trend had continued. The trend break estimated based on the model in Equation 4 is subtracted from each observation. These transformed data are then used to calculate each of the listed outcomes, which are plotted as the dashed red line. The same outcome based on the untransformed, raw data is plotted as a solid line. Panel A shows the percent of infants born at low birth weight. Panel B shows the average across men age 25-54 of median wages for single age-bins. Panels C through F show mortality rates of men and women, age-adjusted assuming a uniform population distribution by single year of age. Data sources for each outcome are the same as for Figure 1.

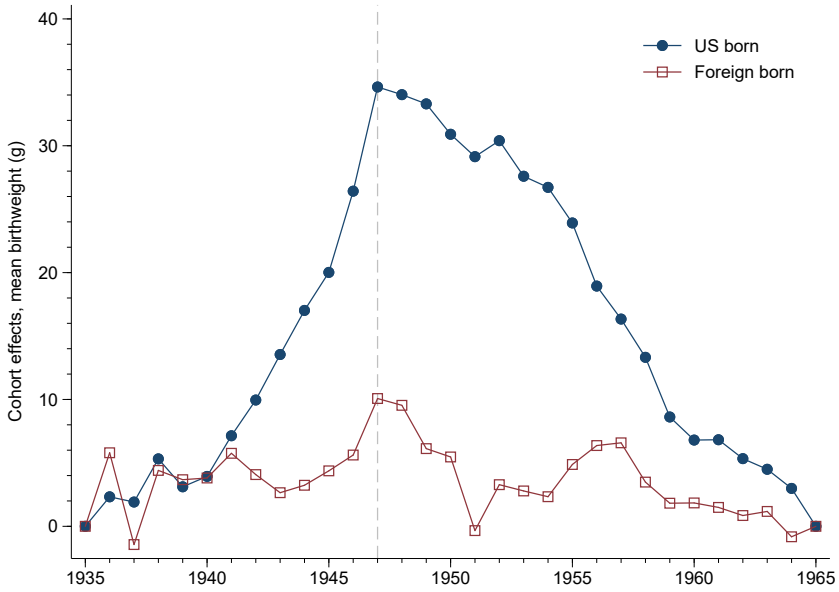
Figure 7: Growth and physical development stagnated (and may have worsened) for approximately same cohorts



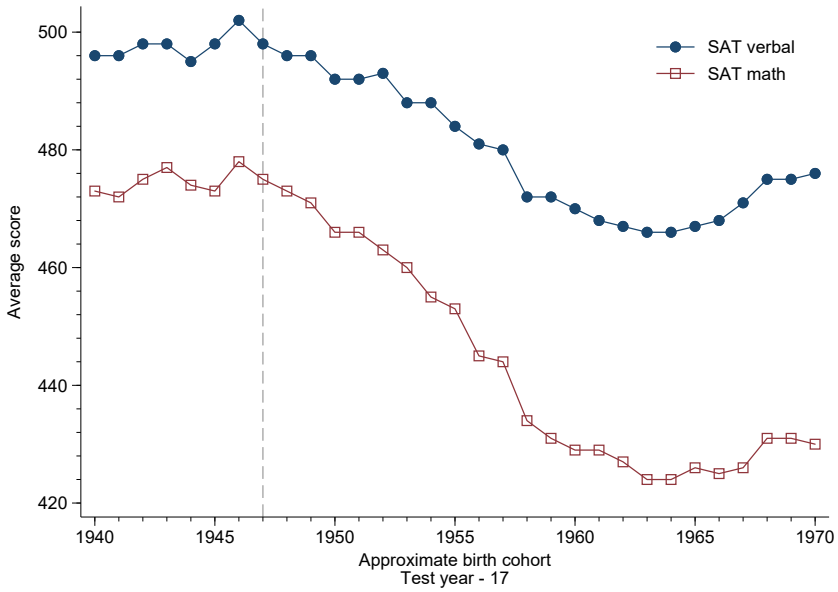
All panels use data from multiple rounds of NHES and NHANES surveys. Panels A and B show binscatter plots following Cattaneo et al. (2019) of heights of men and women ages 23 to 47. Panels C and D plot the height of white boys age 6 and 10 respectively. Pre-1950 data comes from non-representative historical studies compiled by Meredith (1964), with each point representing the mean for an interval of birth cohorts given in Meredith (1964), plotted at the midpoint. Post-1950 data comes from the NHES and NHANES surveys and shows the mean for 5-year birth cohorts. Panels E and F show estimates of the height difference of the listed cohorts at different ages and the difference in implied in implied growth between different ages, estimated from the regressions described Appendix B.

Figure 8: Additional evidence that decline originated early in life

A. Detrended cohort effects, intergenrational infant birth weight, by mother's place of birth



B. Scholastic Aptitude Test (SAT) scores

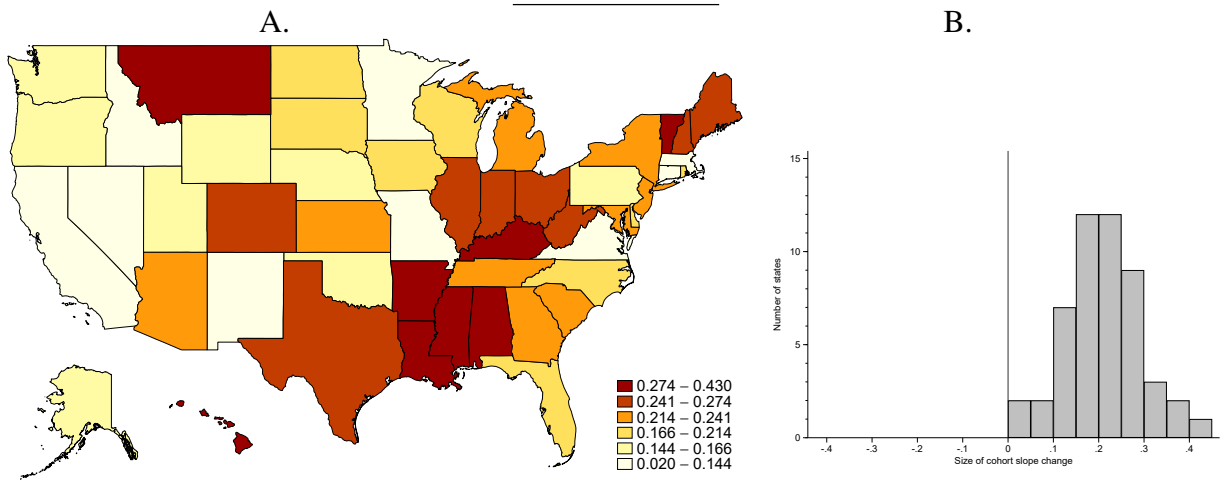


Panel A plots detrended cohort effects from estimation of age-period-cohort models based on Equation 1 — separately for mothers born in one of the 50 states or D.C., and those born outside of the US. Results are based vital statistics natality microdata, 1970-1990, including mothers age 18-40, who were born from 1935-1965.
Panel B shows verbal and mathematics scores on the Scholastic Aptitude Test. Approximate birth cohort is defined as the year the test was given minus 17. Data is from Harnischfeger and Wiley (1975).

Figure 9: Empirical Bayes estimates of geographic variation in size of maternal health decline

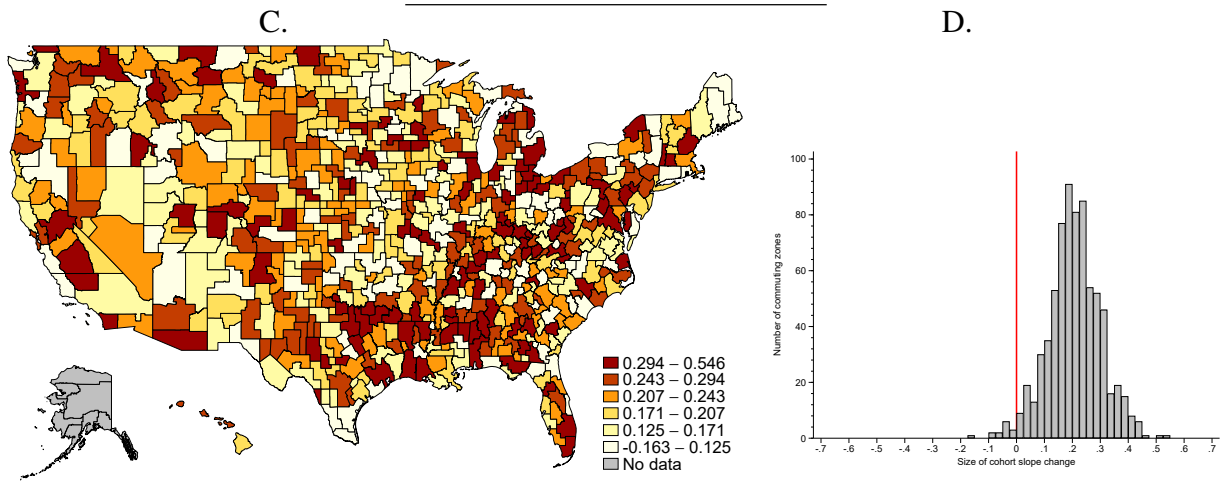
Low birth weight percentage by mother's birth cohort

By state of birth



Posterior estimates
Mean: .212 , Variance: .013

By commuting zone of residence

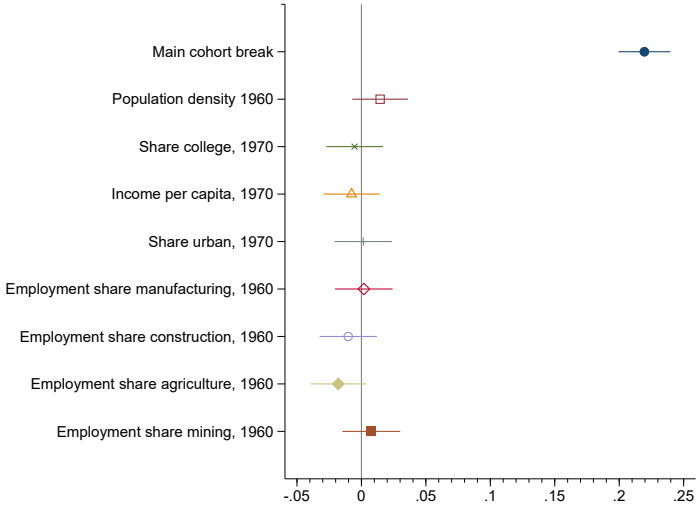


Posterior estimates
Mean: .207 , Variance: .051

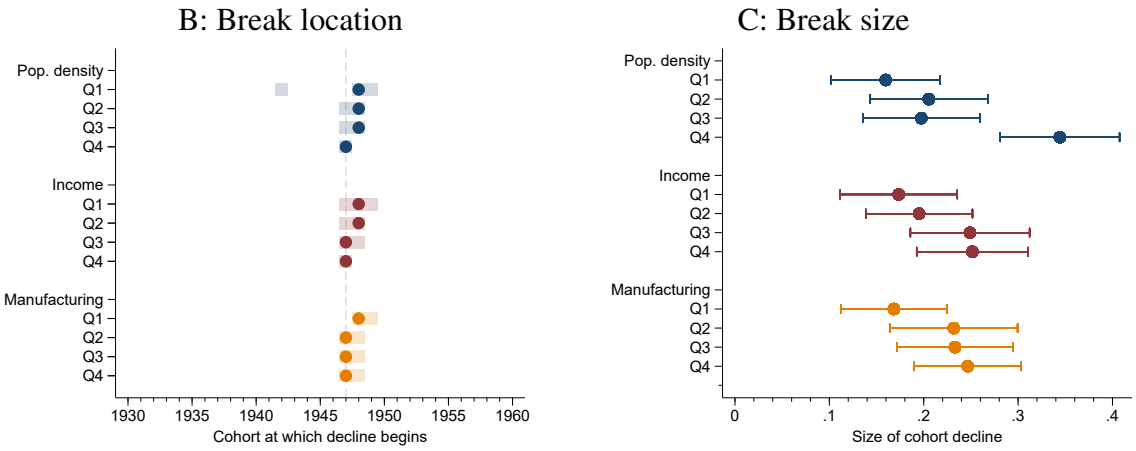
This figures shows estimates of the size of the cohort decline in maternal health by state-of-birth and commuting zone of residence based on a two-step empirical Bayes procedure. All panels are based on vital statistics natality microdata, including mothers age 18-40, who were born from 1935-1965. Panels A and B are based on 1970-1990, and Panels C and D are based on 1968-1988. I estimate the piecewise-linear model of Equation 4, with λ fixed at 1948, separately for each SOB/CZ. I use low birth weight percentage as the dependent variable and include as controls age fixed effects, year fixed effects, and a separate quadratic-in-age in each year. I treat the estimated means and variances by SOB/CZ of the size of the cohort decline, δ as priors. In the second step, I use the Gaussian Empirical Bayes procedure suggested in Morris (1983) to form posterior estimates of: i) the size of the decline by SOB/CZ, ii) the mean and variance in the size of the decline across *all* SOBs/CZs.

Figure 10: Most county characteristics not correlated with size of maternal health decline, but the decline is largest in the densest counties

A: Interaction between county characteristics and cohort break in low birth weight percentage by mother's year of birth



Cohort trend break models with unknown break location, separately by quartile of county characteristic



Panel A shows estimates of the size of the trend break in maternal health, and interactions of this trend break with characteristics of the county where the mother lives. I first take logs and then normalize as a Z-score. I then estimate a separate regression for each county characteristic of the form in Equation 5 with low birth weight percentage as the dependent variable, with a fixed cohort break to occur at 1948 (the estimated break location in the full sample), and interaction terms which interact the normalized county characteristic with the cohort trend break. I include control functions again specified as year fixed effects, age fixed effects, and a separate quadratic in age in each year plus all of these terms interacted with the county characteristic. The figures shows $\hat{\delta}_x$, which estimates how the magnitude of the trend break in cohort effects differs with a 1 standard deviation change in the characteristic.

Panels B and C of Figure 10 show results of cohort trend break models of unknown location, estimated separately for partitions of the data based on quartiles of the listed county characteristics. Separately for each partition, I then estimate cohort trend break models with unknown break location as in Equation 4.

All results are based on the Natality Detail Files, 1968-1988, mother's age 18-40 and mother's born 1935 to 1965. County characteristics are from the County Data Book.

.

References

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4:1043–1171, 2011.

Daron Acemoglu and David Autor. What does human capital do? A review of goldin and katz's "The race between education and technology". *Journal of Economic Literature*, 50(2):426–63, 2012.

Enrique Acosta, Alain Gagnon, Nadine Ouellette, Robert R Bourbeau, Marilia R Nepomuceno, and Alyson A van Raalte. The boomer penalty: Excess mortality among baby boomers in Canada and the United States. Technical report, Max Planck Institute for Demographic Research, Rostock, Germany, 2020.

Mark Aguiar and Erik Hurst. Deconstructing life cycle expenditure. *Journal of Political Economy*, 121(3):437–492, 2013.

Anna Aizer, Janet Currie, Peter Simon, and Patrick Vivier. Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics*, 10(1):307–41, 2018.

Douglas Almond and Kenneth Chay. The long-run and intergenerational impact of poor infant health: Evidence from cohorts born during the civil rights era. *University of California-Berkeley, mimeograph*, 2006.

David Autor and David Dorn. The growth of low-skill service jobs and the polarization of the US labor market. *The American Economic Review*, pages 1553–1597, 2013.

David Autor, Lawrence F. Katz, and Melissa S. Kearney. Rising wage inequality: The role of composition and prices. *NBER Working Paper*, page 11628, 2005.

David Autor, Lawrence F. Katz, and Melissa S. Kearney. Trends in US wage inequality: Revising the revisionists. *The Review of economics and statistics*, 90(2):300–323, 2008.

John Beard, Australian Rural Health Research Collaboration, et al. Ddt and human health. *Science of the total environment*, 355(1-3):78–89, 2006.

Andrew Bell and Kelvyn Jones. The impossibility of separating age, period and cohort effects. *Social Science & Medicine*, 93:163–165, 2013.

José M Bernardo. The concept of exchangeability and its applications. *Far East Journal of Mathematical Sciences*, 4:111–122, 1996.

John H Bishop. Is the test score decline responsible for the productivity growth decline? *The American Economic Review*, pages 178–197, 1989.

Dan A Black, Yu-Chieh Hsu, and Lowell J Taylor. The effect of early-life education on later-life mortality. *Journal of Health Economics*, 44:1–9, 2015.

Margherita Borella, Mariacristina De Nardi, and Fang Yang. The lost ones: the opportunities and outcomes of white, non-college-educated Americans born in the 1960s. *NBER Macroeconomics Annual*, 34(1):67–115, 2020.

John Bound and George Johnson. Changes in the structure of wages in the 1980's: An evaluation of alternative explanations. *American Economic Review*, 82(3):371–92, 1992.

John Bound and Sarah Turner. Cohort crowding: How resources affect collegiate attainment. *Journal of public Economics*, 91(5-6):877–899, 2007.

Audra J Bowlus and Chris Robinson. Human capital prices, productivity, and growth. *American Economic Review*, 102(7):3483–3515, 2012.

William A Brock and Steven N Durlauf. Growth empirics and reality. *World Bank Economic Review*, 15(2):229–272, 2001.

Kasey Buckles, Andreas Hagemann, Ofer Malamud, Melinda Morrill, and Abigail Wozniak. The effect of college education on mortality. *Journal of Health Economics*, 50:99–114, 2016.

David Card. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160, 2001.

David Card and Thomas Lemieux. Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2):705–746, 2001a.

David Card and Thomas Lemieux. Dropout and enrollment trends in the postwar period: What went wrong in the 1970s? In *Risky behavior among youths: An economic analysis*, pages 439–482. University of Chicago Press, 2001b.

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.

Anne Case and Angus Deaton. Mortality and morbidity in the 21st century. *Brookings Papers on Economic Activity*, 2017:397, 2017.

Matias D Cattaneo, Richard K Crump, Max H Farrell, and Yingjie Feng. On binscatter. *arXiv preprint arXiv:1902.09608*, 2019.

Centers for Disease Control. Increasing incidence of low birthweight – United States, 1981-1991. *Morbidity and Mortality Weekly Report*, 43(18):335–339, 1994.

Gary Chamberlain. Quantile regression, censoring, and the structure of wages. In *Advances in econometrics: sixth world congress*, volume 2, pages 171–209, 1994.

Amitabh Chandra, Amy Finkelstein, Adam Sacarny, and Chad Syverson. Health care exceptionalism? Performance and allocation in the US health care sector. *American Economic Review*, 106 (8):2110–44, 2016.

Louis Chauvel. Age-period-cohort detrended apc-d model. *Working paper*, 2011.

Kenneth Chay, Jonathan Guryan, and Bhashkar Mazumder. Birth cohort and the black-white achievement gap: The roles of access and health soon after birth. *National Bureau of Economic Research, Working Paper*, 2009.

Kenneth Chay, Jonathan Guryan, and Bhashkar Mazumder. Early life environment and racial inequality in education and earnings in the United States. *NBER Working Paper*, 2014.

Raj Chetty, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. The association between income and life expectancy in the United States, 2001-2014. *Journal of the American Medical Association*, 315(16):1750–1766, 2016.

Denis Chetverikov, Bradley Larsen, and Christopher Palmer. IV quantile regression for group-level treatments, with an application to the distributional effects of trade. *Econometrica*, 84(2): 809–833, 2016.

Damon Clark and Heather Royer. The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, 103(6):2087–2120, 2013.

T.J. Cole. The secular trend in human physical growth: a biological view. *Economics Human Biology*, 1(2):161–168, 2003.

Dora Costa and Richard H Steckel. Long-term trends in health, welfare, and economic growth in the United States. In *Health and welfare during industrialization*, pages 47–90. University of Chicago Press, 1997.

Dora L Costa. Health and the economy in the United States from 1750 to the present. *Journal of Economic Literature*, 53(3):503, 2015.

Flavio Cunha, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. Interpreting the evidence on life cycle skill formation. volume 1 of *Handbook of the Economics of Education*, pages 697–812. 2006.

Federico Curci and Federico Masera. Flight from urban blight: lead poisoning, crime and suburbanization. 2018.

Janet Currie and Douglas Almond. Human capital development before age five. In *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier, 2011.

Janet Currie and Jonathan Gruber. Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy*, 104(6):1263–1296, 1996.

Angus Deaton. *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications, 1997.

John DiNardo, Nicole M. Fortin, and Thomas Lemieux. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5):1001–1044, 1996.

David Draper, James S Hodges, Colin L Mallows, and Daryl Pregibon. Exchangeability and data analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 156(1):9–28, 1993.

Richard A Easterlin. *Birth and fortune: The impact of numbers on personal welfare*. University of Chicago Press, 1987.

Kathryn B Egan, Cheryl R Cornwell, Joseph G Courtney, and Adrienne S Ettinger. Blood lead levels in US children ages 1–11 years, 1976–2016. *Environmental health perspectives*, 129(3): 037003, 2021.

Irma T Elo and Samuel H Preston. Educational differentials in mortality: United States, 1979–1985. *Social science & medicine*, 42(1):47–57, 1996.

Michael D Ernst. Permutation methods: a basis for exact inference. *Statistical Science*, pages 676–685, 2004.

Brenda Eskenazi, Jonathan Chevrier, Lisa Goldman Rosas, Henry A Anderson, Maria S Bornman, Henk Bouwman, Aimin Chen, Barbara A Cohn, Christiaan De Jager, Diane S Henshel, et al. The pine river statement: human health consequences of ddt use. *Environmental Health Perspectives*, 117(9):1359–1367, 2009.

Roderick Floud, Robert W Fogel, Bernard Harris, and Sok Chul Hong. *The changing body: health, nutrition, and human development in the western world since 1700*. Cambridge University Press, 2011.

James R Flynn. The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological bulletin*, 95(1):29, 1984.

Robert W Fogel. Nutrition and the decline in mortality since 1700: some preliminary findings. In

Long-term factors in American economic growth, pages 439–556. University of Chicago Press, 1986.

Robert W Fogel. *Explaining long-term trends in health and longevity*. Cambridge University Press, 2012.

Robert W Fogel and Dora L Costa. A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. *Demography*, 34(1):49–66, 1997.

"Center for Economic and Policy Research". Cps org uniform extracts, version 2.5. 2020.

Ethan Fosse and Christopher Winship. Analyzing age-period-cohort data: A review and critique. *Annual Review of Sociology*, 45:467–492, 2019.

Richard B Freeman. The effect of demographic factors on age-earnings profiles, 1979.

Titus Galama, Adriana Lleras-Muney, and Hans van Kippersluis. The effect of education on health and mortality: A review of experimental and quasi-experimental evidence. In *Oxford Research Encyclopedia of Economics and Finance*. 2018.

Peter Ganong and Simon Jäger. A permutation test for the regression kink design. *Journal of the American Statistical Association*, 113(522):494–504, 2018.

Christina Gathmann, Hendrik Jürges, and Steffen Reinhold. Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science & Medicine*, 127:74–82, 2015.

Claudia Goldin. The quiet revolution that transformed women's employment, education, and family. *Richard T. Ely Lecture — American economic review*, 96(2):1–21, 2006.

Ahmed Gomaa, Howard Hu, David Bellinger, Joel Schwartz, Shirng-Wern Tsaih, Teresa Gonzalez-Cossio, Lourdes Schnaas, Karen Peterson, Antonio Aro, and Mauricio Hernandez-Avila. Maternal bone lead as an independent risk factor for fetal neurotoxicity: a prospective study. *Pediatrics*, 110(1):110–118, 2002.

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.

Elise Gould. Why America's workers need faster wage growth — and what we can do about it. *Economic Policy Institute Briefing Paper*, 382, 2014.

Michael Grossman and Theodore J. Joyce. Unobservables, pregnancy resolutions, and birth weight production functions in New York City. *Journal of Political Economy*, 98(5, Part 1):983–1007, 1990.

Fatih Guvenen, Greg Kaplan, Jae Song, and Justin Weidner. Lifetime earnings in the United States over six decades. *American Economic Journal: Applied Economics*, 14(4):446–79, 2022.

Robert E Hall. The measurement of quality change from vintage price data. In Zvi Girilches, editor, *Dynamic Demographic Analysis*, pages 240–271. Harvard University Press, 1968.

Peter V Hamill, Terence A Drizd, Clifford L Johnson, Robert B Reed, and Alex F Roche. NCHS growth curves for children birth-18 years. Technical report, Department of Health Education and Welfare Washington DC, 1977a.

Peter V Hamill, Johnston E Frances, and Stanley Lemeshow. Height and weight of children: Socioeconomic status. Technical report, Department of Health Education and Welfare Washington DC, 1977b.

Christopher Handy and Katharine L Shester. The baby boom and educational attainment. 2019.

Bruce E. Hansen. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, pages 413–430, 1996.

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.

Annegret Harnischfeger and David E. Wiley. *Achievement Test Score Decline: Do We Need to Worry?* CEMREL, Inc., 1975.

Jerry A. Hausman, Haoyang Liu, Ye Luo, and Christopher Palmer. Errors in the dependent variable of quantile regression models. *NBER Working Paper*, 2019.

RC Hauspie, Martine Vercauteren, and Charles Susanne. Secular changes in growth and maturation: an update. *Acta Paediatrica*, 86(S423):20–27, 1997.

James J Heckman. The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255, 2007.

James J Heckman. Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature*, 48(2):356–398, 2010.

James J. Heckman and Paul A. LaFontaine. The American high school graduation rate: Trends and levels. *The Review of Economics and Statistics*, 92(2):244–262, 2010.

James J Heckman, Lance Lochner, and Christopher Taber. Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics*, 1(1):1–58, 1998.

Theodore R Holford, David T Levy, Lisa A McKay, Lauren Clarke, Ben Racine, Rafael Meza, Stephanie Land, Jihyoun Jeon, and Eric J Feuer. Patterns of birth cohort–specific smoking histories, 1965–2009. *American Journal of Preventive Medicine*, 46(2):e31–e37, 2014.

Paul W Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

Alex Hollingsworth, Jiafang Mike Huang, Ivan Rudik, and Nicholas J Sanders. A thousand cuts: Cumulative lead exposure reduces academic achievement. *Journal of Human Resources*, pages 0222–12169R2, 2022.

Mark Huggett, Gustavo Ventura, and Amir Yaron. Sources of lifetime inequality. *American Economic Review*, 101(7):2923–2954, 2011.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475, 1994.

Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Thomas J. Kane and Douglas O. Staiger. Improving school accountability measures. *NBER Working Paper*, 2001.

Lawrence F. Katz and Kevin M. Murphy. Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

Lawrence F Katz et al. Changes in the wage structure and earnings inequality. In *Handbook of labor economics*, volume 3, pages 1463–1555. Elsevier, 1999.

John Komlos. The height and weight of west point cadets: dietary change in antebellum america. *The Journal of Economic History*, 47(4):897–927, 1987.

John Komlos. The recent decline in the height of African-American women. *Economics & Human Biology*, 8(1):58–66, 2010.

John Komlos and Benjamin E Lauderdale. The mysterious trend in American heights in the 20th century. *Annals of human biology*, 34(2):206–215, 2007a.

John Komlos and Benjamin E Lauderdale. Underperformance in affluence: the remarkable relative decline in US heights in the second half of the 20th century. *Social Science Quarterly*, 88(2): 283–305, 2007b.

Daniel Koretz. *Educational Achievement: Explanations and Implications of Recent Trends*. Congressional Budget Office, 1987.

Edward L Korn and Barry I Graubard. *Analysis of health surveys*. John Wiley & Sons, 1999.

David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849, 2018.

BYH Lam, A Williamson, S Finer, FR Day, JA Tadross, A Gonçalves Soares, K Wade, P Sweeney, MN Bedenbaugh, DT Porter, et al. Mc3r links nutritional state to childhood growth and the timing of puberty. *Nature*, 599(7885):436–441, 2021.

Thomas Lemieux. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498, 2006.

Adriana Lleras-Muney. The relationship between education and adult mortality in the United States. *The Review of Economic Studies*, 72(1):189–221, 2005.

Adriana Lleras-Muney. Comment on "mortality and morbidity in the 21st century". *Brookings Papers on Economic Activity*, pages 444–476, 2017.

Adriana Lleras-Muney, Alessandro Tarozzi, Elisabetta Aurino, and Brendan Tinoco. The rise and fall of SES gradients in heights around the world. *Available at SSRN 4127021*, 2022.

Michael C. Lu and Neal Halfon. Racial and ethnic disparities in birth outcomes: a life-course perspective. *Maternal and child health journal*, 7(1):13–30, 2003.

Liying Luo and James S Hodges. Constraints in random effects age-period-cohort models. *Sociological Methodology*, 50(1):276–317, 2020.

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.

Ryan K. Masters, Andrea M. Tilstra, and Daniel H. Simon. Explaining recent mortality trends among younger and middle-aged White Americans. *International Journal of Epidemiology*, 47 (1):81–88, 2017.

Bhashkar Mazumder. Does education improve health? a reexamination of the evidence from compulsory schooling laws. *Economic Perspectives*, 32(2), 2008.

Bhashkar Mazumder. Erratum: Does education improve health? a reexamination of the evidence from compulsory schooling laws. *Chicago Fed Working Paper*, 2010.

Bhaskar Mazumder. The effects of education on health and mortality. *Nordic Economic Policy Review*, 1(2012):261–301, 2012.

Mark McClellan and Douglas Staiger. Comparing the quality of health care providers. In *Forum for Health Economics & Policy*, volume 3. De Gruyter, 2000.

Michael J McFarland, Matt E Hauer, and Aaron Reuben. Half of US population exposed to adverse

lead levels in early childhood. *Proceedings of the National Academy of Sciences*, 119(11): e2118631119, 2022.

David J. McKenzie. Disentangling age, cohort and time effects in the additive model. *Oxford Bulletin of Economics and Statistics*, 68(4):473–495, 2006.

Anthony J McMichael, Graham V Vimpani, Evelyn F Robertson, Peter A Baghurst, and Peter D Clark. The Port Pirie cohort study: maternal blood lead and pregnancy outcome. *Journal of Epidemiology & Community Health*, 40(1):18–25, 1986.

Howard V Meredith. Change in the stature and body weight of North American boys during the last 80 years. In *Advances in child development and behavior*, volume 1, pages 69–114. Elsevier, 1964.

Keith Meyers. In the shadow of the mushroom cloud: Nuclear testing, radioactive fallout, and damage to US agriculture, 1945 to 1970. *The Journal of Economic History*, 79(1):244–274, 2019.

Howard W Mielke, Mark AS Laidlaw, and Chris Gonzales. Lead (pb) legacy from vehicle traffic in eight California urbanized areas: continuing influence of lead dust on children's health. *Science of the total environment*, 408(19):3965–3975, 2010.

Jacob Mincer. *Schooling, Experience, and Earnings*. National Bureau of Economic Research, 1974.

Engineering National Academies of Sciences, Medicine, et al. High and rising mortality rates among working-age adults. 2021.

National Commission on Excellence in Education. A nation at risk: The imperative for educational reform. 1983.

Herbert Needleman. Lead poisoning. *Annual Review of Medicine*, 55(1):209–222, 2004.

P. R O'Rourke, S. J. Smith, A. Mott, H. Ahsan, E. E. McDuffie, M. Crippa, S. Klimont, B. McDonald, Z. Wang, M. B. Nicholson, L. Feng, and R. M. Hoesly. Ceds v-2021-02-05 emission data 1975-2019 (2021, february 05). *Zenodo, http://doi.org/10.5281/zenodo.4509372.*

Gil Oudijk. The rise and fall of organometallic additives in automotive gasoline. *Environmental Forensics*, 11(1-2):17–49, 2010.

Belinda Phipson and Gordon K Smyth. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. *Statistical applications in genetics and molecular biology*, 9(1), 2010.

Gary G. Price and Thomas P. Carpenter. On further examination: Report of the advisory panel on the Scholastic Aptitude Test score decline, 1978.

Jessica Wolpaw Reyes. Environmental policy as social policy? the impact of childhood lead exposure on crime. *The BE Journal of Economic Analysis & Policy*, 7(1), 2007.

Nicholas Reynolds. Increased mortality of white Americans and a decline in the health of cohorts born after World War II. *Working paper*, 2023.

Mark R. Rosenzweig and T. Paul Schultz. Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91(5):723–746, 1983.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.

Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. Integrated public use microdata series: Version 6.0 [dataset]. 2015.

Norman B. Ryder. The cohort as a concept in the study of social change. *American Sociological Review*, 30(6):843–861, 1965.

Sam Schulhofer-Wohl. The age-time-cohort problem and the identification of structural parameters in life-cycle models. *Quantitative Economics*, 9(2):643–658, 2018.

Sherry G Selevan, Deborah C Rice, Karen A Hogan, Susan Y Euling, Andrea Pfahles-Hutchens, and James Bethel. Blood lead concentration and delayed puberty in girls. *New England Journal of Medicine*, 348(16):1527–1536, 2003.

Ella Mae Shelton, Marvin L Whisman, and Paul W Woodward. Trends in motor gasolines: 1942-1981. Technical report, Department of Energy, 1982.

Ellen K Silbergeld. Lead in bone: implications for toxicology during pregnancy and lactation. *Environmental health perspectives*, 91:63–70, 1991.

Richard H. Steckel and Donald R. Haurin. Height, nutrition, and mortality in ohio, 1870-1900. *Working paper*, 1982.

James M Tanner. Growth as a target-seeking function: catch-up and catch-down growth in man. *Human Growth: A Comprehensive Treatise Volume 1 Developmental Biology Prenatal Growth*, pages 167–179, 1986.

James Mourilyan Tanner. *Foetus into man: Physical growth from conception to maturity*. Harvard University Press, 1990.

Andrea M Tilstra and Ryan K Masters. Worth the weight? recent trends in obstetric practices, gestational age, and birth weight in the United States. *Demography*, 57(1):99–121, 2020.

Vladimir Turusov, Valery Rakitsky, and Lorenzo Tomatis. Dichlorodiphenyltrichloroethane (ddt): ubiquity, persistence, and risks. *Environmental health perspectives*, 110(2):125–128, 2002.

United States Bureau of Mines. Minerals yearbook. Technical report, 1941-1970.

US Census Bureau. *Historical statistics of the United States, colonial times to 1970*. Number 78. US Department of Commerce, Bureau of the Census, 1975.

US Environmental Protection Agency. Ddt regulatory history: A brief survey (to 1975). Technical report, 1975.

Finis Welch. Effects of cohort size on earnings: The baby boom babies' financial bust. *Journal of Political Economy*, 87(5, Part 2):S65–S97, 1979.

Tiejian Wu, Germaine M Buck, and Pauline Mendola. Blood lead levels and sexual maturation in US girls: the third national health and nutrition examination survey, 1988-1994. *Environmental health perspectives*, 111(5):737–741, 2003.

Grace Wyshak and Rose E Frisch. Evidence for a secular trend in age of menarche. *New England journal of medicine*, 306(17):1033–1035, 1982.

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

Alwyn Young. Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics*, 134(2):557–598, 2019.

Emma Zang, Hui Zheng, Yang Claire Yang, and Kenneth C. Land. Recent trends in US mortality

in early and middle adulthood: racial/ethnic disparities in inter-cohort patterns. *International Journal of Epidemiology*, 2018.

Hui Zheng. A new look at cohort trend and underlying mechanisms in cognitive functioning. *The Journals of Gerontology: Series B*, 76(8):1652–1663, 2021.

Hui Zheng and Paola Echave. Are recent cohorts getting worse? trends in US adult physiological status, mental health, and health behaviors across a century of birth cohorts. *American Journal of Epidemiology*, 190(11):2242–2255, 2021.

Hui Zheng, Yoonyoung Choi, Jonathan Dirlam, and Linda George. Rising childhood income inequality and declining Americans' health. *Social Science and Medicine*, 303:115–116, 2022.

Supplementary Appendix for:

"The Broad Decline of Health and Human Capital of Americans Born after 1947"

by Nicholas Reynolds

September 19, 2023

Appendix A: Permutation-based inference

This appendix provides more details on the permutation-based inference described in Section 7. I present the test and show that it has correct size in expectation. I then describe how I use this test to construct confidence intervals.

I will test the so-called "sharp null" hypothesis that the second-difference in cohort effects is equal to zero in all years:

$H_0 : \Delta\Delta^{ap}\psi_c = 0$ for all a, p s.t. $c = p - a$

I take the sequence of observed first difference in outcomes for the neighboring cohorts $\tilde{c}+1$ and $\tilde{c}$ as data: $\left\{\Delta^p Y_{a-1,p,\tilde{c}+1}, \Delta^p Y_{a,p,\tilde{c}}\right\}_{p\in\mathcal{P},a=p-\tilde{c}}$.

Consider then all possible sequences which adjust this sequence by permuting any number of the neighboring first differences, and keeping the other first differences in the order observed. Ideally I would use this entire set of possible sequences to conduct the permutation test. As a computationally feasible approximation, I instead make K repeated draws where in each draw: for each pair of neighboring first differences I randomly either permute them or keep them as observed.

I will use as a test statistic the average second difference estimator, $\hat{\Psi}_c$, defined in Section 7. Denote this test statistic defined based on the observed data as $\hat{\Psi}_{\tilde{c}}$. Define the corresponding test statistics calculated from the permuted data from each draw 1 to K as $\hat{\Psi}^{(1)}$, $\hat{\Psi}^{(2)}$ ... $\hat{\Psi}^{(K)}$.

Note that under the dual assumptions of exchangeability and the null hypothesis, the sequence of test statistics calculated on the observed data and based on the K draws, $\{\hat{\Psi}_{\tilde{c}}, \hat{\Psi}^{(1)}, \hat{\Psi}^{(2)}$ ... $\hat{\Psi}^{(K)}\}$, is a sequence of exchangeable random variables. Therefore any ordering of the sequence from smallest to largest is equally likely. Assuming the test statistic is continuous and therefore one can ignore ties, then the probability that $\hat{\Psi}_{\tilde{c}}$ is greater than m of the permuted test statistics is simply $\frac{m+1}{K+1}$.[55] One can therefore calculate a 1-sided permutation p-value as:

$$\hat{p}_{\tilde{c},1} = \frac{1 + \sum_{k=1}^{K} \mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1}$$

And such a p-value has correct size:

[55] See Phipson and Smyth (2010) for a similar argument in a different setting.

$$P(\hat{p}_{\tilde{c},1} \geq \alpha) = P\left(\frac{1 + \sum_{k=1}^{K} \mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1} \geq \alpha\right)$$
$$= P\left(\sum_{k=1}^{K} \mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right) \geq \alpha(K+1) - 1\right) = \alpha$$

In practice I use the following two-sided p-value:

$$\hat{p}_{\tilde{c}} = 2 \cdot \min\left(\frac{1 + \sum_{k=1}^{K} \mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1}, \frac{1 + \sum_{k=1}^{K} \mathbb{1}\left(\Psi_{\tilde{c}} \leq \hat{\Psi}^{(k)}\right)}{K+1}\right)$$

I also construct confidence intervals by inverting the permutation test following the approach in Imbens and Rubin (2015); Ganong and Jäger (2018). The confidence interval can be interpreted as the set of constant second-differences in local cohort effects which the test fails to reject. In particular I find interval the B such that for all $b \in B$ would fail to reject null hypothesis that: $H_0 : \Delta\Delta^{ap}\psi_c = b$ for all a, p s.t. $c = p - a$

I follow a similar approach to that described in Ganong and Jäger (2018) to construct the confidence intervals. I test the above null for a given value b by reconstructing the data adding in a hypothetical effect b to one of the first differences. That is I replace the sequence listed above with the following reconstructed data: $\{\Delta^p Y_{a-1,p,\tilde{c}+1}, \Delta^p Y_{a,p,\tilde{c}} + b\}_{p \in \mathcal{P}, a=p-\tilde{c}}$. I then construct a permutation p-value following the approach described above. I use a bisection algorithm to find the endpoints of the confidence interval.

Appendix B: More details on growth and physical development data, analysis, and results

Stagnating adult height

To study adult height, I combine NHES I, NHANES I to III, and the Continuous NHANES from 1999 to 2018. Year of birth is not directly recorded in some of the surveys so I impute it with some error based on age and (approximate) year of survey. For NHES I, I define year of birth as 1960 minus age (the survey was conducted in 1959-1962). For NHANES III, I define year of birth as 1990 minus age for those surveyed in Phase 1 (October 18, 1988, through October 24, 1991) and 1993 minus age for those surveyed in phase 2 (September 20, 1991, through October 15, 1994). The continuous NHANES is organized into two-year waves and the exact year of survey within the wave is not recorded. For these surveys, I define year of birth as the later of the two survey years minus age. Native born is not reported in NHES I, so I include all individuals.

Smaller adolescent growth spurt and later leg growth

I combine data from the NHES II, NHES III, and NHANES I to III samples to study changes in height, leg length, sitting height in childhood, adolesence, and adulthood. Height and sitting height are directly measured and leg length is calculated as their difference. I focus on individuals born in the US. For the NHES surveys this is defined based on a linkage to birth certificates and for the later surveys is self-reported.

I use regressions of the following form to estimate differences in height (or other outcomes) for different age groups:

$$Y_{i,ac} = \gamma_{c=1960-66} + \mu + f(a) + \epsilon_{i,ac} \tag{6}$$

where $Y_{i,ac}$ denotes height of individual i at age a in cohort c. I run the above regression separately for different age groups: 6 to 11, 12 to 17, and 23 to 40. In each regression $\hat{\gamma}_{c=1960-66}$ estimates height differences between cohorts, and $f(a)$ controls for small differences in the distribution of ages when the cohorts are observed in the different surveys. For children I include dummies for age in years and a quadratic in age in months. I restrict to the native-born population and use sampling weights in all analysis.

I estimated additional models to directly estimate and gauge the statistical significance of the differential growth in height implied by these patterns. For the sample of 6 to 17 year olds I estimate:

$$Y_{i,ac} = \gamma^{a=12-17}_{c=1960-66} + \mu + f(a) + \epsilon i, ac \tag{7}$$

And for the sample of 12 to 17 and 23 to 40 year olds I estimate:

$$Y_{i,ac} = \gamma^{a=23-40}_{c=1960-66} + \mu + f(a) + \epsilon i, ac \tag{8}$$

To study age at menarche I use restrospective questions asked to women about the age when their periods first began. I use NHANES I to III and the continuous NHANES up to 2018, and include native-born women aged 20 to 75. Year of birth is imputed where necessary as described above.

Appendix C: Connection between educational declines and other declines

A. Implied causal effects of education — assuming no change in unobservables

I present a first test of the "education-only" hypothesis, by considering the implied causal effects of education on the other outcomes, under the assumption that unobservables remained unchanged across cohorts. Under the assumption that education is the only factor that changed across cohorts, one can use the post-late-1940s cohort trend break as an instrument for education.

For example, the estimated cohort break in log wages over the change in the cohort slope in years of schooling can be used as a two-stage least squares estimator of the return to a year of schooling. Analogous ratios for infant health and mortality can generate implied causal effects of mother's schooling on infant health and of education on mortality. By comparing these estimates to prior results in the literature I can assess the plausibility that the declines in other outcomes were caused by education alone.

I perform such a two-stage least squares exercise for log wage, the low birth weight rate, and the log mortality of men and women. I calculate a separate "first-stage" estimate of the cohort trend break in years of schooling for each of the outcomes to address the slightly different selection in each sample.[56]

To calculate the "reduced-form" I simply apply Equation 4 to each outcome, and estimate the location of the cohort break as above following Hansen (1999, 2000). For the first-stage I then impose the location of the cohort trend break to occur at the same location as in the corresponding reduced-form. I then estimate by least squares a model similar to Equation 4 with years of schooling on the left-hand side, and with the location of the trend break, λ, treated as known. Finally, the two-stage least squares estimate is simply the ratio of cohort break, delta, from the "reduced-form" over that in the "first-stage." [57]

Table A3 reports the results. Interestingly, the implied return for wages to a year of schooling is .110 which is only slightly larger than OLS estimates and closely matches many of the IV estimates summarized in Card (2001).

Turning to low birth weight, my two-stage-least-squares estimate of -.92 implies that a year of maternal education reduces the incidence of low birth weight by .9 percentage points. This estimate is nearly twice the cross-sectional correlation between low birth weight and years of schooling calculated from the same data. However, it is remarkably similar to estimates from Currie and Moretti, who use the founding of colleges in a mother's county at age 17 as an instrument, and range from -0.96 to -.99.

The fact that the IV estimates are *larger* than the OLS estimate could be the result of measurement error in maternal years of schooling. Or as Currie and Moretti emphasize, it could be, following the logic regarding wages of Card (2001), because those "marginal women" induced

[56]For wages I calculate years of schooling using the CPS MORG sample of employed men with non-missing wages. For infant health I calculate mother's years of schooling directly from the natality files. Because mother's education is only available after 1969 and is missing for a non-trivial fraction of mothers, I also re-estimate the cohort decline in a restricted sample of births with non-missing maternal education, and find a similar cohort trend break to that in the full sample. For the mortality "first-stage" I use Decennial Census microdata from 1970-2000 to estimate education levels for the full US population.

[57]For easier interpretation of the two-stage-least squares estimates for log wages I run all models on the individual microdata (rather than calculating medians in bins), so the models are for the mean log wage rather than the median. I also only include age and year fixed effects as controls. I use CEPR's imputation of topcoded wages.

to change schooling levels by the instrument have a larger causal effect of education on maternal health than the average in the population. However, if one assumes that the direction of more traditional "ability bias" is negative, then this "selection on gains" would have to be very large to explain the IV estimates I find. In particular, the returns to schooling of those induced to change schooling levels by changes in education "supply" across cohorts would need to be double that of the average in the population. Given the large changes in the educational distribution across cohorts, this would imply that the causal effect of education in the population is *extremely* variable.

In contrast to the above results for wages and maternal health, the implied log mortality effect of education appears implausibly large — in comparison to both the cross-sectional relationship and *most* past estimates of the causal effect of schooling on mortality. My two-stage-least-squares estimates, using the cohort-slope change as an instrument for years of schooling and log mortality as the dependent variable, are -.27 and -.4 for men and women respectively. This would imply that earning a 4-year Bachelor's degree causes a male's mortality risk in a given year to be reduced by *nearly two-thirds* (specifically it would fall to 34 percent of it's prior rate). The implied effect of a 4-year Bachelor's degree for women is even larger — suggesting it would decrease the mortality rate to *nearly 20 percent* of it's prior rate.

These estimates are larger than cross-sectional differences. For example, in the 1980s the mortality rate at ages 25 to 64 of men with 16+ years of education was 49 percent lower than for men with 12 years of education. Mortality of women with 16+ years of education was just 31 percent lower than for women with 12 years of education (Elo and Preston, 1996).

Further, instrumental variable estimates of the causal effect of education on mortality using changes in compulsory schooling laws generally find *much* smaller causal effects than the cross-sectional relationship. The estimates generally suggest that a year of schooling reduces mortality by 0 to 6 percent.[58] In a notable outlier, Buckles et al. (2016) use a different research design based on Vietnam draft avoidance in the US and find IV estimates of the impact of a year of college schooling on men's mortality that are nearly identical to the cross-sectional association.[59]

The size of the cohort trend break in mortality appears likely *much* too large to be explained by the decline in education alone, and therefore strongly suggests that there was a broader decline

[58] See Galama et al. (2018); Mazumder (2012) for reviews. For example, Clark and Royer (2013) estimate a precise zero effect of schooling on adult mortality in a particularly credible study based on UK schooling reforms. Gathmann et al. (2015) pool data from 19 European countries and estimate that a year of schooling reduces the mortality rate of men by 2.8 percent, but find no statistically significant effects for women; the largest individual country estimate for men from that paper is only 5.6 percent The credibility of the approximately 6 percentage point estimate of Lleras-Muney (2005) for the US is debated by Mazumder (2010, 2008) and Black et al. (2015).

[59] It is hard to directly compare Buckles et al. (2016) causal effect estimate to my result as their dependent variable is cumulative mortality over a 26 year period. As their IV estimates are indistinguishable from OLS, I treat their causal effect estimate as near the cross-sectional association listed above. It is also hard to know more generally how to treat their estimate, as it is such an outlier in the literature and their design differs considerably in using quadratic functions of state and national draft risk as instruments for both veteran status and educational attainment.

in health and human capital for these cohorts.

B. Cohort declines in earnings ability and health conditional on education

I now show evidence of cohort declines in wages and maternal health within some narrowly defined educational bins — health and earnings ability appears to have declined for post-1947 cohorts even *conditional* on education. First, I show robust evidence of a trend break in cohort effects of median wages of those without a bachelor's degree. Then, I show evidence of similar trend break of cohort effects for maternal health, as measured by infant birth weight, at many levels of maternal education, including exactly 12 years of education and exactly 16 years of education.

These findings suggest either that the cohort break was driven by a decline in latent health and human capital *broader* than the educational declines alone; or that the change in educational attainment for late-1940s cohorts also involved large and unusual changes in selection effects, eg. the relationship between unobservables and years of schooling.

Wages by education

Table A4 shows estimates of the trend break in cohort effects in median log wages of employed men separately for those with and without a bachelor's degree. I estimate the models with piecewise linear cohort effects based on Equation 4. All models include age and year fixed effects, and from left to right each column includes higher order polynomials in age separately for each year — up to a quartic polynomial.

Panel A shows remarkably robust results for the wages of men without a bachelors degree: a cohort decline beginning with the 1947 cohort of similar size to that estimated for unconditional wages above. For all specifications of the control function the estimated location of the cohort trend break is 1947 with only that cohort included in the 99 percent confidence interval, and the F-type bootstrap tests all imply a p-value of less than .001 for the null of no break. The size of the estimated trend break varies from -.0124 to -.0207 — quite similar to the estimate of -.016 found above for unconditional wages for the entire sample of employed men. [60]

Panel B of Table A7 shows that there is much less evidence of a change in the cohort slope of earnings ability for college-educated men. In contrast to the non-college educated results in Panel A, the results for those with at least a Bachelor's degree vary widely across different specifications. The point estimate for the location of the trend break vary from 1941 to 1951, and a number of the confidence intervals are quite large. Further the estimated size of the cohort break varies widely and even reverses sign. The model with just age and year fixed effects shown in column 1 implies a small decline in ability beginning with the 1941 cohort — with a trend break of -.0057. In contrast,

[60] Because average years of schooling *within* the bin of men without a bachelor's degree also exhibits a trend break I apply the two-stage-least-squares procedure from the previous section. The implied return necessary for schooling alone to explain the decline is .157; much larger than the cross-sectional relationship, and larger than most, though not all, of the IV estimates reviewed in Card (2001).

with the addition of quadratic age polynomials in each year the estimated location of the cohort trend break moves to 1951, and the sign reverses with a size estimate of .0237.

The existence of a cohort decline for those *without* a bachelor's degree and no similar decline for those *with* a bachelor's degree implies a cohort-specific increase in the college-high-school wage gap — likely the same phenomenon documented in Card and Lemieux (2001a). Card and Lemieux (2001a) use census data and pool cohorts and ages into larger 5-age/cohort groupings. They therefore do not detect the sharpness by cohort of the decline in wages of those without a Bachelor's degree (and of the wage-gap). The discontinuous nature of the trend break appears hard to reconcile with their explanation based on falling relative supply of the college educated and imperfect substitutability of workers of different ages. For example, it would require workers only a year or two apart in age to be very poor substitutes.

The evidence presented in this paper of a broad decline in health and human capital may imply that a reassessment is due regarding the cause of cohort patterns in wages by education — and therefore of changes in the college high-school wage gap more generally. If there was a broad decline in health and human capital, it is not obvious why those with a bachelor's degree would not also have a decline in wages. One explanation would be that the factor which declined across cohorts is a *substitute* in the labor market with a college education, in other words that a bachelor's degree has a "protective effect" in the labor market against the broader decline in human capital.

As noted above, the decline in wages for those without a bachelor's degree could conceivably be explained by a large change in the selection of those who earn a bachelor's degree — even if the unconditional ability distribution remained unchanged across cohorts. However, these selection effects would need to be highly unusual. First, note that the share of each cohort who comprise the non-college-educated group *grew* substantially after the late 1940s cohorts, from near 65 percent for the 1946 cohort to near 73 percent by the 1960 cohort. For changing selection to explain the wage decline for this group, the marginal college-goers induced to not get a college-degree would need to be lower ability than the median "never-taker" — those who would not get a college-degree whether they were born in the late 1940s or the 1960s. That is nearly a quarter of the college-educated men in the 1946 cohort would need to be lower ability than the median man *without* a college degree.

Intergenerational infant health by mother's education

I now show evidence of similar trend break in cohort effects for maternal health, as measured by infant birth weight, at many levels of maternal education, including exactly 12 years of education and exactly 16 years of education.

I first estimate models on infant birth weight for the two maternal education levels which comprise the largest share of the population: exactly 12 and 16 years of schooling. I use reported mother's years of schooling in the natality detail files, and focus on mean birth weight to improve

precision.[61]

Table A5 shows that the piecewise linear cohort effect method similarly detects robust evidence of a change in the cohort slope of maternal health for mothers with exactly 12 and 16 years of education. Across specifications of the age-by-year control function, the location of the estimated trend break is either the 1946 or 1947 cohort for both groups. The estimated trend breaks across specification are all negative for both groups, and all tests for the existence of a trend break have implied p-values lower than .001.

Table A6 applies the piecewise linear cohort effect method to 5 educational categories, and shows evidence of a similar change in the cohort slope of maternal health across the maternal educational distribution. I apply the model based on Equation 4 with age and year fixed effects separately for the following 5 categories of maternal education: less than high school, high school, some college, 4 years college, and 5+ years of college. For each of the education levels the models detect a cohort break between the 1946 and 1948 cohorts. The size of the estimated trend breaks are all negative — implying a relative decline in cohort health — and vary from -.99 grams for the high school group to -3.31 for the some college group. Based on the bootstrap-based test of existence, all the changes in slope are significant. Appendix Table A7 shows that controlling for a separate quadratic-in-age in each year suggests declines in cohort slope at a similar location and at least twice the magnitude, for all groups except those with 5+ years of college.

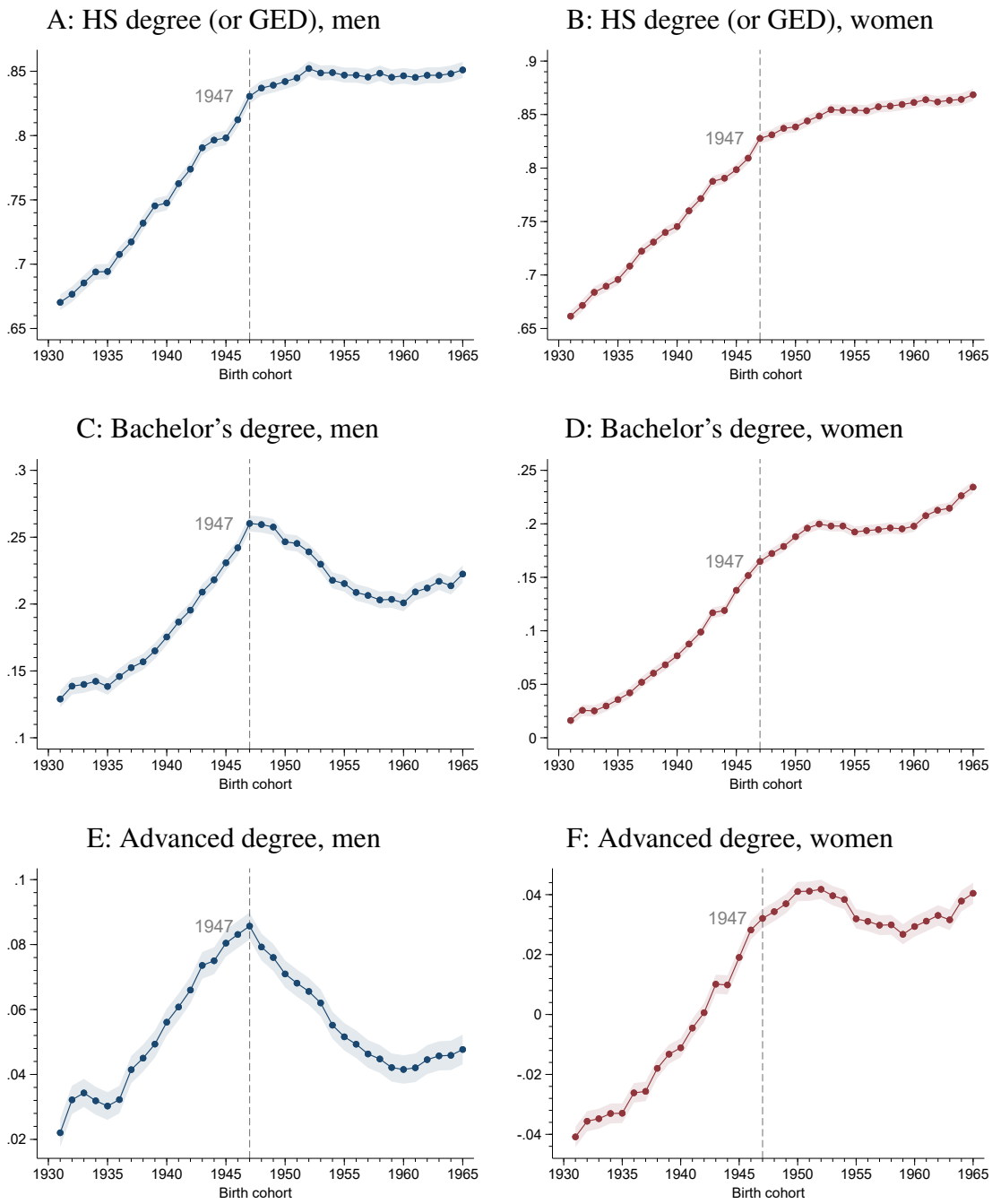
The above results show strong evidence of cohort declines in maternal health *conditional* on mother's education across the education distribution. These results would rule out an explanation in which a change in the supply of education alone has driven maternal health declines, if that change in supply affected individuals with differing latent health equally. Further, a selection based explanation would require that those who were induced to reduce their education level in the later born cohorts were less healthy than the average mother in the *lower* educational category to which they fell. As described above the changes in cohort educational shares were quite large, so this would imply a large portion of mothers of higher education had lower latent health than the average mother of a lower educational category.

One final piece of evidence against an "education-only" explanation is the different timing by cohort of the declines in women's college graduation rate and of the conditional maternal health declines. Recall that Panel C of Appendix Figure A1 clearly shows that the share of women with a bachelor's degree continued to increase until the 1951 cohort. This later change in the cohort slope for this outcome is confirmed with estimation of piecewise linear cohort effect models, in both the

[61] In addition to the reduced cell-size, note that the incidence of low birth weight births is much lower for highly educated mothers. Mother's education is missing for a non-trivial fraction of mothers, however the cohort decline in infant health is of similar size in the restricted sample of births to mothers with non-missing education information. Mother's education is only recorded beginning in 1969. I further restrict the sample to include mothers over the age of 22, the years 1969 to 1990, and the maternal cohorts 1938 to 1960.

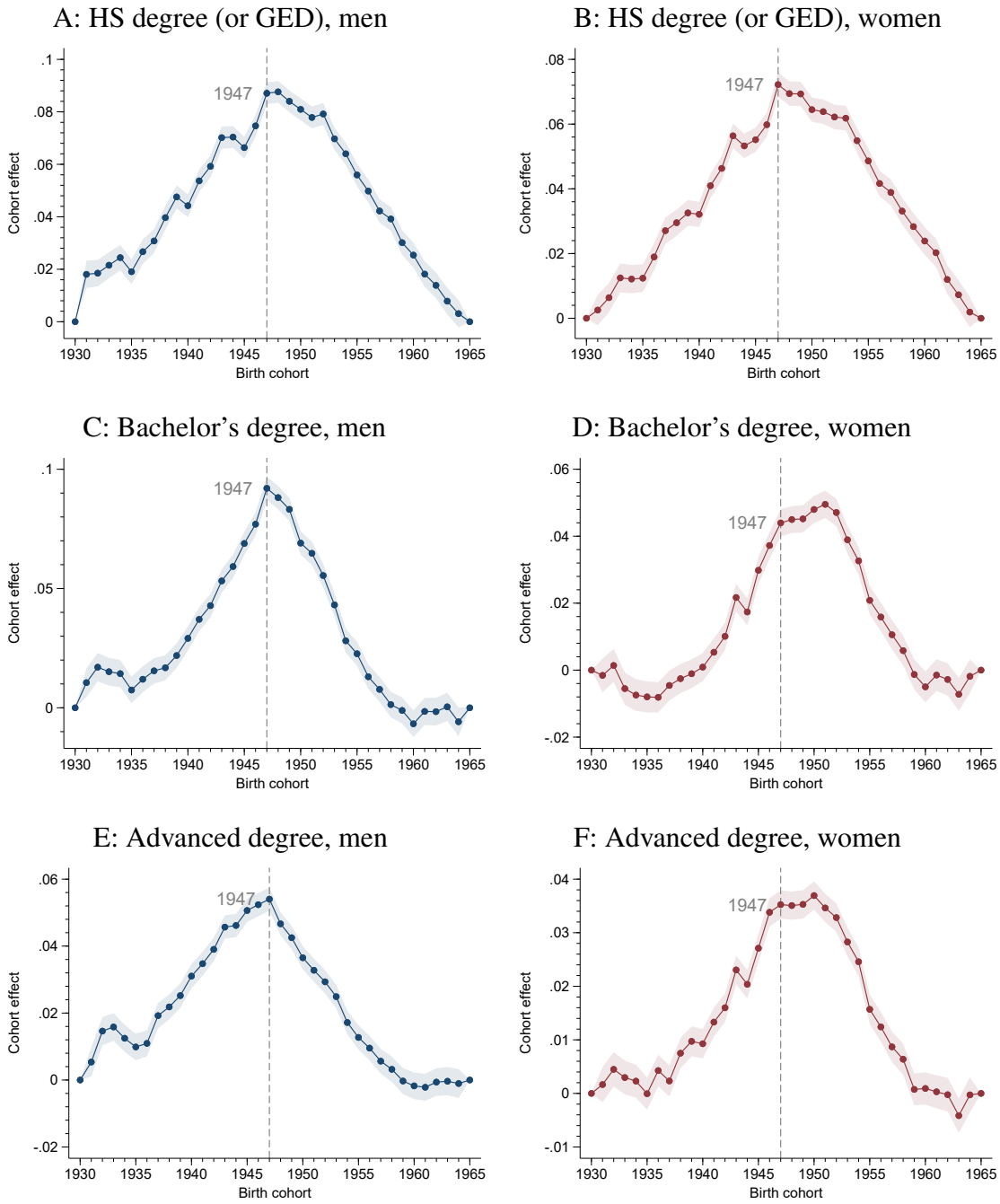
CPS and natality detail files. Recall that the declines in maternal health for those with exactly a bachelor's degree, some college, and exactly HS all declined precisely and sharply beginning with the 1947 or 1948. Reconciling this differential timing would require a complex selection story in which there was a sudden change in trend near the 1947 cohort in the relationship between latent health and bachelor's degree attainment, but no corresponding change in the trend in the total share of population who earn a bachelor's degree, until four cohorts later. This would require that the post-1947 change in selection was "non-monotonic" in the sense of Imbens and Angrist (1994) — that some individuals were induced to get a college degree after 1947, while others who would have previously were induced to *not* earn a college degree.

Figure A1: Age-adjusted educational attainment by birth cohort



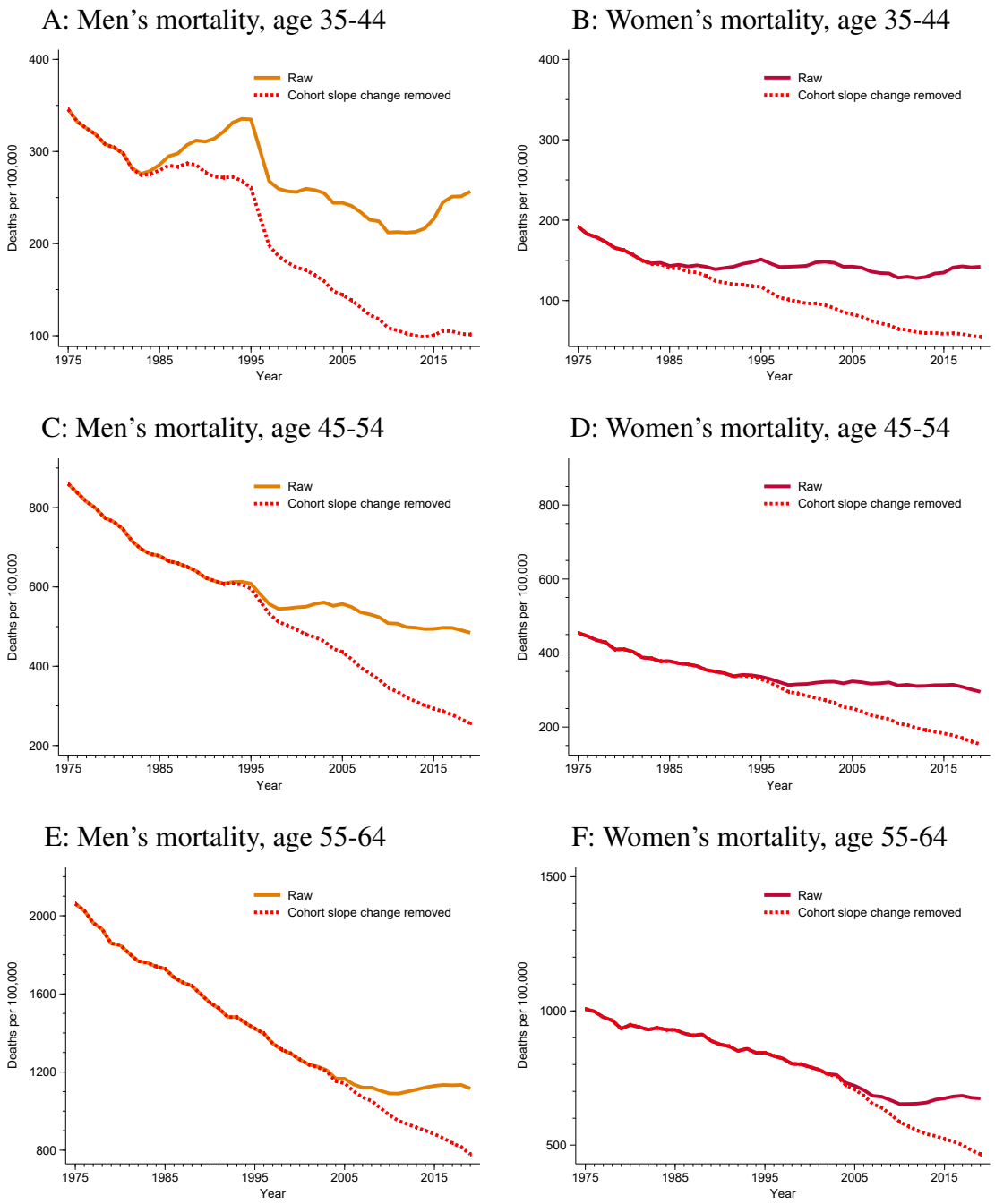
Each panel plots age-adjusted outcomes by birth cohort. Age adjustment is done by regressing the outcome on a full set of cohort fixed effects and age effects, and plotting predicted values. All panels are based on CPS-MORG data, 1979-2016, and includes men and women aged 25-75, who were born from 1930 to 1965.

Figure A2: Detrended cohort effects, educational attainment



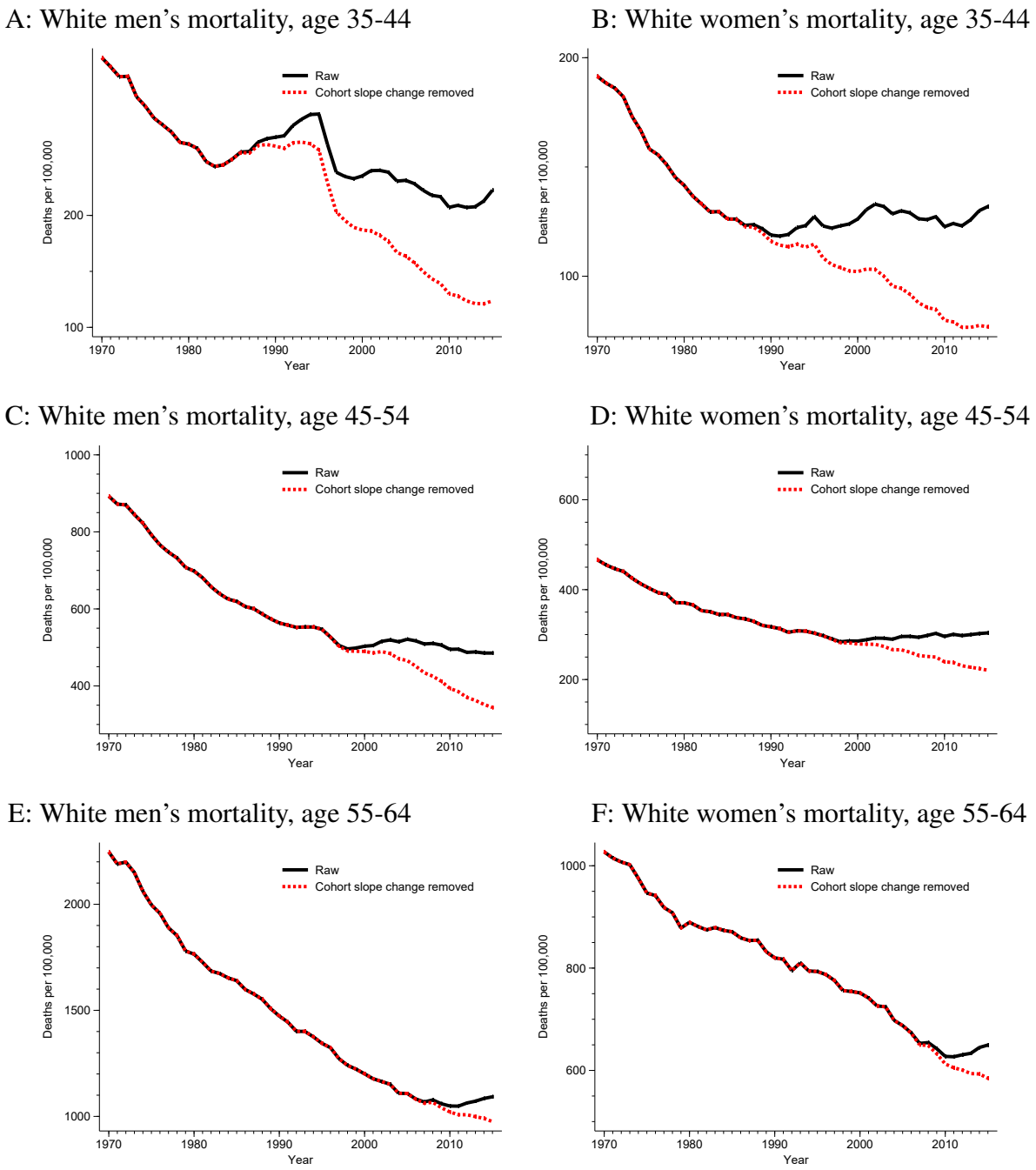
Each panel plots detrended cohort effects from estimation of age-period-cohort models based on Equation 1. All panels are based on CPS-MORG data, 1979-2016, and includes men and women aged 25-75, who were born from 1930 to 1965.

Figure A3: Impact of cohort decline on year-over-year trends in mortality



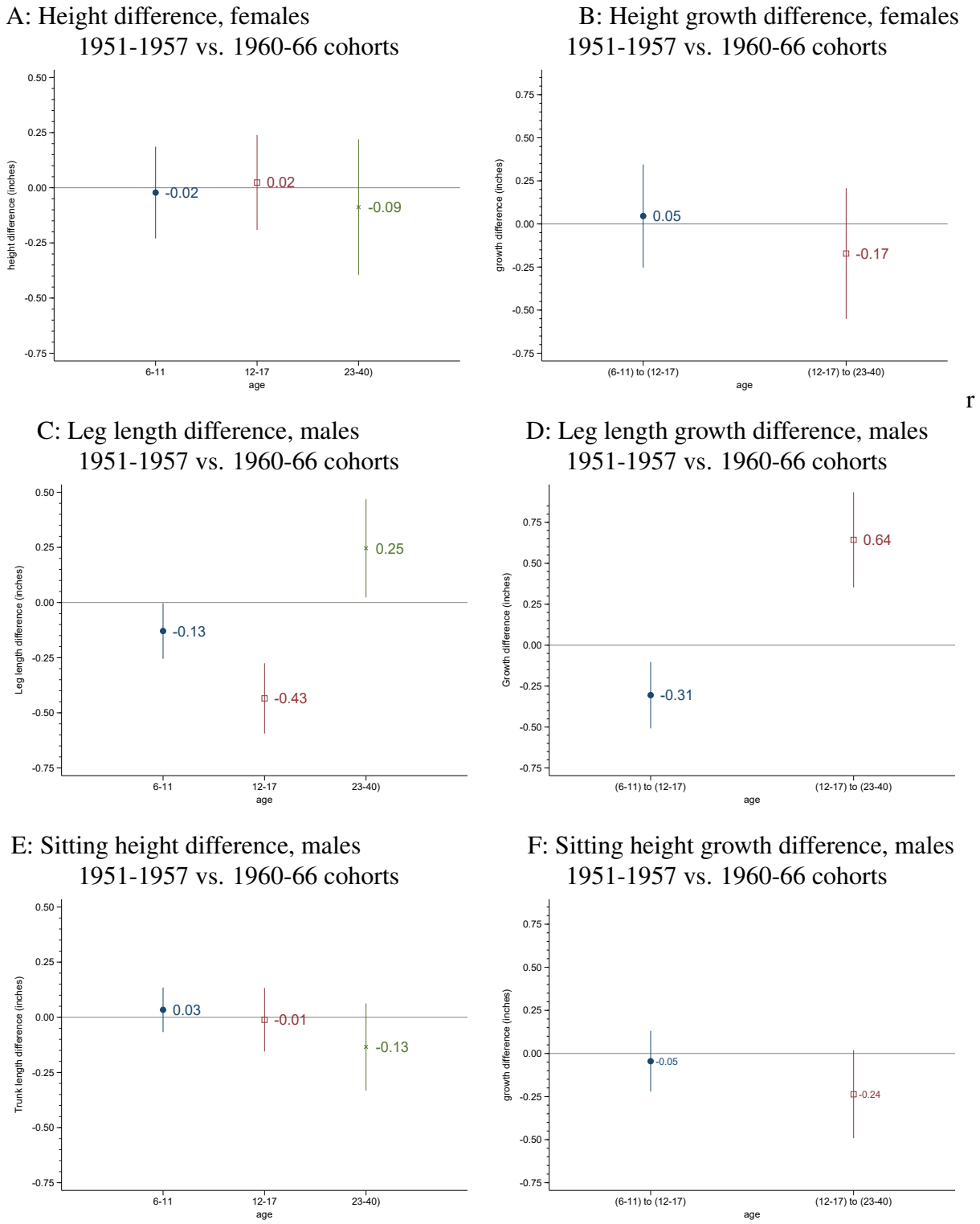
This figure shows simulated counterfactual year-over-year trends in the mortality rate had the trend break in cohort effects not occured, ie. if the pre-slope-change cohort trend had continued. The trend break estimated based on the model in Equation 4 is subtracted from each log mortality observation. These transformed data are then used to calculate each of the listed outcomes, and is plotted as the dashed red line. The same outcome based on the untransformed, raw data is plotted as a solid line. Each panel shows mortality for the listed age group, age-adjusted assuming a uniform population distribution by age. Underlying data is the Multiple Cause of Death file and SEER population estimates.

Figure A4: Impact of cohort decline on year-over-year trends in white mortality



This figure shows simulated counterfactual year-over-year trends in the white mortality rate had the trend break in cohort effects not occured, ie. if the pre-slope-change cohort trend had continued. The trend break estimated based on the model in Equation 4 is subtracted from each log mortality observation. These transformed data are then used to calculate each of the listed outcomes, and is plotted as the dashed red line. The same outcome based on the untransformed, raw data is plotted as a solid line. Each panel shows white for the listed age group, age-adjusted assuming a uniform population distribution by age. Underlying data is the Multiple Cause of Death file and SEER population estimates.

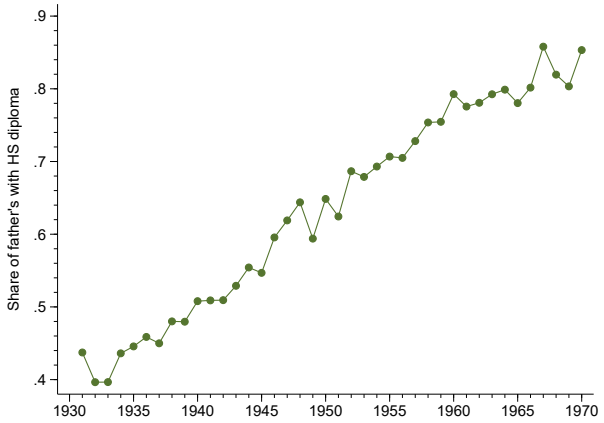
Figure A5: Growth and physical development, additional results



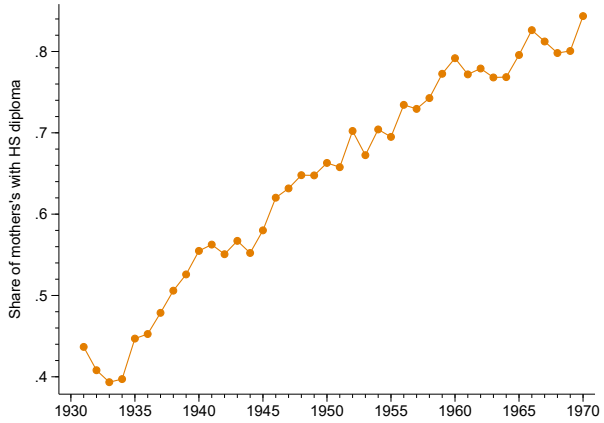
All panels use data from multiple rounds of NHES and NHANES surveys. Panels A and B show binscatter plots following Cattaneo et al. (2019) of heights of men and women ages 23 to 47. Panels C and D plot the height of white boys age 6 and 10 respectively. Pre-1950 data comes from non-representative historical studies compiled by Meredith (1964), with each point representing the mean for an interval of birth cohorts given in Meredith (1964), plotted at the midpoint. Post-1950 data comes from the NHES and NHANES surveys and shows the mean for 5-year birth cohorts. Panels E and F show estimates of the height difference of the listed cohorts at different ages and the difference in implied in implied growth between different ages, estimated from the regressions described in the text.

Figure A6: Parental education by birth cohort

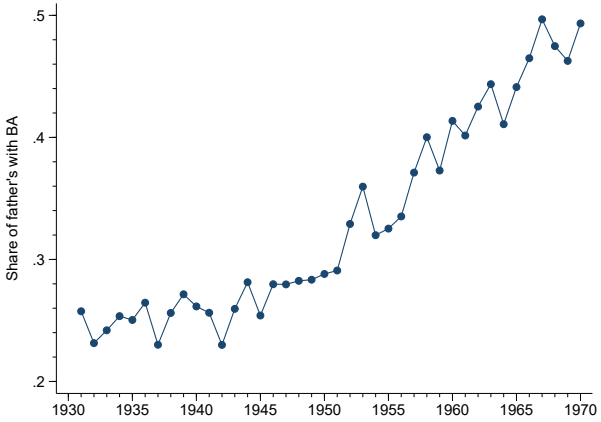
A: Share of fathers with high school diploma



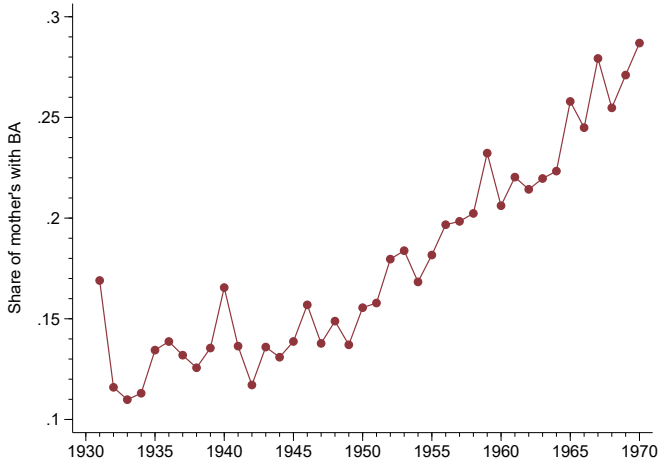
B: Share of mothers with high school diploma



A: Share of fathers with Bachelor's degree

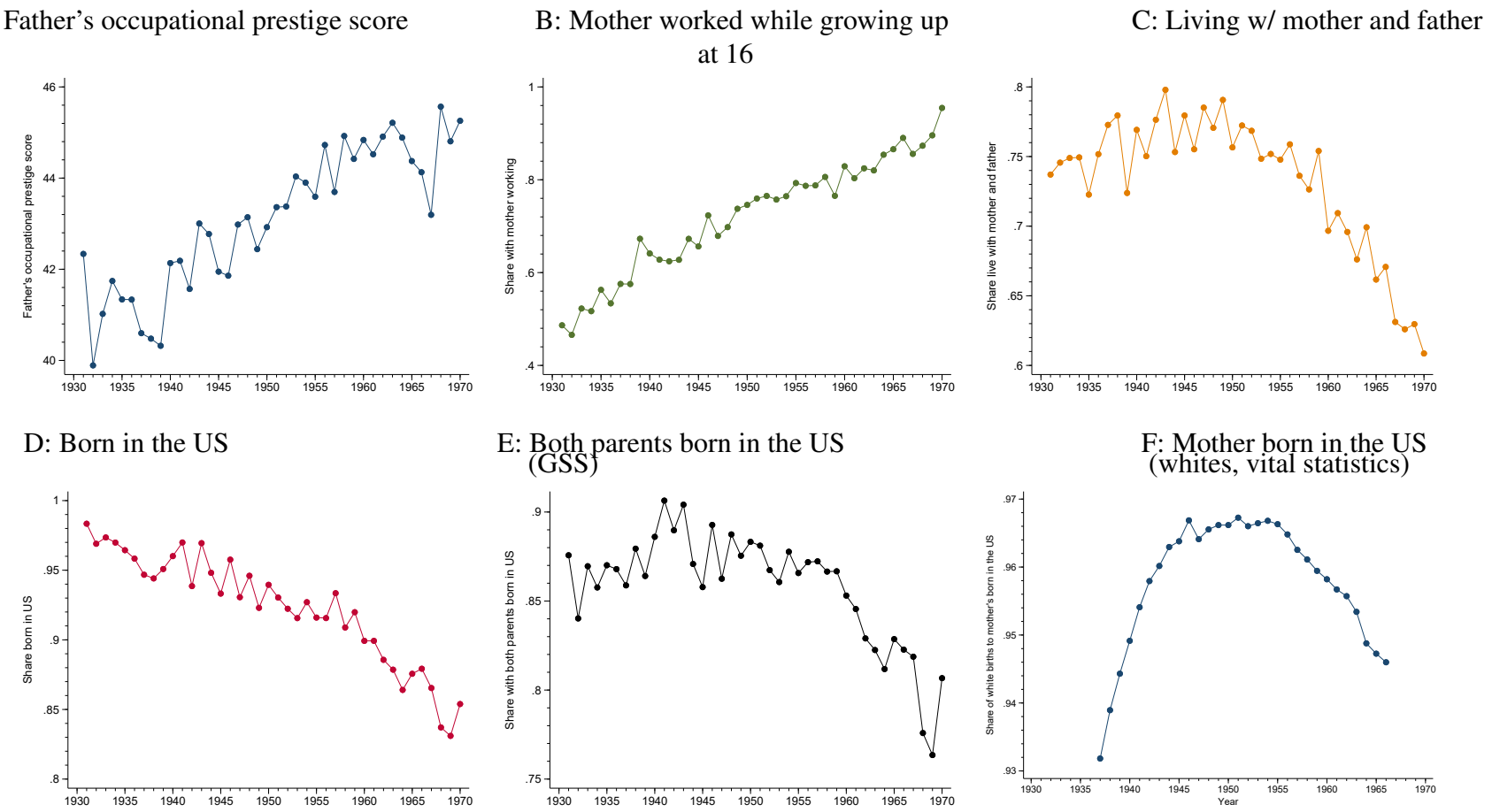


B: Share of mothers with Bachelor's degree



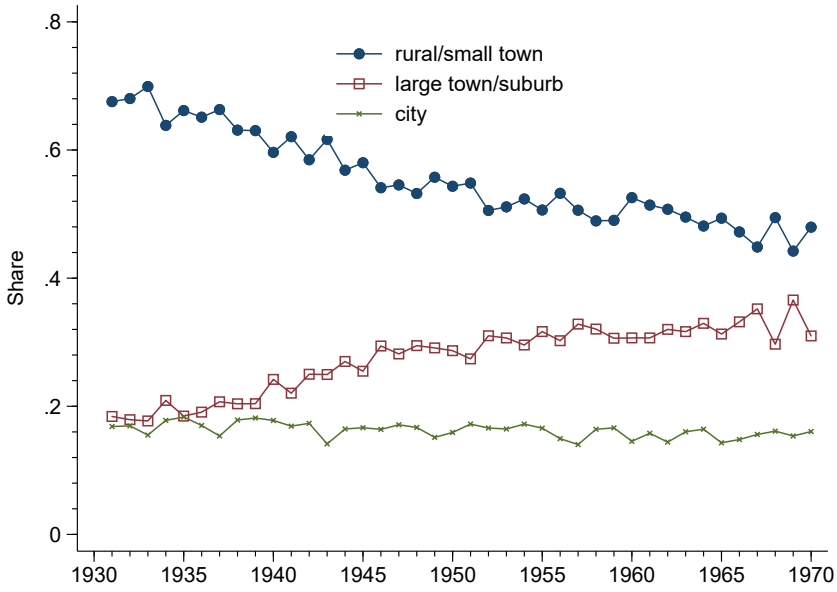
Each graph shows an estimate of parental educational attainment by individuals' birth cohort, estimated from the 1972-2016 waves of the General Social Survey. Each outcome is age-adjusted, by running a regression with cohort fixed effects and a quartic-in-age. The plots then show the estimated cohort effects, plus the estimated age effect for age 35. All regressions use sampling weights.

Figure A7: Family background and childhood circumstances by birth cohort



Panels A-E shows an estimate of the average value of the listed variable for Americans by year of birth, estimated from the General Social Survey. Father's occupational prestige score is based on 1980 occupational classifications and is only available in 1988-2010, so Panel A is based on those years. The question on whether an individual's mother was working is only available 1994-2016, so panel B is based on only those years. Panel C is based on 1972-2016. Panel D and Panel E are based on 1977-2016. Each outcome is age-adjusted, by running a regression with cohort fixed effects and a quartic-in-age. The plots then show the estimated cohort effects, plus the estimated age effect for age 35. All regressions use sampling weights. Panel F reports estimates directly from vital statistics volumes which report parent's nativity.

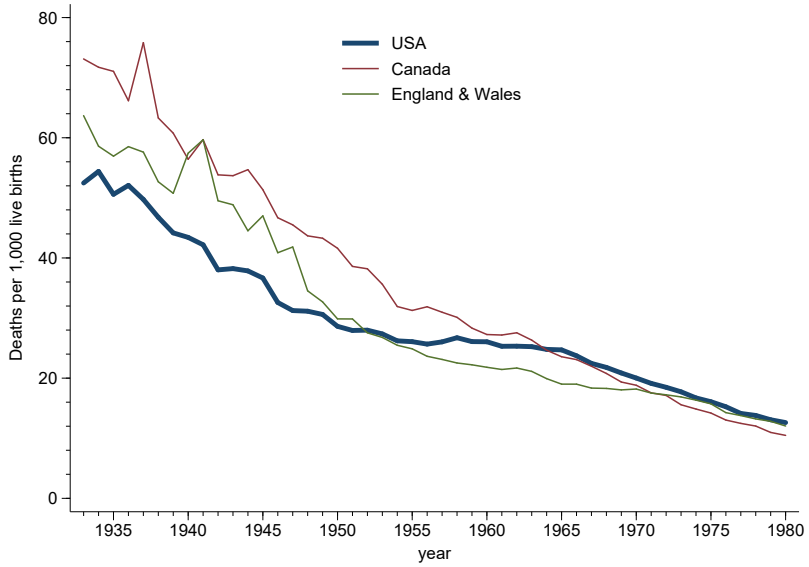
Figure A8: Where Americans lived at age 16, by birth cohort



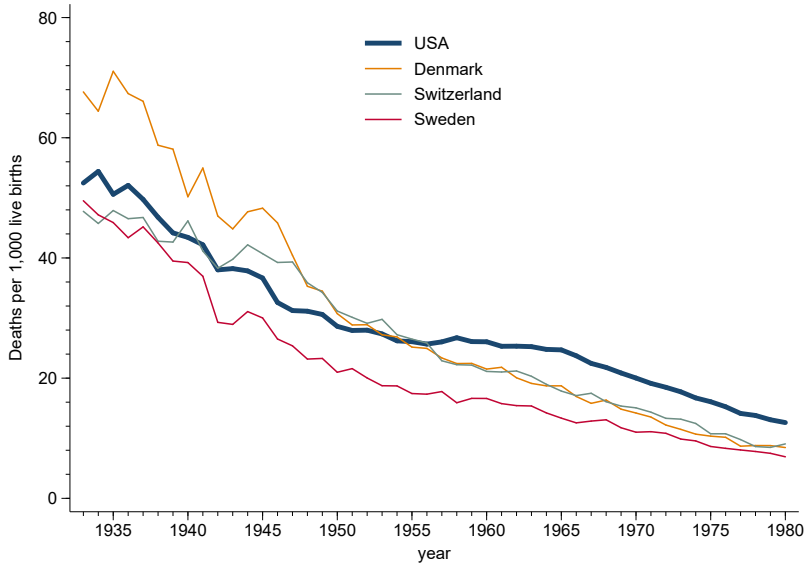
The figure shows estimates from the General Social Survey of the share of white Americans who lived in the listed type of place at age 16 by year of birth. ”City” refers to large cities over 250,000 people. ”Suburb” refers to a suburb near a large city. “Large town” refers to a city/town of 50,000 to 250,000. “rural/small town” includes smaller towns and rural areas. Each outcome is age-adjusted, by running a regression with cohort fixed effects and a quartic-in-age. The plots then show the estimated cohort effects, plus the estimated age effect for age 35. All regressions use sampling weights.

Figure A9: Infant mortality rate in US and comparison countries

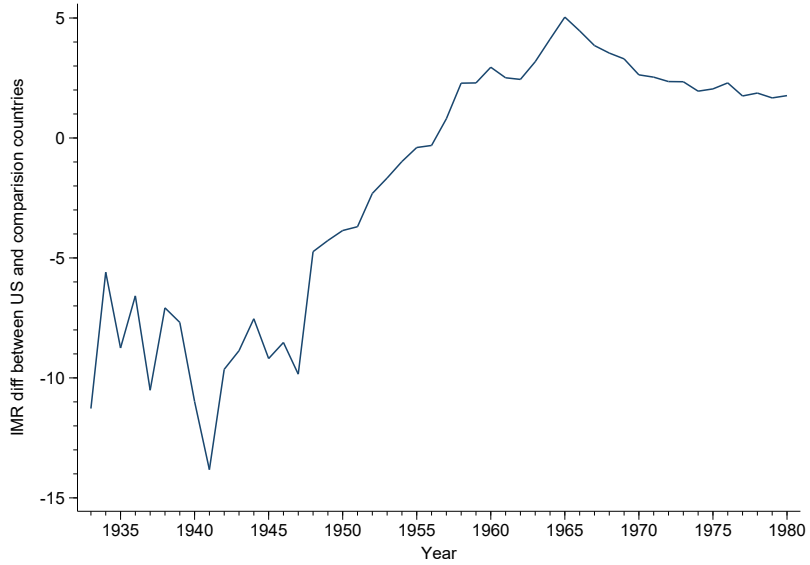
A: English-speaking countries



B: Scandinavian countries

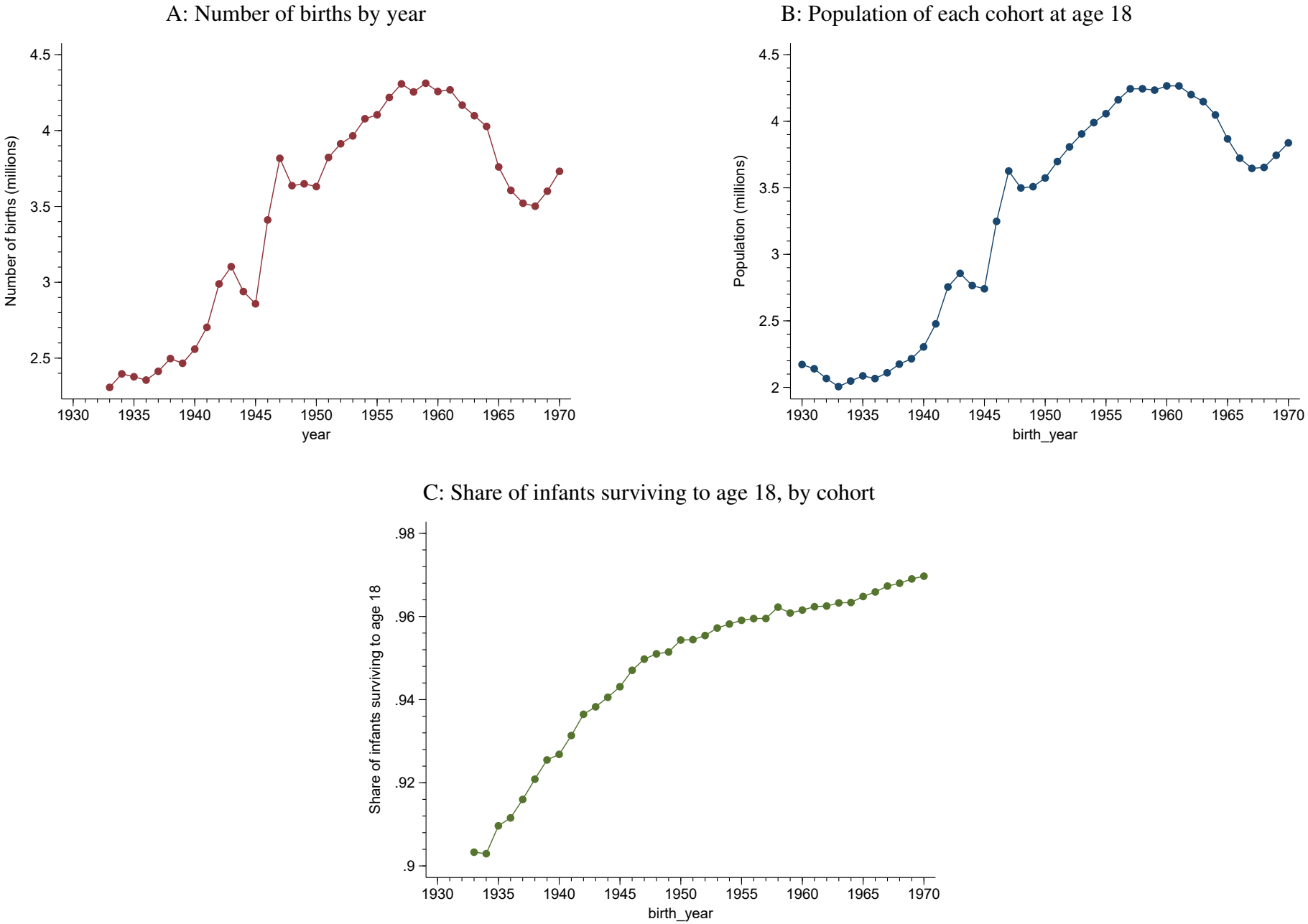


C: Difference between US and mean across Canada, England and Wales, Denmark, Switzerland, and Sweden



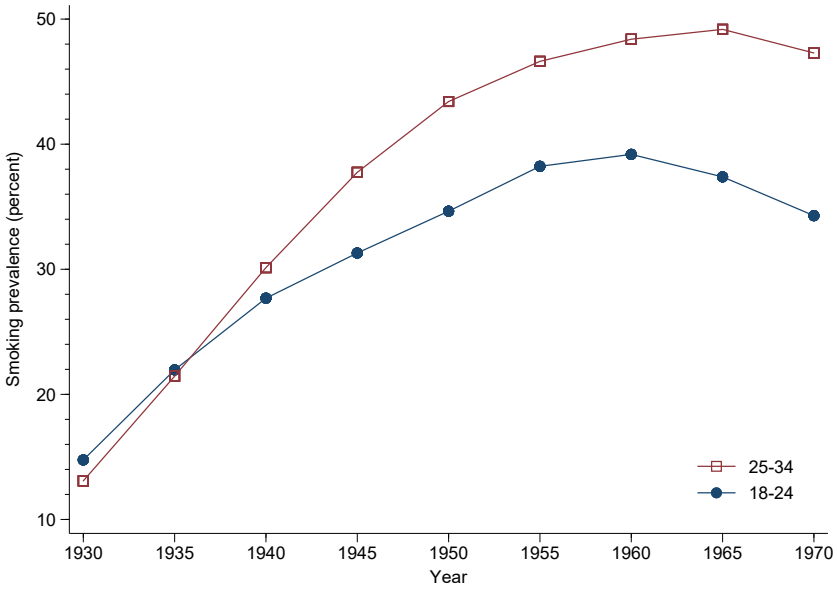
Panels A and B show the infant mortality rates in the listed countries. Panel C shows the difference between the infant mortality rate in the US and the mean across Canada, England and Wales, Denmark, Switzerland, and Sweden. The mean in the above figure is weighted by the number of births. Results are similar using the unweighted average. All data is from the Human Mortality Database.

Figure A10: Cohort size and survival



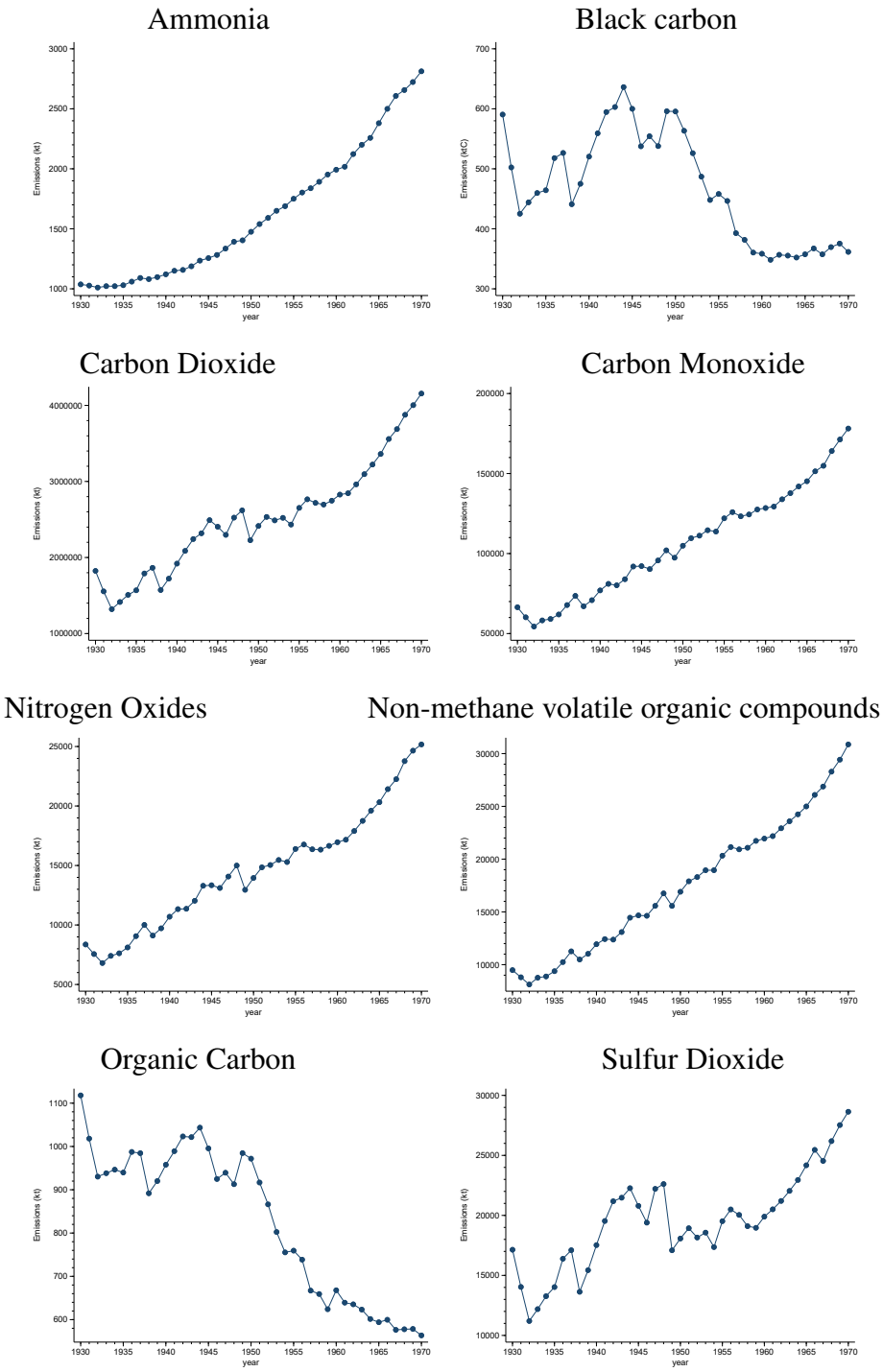
All data from the Human Mortality Database. Panel A shows the number of births in the US by year. Panel B shows the population of each cohort when they were age 18. Panel C shows the share of infants surviving to age 18 for each cohort — calculated by decrementing the cohort life tables.

Figure A11: Estimates of smoking prevalence of American women of childbearing age



Based on estimates of smoking prevalence by age and year (every 5 years) from Holford et al. (2014), derived from survey data on retrospective smoking history. The figure plots age-adjusted smoking prevalence for women age 18-24 and 25-34 separately, assuming a uniform distribution of ages within age bins (ie. the unweighted average in each age bin across smoking rates by single year of age).

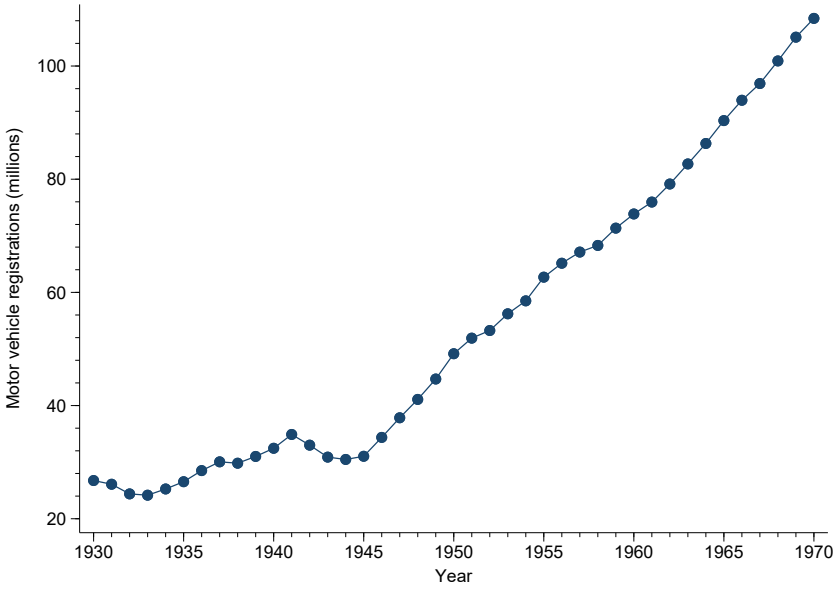
Figure A12: Air pollution trends, US



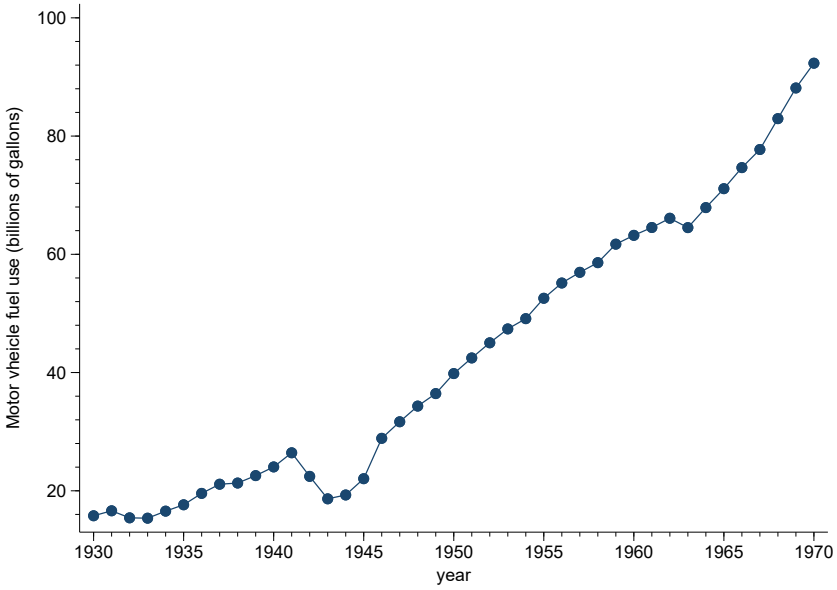
Each panel shows estimates from the Community Emissions Data System (O'Rourke et al.) of the trend in emissions of the listed air pollutant in the US.

Figure A13: Motor vehicle registrations and fuel use

A: motor vehicle registrations



B: Fuel usage by motor vehicles



This figure shows annual time series of a) motor vehicle registrations and b) fuel usage by motor vehicles, for 1930 to 1970 in the US. Both series come from Historical Statistics of the United States US Census Bureau (1975).

Table A1: Piecewise linear cohort effect models

robustness to varying age-by-year control function

	(1)	(2)	(3)	(4)
Panel A: Mean birth weight				
Size	-1.83	-6.35	-5.52	-6.91
	(0.13)	(0.35)	(0.54)	(0.70)
Location	1948	1948	1947	1947
	[1947, 1949]	[1948, 1948]	[1947, 1947]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel B: Low birth weight percentage				
Size	0.085	0.241	0.195	0.200
	(0.003)	(0.014)	(0.022)	(0.029)
Location	1947	1948	1947	1947
	[1947, 1948]	[1948, 1948]	[1937], [1948]	[1948, 1948]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel C: Median log wage				
Size	-0.016	-0.016	0.015	-0.017
	(0.0005)	(0.001)	(0.002)	(0.003)
Location	1947	1947	1953	1947
	[1947, 1947]	[1946, 1947]	[1946, 1947], [1952, 1953]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel D: Log mortality, men				
Size	0.016	0.028	0.029	0.029
	(0.001)	(0.001)	(0.001)	(0.001)
Location	1942	1947	1947	1947
	[1941, 1943]	[1947, 1947]	[1947, 1947]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel E: Log mortality, women				
Size	0.017	0.029	0.021	-0.019
	(0.000)	(0.001)	(0.001)	(0.002)
Location	1948	1949	1950	1950
	[1948, 1948]	[1949, 1949]	[1950, 1950]	[1949, 1950]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Quadratic-age-by-year	No	Yes	No	No
Cubic-age-by-year	No	No	Yes	No
Quartic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on Equation 4, with the listed outcome in single age-by-year bins as the dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). The row titled "Size" reports the size of the trend break in cohort effects, δ, with the standard error in parentheses. The row titled "Location" reports the estimated cohort at which a trend break occurs, with a 99 % confidence region in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear.

Table A2: Trend break in cohort effects for share of employed men working in white-collar occupations — native-born versus foreign-born

robustness to varying age-by-year control function

	(1)	(2)	(3)	(4)
Panel A: Born in US				
Size	-0.007	-0.017	-0.017	-0.014
	(0.0004)	(0.0009)	(0.002)	(0.002)
Location	1946	1946	1946	1946
	[1944, 1946]	[1946, 1946]	[1946, 1946]	[1946, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Born outside US				
Size	-0.008	0.027	0.031	0.018
	(0.001)	(0.005)	(0.008)	(0.006)
Location	1941	1932	1932	1956
	[1940, 1943]	[1932, 1933], [1943, 1944]	[1932, 1933]	[1932], [1936, 1937], [1956, 1957]
P-value for existence	$< .001$	$< .001$	$< .001$	0.007
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Quadratic-age-by-year	No	Yes	No	No
Cubic-age-by-year	No	No	Yes	No
Quartic-age-by-year	No	No	No	Yes

Based on Integrated Public Use Microdata Samples from the Decennial Censuses of 1970, 1980, 1990 and 2000 (Ruggles et al., 2015). The share of employed men in white-collar occupations corresponds to the "managers /professionals /technicians /finance /public safety" category in Autor and Dorn (2013), and I use the occupational crosswalk from these authors to obtain consistent occupational categories. Each column shows the results of estimation of a model based on equation 4, with the the share of employed men in white-collar occupations as the dependent variable in single age-by-year bins as the dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). The row titled "Size" reports the size of the trend break of cohort effects, δ, with the standard error in parentheses. The row titled "Location" reports the estimated cohort at which a trend break occurs, with a 99 $\%$ confidence region in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear.

Table A3: Implied causal effect of schooling on wages, maternal health, and mortality assuming no change in unobservables across cohorts

	Outcome *(reduced-form)*	cohort break Years of schooling *(first stage)*	Implied causal effect *(2sls)*
Unconditional			
Log wage	-0.013 (0.0003)	-0.119 (0.002)	0.110 (0.002)
Percent low birth weight	0.16 (0.02)	-0.173 (.027)	-0.92 (.09)
Male log mortality	.029 (0.001)	-.104 (.001)	-.270
Female log mortality	.028 (0.001)	-.072 (.001)	-.399
Conditional on no Bachelor's degree			
Log wage	-0.011 (0.0003)	-0.074 (0.002)	0.152 (0.005)

This table reports the results of two-stage-least-squares estimation of the causal effect of a year of schooling on wages and health, under the assumption that unobservables are unchanged across cohorts. Each row shows the results for the listed outcome. The column titled "Outcome (reduced-form)" reports the estimated size of a trend break in cohort effects, δ, from a model based on Equation 4 for the listed outcome. The column titled "Years of schooling (first stage)" reports the estimated size of a trend break in cohort effects from estimation of a similar model with years of schooling on the left-hand side, and with the location of the trend break, λ, set equal to the estimated location in the reduced-form and treated as known. The column titled "Implied causal effect (2sls)" uses these estimates to form a two-stage-least-squares estimate of the causal effect of years of schooling on the outcome. Log wage results are based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1965. Share low birth weight results are based on vital statistics natality microdata, 1968-1995, mothers age 18-40 who were born between 1930 to 1970. Log mortality results are based on data from the Human Mortality Database, and include the years 1975-2019, ages 25-85, cohorts born from 1930 to 1965.

Table A4: Trend break in cohort effects for median log wage of employed men — separately for those with and without a Bachelor's degree

robustness to varying age-by-year control function

	(1)	(2)	(3)	(4)
Panel A: Without Bachelor's Degree				
Size	-0.0132	-0.0124	-0.0183	-0.0207
	(0.0005)	(0.0015)	(0.0029)	(0.0031)
Location	1947	1947	1947	1947
	[1947, 1947]	[1947, 1947]	[1947, 1947]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel B: With Bachelor's degree				
Size	-0.0057	0.0237	-0.0265	-0.0235
	(0.0010)	(0.0028)	(0.0034)	(0.0056)
Location	1941	1951	1943	1943
	[1938, 1943]	[1950, 1952]	[1942, 1951]	[1936, 1962]
P-value for existence	$< .001$	$< .001$	$< .001$	.002
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Quadratic-age-by-year	No	Yes	No	No
Cubic-age-by-year	No	No	Yes	No
Quartic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on Equation 4, with the listed outcome in single age-by-year bins as the dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). Results are based on CPS-MORG data, 1979-1993, and includes men age 25-54, who were born from 1930 to 1965. The row titled "Size" reports the size of the trend break in cohort effects, δ, with the standard error in parentheses. The row titled "Location" reports the estimated cohort at which a trend break occurs, with a 99 $\%$ confidence region in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear.

Table A5: Trend break in cohort effects for mean birth weight of infants by mother's birth cohort — separately for mothers with exactly 12 and 16 years of schooling

robustness to varying age-by-year control function

	(1)	(2)	(3)	(4)
Panel A: 12 years of schooling				
Size	-0.99	-5.22	-7.32	-10.44
	(0.22)	(0.79)	(1.02)	(1.37)
Location	1947	1947	1947	1947
	[1946, 1949]	[1947, 1947]	[1947, 1947]	[1947, 1947]
P-value for existence	$< .001$	$< .001$	$< .001$	$< .001$
Panel B: 16 years of schooling				
Size	-2.15	-6.57	-6.61	-4.07
	(0.43)	(1.37)	(1.72)	(2.40)
Location	1946	1947	1947	1947
	[1944, 1948]	[1946, 1953]	[1947, 1947]	[1938, 1963]
P-value for existence	$< .001$	$< .001$	$< .001$	.002
Year FEs	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Quadratic-age-by-year	No	Yes	No	No
Cubic-age-by-year	No	No	Yes	No
Quartic-age-by-year	No	No	No	Yes

Each column shows the results of estimation of a model based on Equation 4, with the listed outcome in single age-by-year bins as the dependent variable. All models are estimated by least squares, following the approach outlined in Hansen (2000). The row titled "Size" reports the size of the trend break in cohort effects, δ, with the standard error in parentheses. The row titled "Location" reports the estimated cohort at which a trend break occurs, with a 99 $\%$ confidence region in brackets calculated by inverting the likelihood ratio statistic. The row titled "P-value for existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear. Results are based on vital statistics natality microdata, 1969-1985, mothers age 22-40, cohorts born 1938-1965.

Table A6: Evidence of cohort decline conditional on educational attainment — piecewise linear cohort effect models of intergenerational infant birth weight

controlling for year FEs and age FEs

	cohort break		
	Size	Location	Existence
	δ	λ	*p-value*
Maternal education level			
Less than HS	-1.43	1948	0.022
	(0.35)	[1945, 1950], [1958, 1962]	
High school	-0.99	1947	$< .001$
	(0.22)	[1946, 1949]	
Some college	-3.31	1948	$< .001$
	(0.28)	$[1947, 1948]$	
4 years college	-2.15	1946	$< .001$
	(0.43)	[1944, 1948]	
5+ years college	-2.42	1946	$< .001$
	(0.63)	[1941, 1948]	

Each row shows the results of estimation of a model based on Equation 4, with birth weight by mother's birth cohort as the dependent variable and the sample restricted to mothers with the listed level of education. All models are estimated by least squares, following the approach outlined in Hansen (2000). Results are based on vital statistics natality microdata, 1969-1985, mothers age 22-38, cohorts born from 1938 to 1965. The column titled "Size" reports the estimated size of the cohort break δ, with the standard error in parentheses. The column titled "Location " reports the cohort at which the trend break is estimated to occur, with a 99 $\%$ confidence interval in brackets calculated by inverting a likelihood ratio statistic. The row titled "Existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear.

Table A7: Evidence of cohort decline conditional on educational attainment — piecewise linear cohort effect models of intergenerational infant birth weight

controlling for year FEs, age FEs, and quadratic age-by-year

		cohort break	
	Size	Location	Existence
	δ	λ	*p-value*
Maternal education level			
Less than HS	-6.90	1948	0.014
	(1.69)	[1942, 1944], [1947, 1949]	
High school	-5.22	1947	0.023
	(0.79)	[1947, 1947]	
Some college	-7.19	1947	$< .001$
	(1.28)	$[1947, 1948]$	
4 years college	-6.57	1947	0.067
	(1.37)	[1946, 1953]	
5+ years college	9.52	1951	0.305
	(2.05)	[1951, 1952]	

Each row shows the results of estimation of a model based on equation 4. All models are estimated by least squares, following the approach outlined in Hansen (2000). The column titled "Size" reports the estimated size of the cohort break δ, with the standard error in parentheses. The column titled "Location " reports the cohort at which the trend break is estimated to occur, with a 99 $\%$ confidence interval in brackets calculated by inverting a likelihood ratio statistic. The row titled "Existence" reports a p-value from an F-type test, based on 1000 bootstrap samples, for the null hypothesis that no trend break occurs, ie. that cohort effects are linear. Results are based on vital statistics natality microdata, 1969-1985, mothers age 22-40, cohorts born 1938-1965.