Is the 1918 Influenza Pandemic Over?

Long-term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population

Douglas Almond

Columbia University and NBER

July 2005

Abstract

This paper studies the long-term impact of the fall 1918 Influenza Pandemic. In the 1960-1980 Decennial U.S. Census data, cohorts *in utero* during the height of the Pandemic typically display reduced educational attainment, increased rates of physical disability, lower income, lower socioeconomic status, as well as accelerated adult mortality compared with other birth cohorts. In addition, persons born in states with more severe exposure to the Pandemic experienced worse outcomes than those born in states with less severe Pandemic exposures. These results demonstrate that investments aimed at improving fetal health can have substantial long-term effects on subsequent health and economic outcomes.

Financial Support by the National Institute of Aging through Grant Number T32-AG00186 to the National Bureau of Economic Research and NIH/National Institute of Aging Grant: R03AG023939-01 is gratefully acknowledged. Comments and suggestions from Hoyt Bleakley, David Card, Ken Chay, Dora Costa, Lena Edlund, Amy Finkelstein, Michael Greenstone, Bhash Mazumder, Marie McCormick, Ellen Meara, Kevin Milligan, Tara Watson, and seminar participants at several universities are gratefully acknowledged. Timely and extensive information on current pandemic influenza risks, provided by Daniel Sharp and his associates at the Royal Institution World Science Assembly, is gratefully acknowledged. Sharon Schaff provided excellent research assistance. Data used include those provided by the National heart, Lung and Blood Institute, NIH, and DHHS from the National Longitudinal Mortality Study and do not necessarily reflect the views of the National Heart, Lung, and Blood Institute, the Bureau of the Census, or the National Center for Health Statistics. I am responsible for all errors.

This paper examines whether the Influenza Pandemic that struck the United States in the fall of 1918 had long-term impacts on the economic status and health of birth cohorts *in utero* during the Pandemic. Recent research has suggested that the effects of environmental conditions on health may be particularly strong during fetal development and that damage during this period can have lasting consequences for health over the life course (Barker 1998). The "fetal origins" hypothesis has significant implications for several areas of health economics and human capital research, including whether the rate of return to investments in fetal health may be larger than other traditional investments, such as schooling. However, much of the evidence supporting the hypothesis comes from estimated correlations between measures of infant health and adult outcomes where omitted variables bias is a real concern.

Adult health has experienced tremendous improvement in the United States over recent decades, with life expectancy increasing seven years since 1960. The sources of this improvement remain unclear. In the context of a health production function, factors can be divided into a) changes in the health endowment at birth and b) changes in health investments after birth. Both of these factors have shown tremendous gains. Measures of maternal health, and thereby the health endowments of successive cohorts, have experienced tremendous improvement in the 20th Century. The inception of the Medicare program in 1966, improvements in medical technology and pharmaceuticals would each constitute major expansions in post-birth health investments; health spending has risen from 5 percent of GDP in 1960 to around 14 percent in recent years. As these inputs move in the same direction, the empirical challenge is to disentangle which factors have driven the major improvement in adult health.

The Influenza Pandemic may help resolve this identification problem. Influenza struck without warning in the fall of 1918 and with catastrophic effect. 550,000 American died over the next few months, a casualty toll exceeding U.S. combat deaths during World War I, World War II, and the Korean and Vietnam Wars combined. Approximately fifty times this number, around twenty-five million persons in the United States, contracted the virulent influenza strain and survived. Some of the highest infection rates were observed among women of childbearing age, one third of whom contracted influenza.

The Pandemic has two distinct features, each of which reduce the biases of positively-associated health investments. First, the Pandemic struck in October of 1918 and had largely dissipated by the beginning of 1919 (Figures 1a and 1b), implying that cohorts born just months apart experienced markedly different *in utero* conditions. This presents a severe test of the fetal origins hypothesis as the design generates sharp predictions for differences in long-run outcomes among individuals born within months of each other. Second, the incidence of the Pandemic varied widely and idiosyncratically across

states: pregnant mothers in Kansas, for example, experienced more then ten times the increase in mortality rates as mothers in Wisconsin (Table 1 and Figure 2). Rather than using temporal differences between the *in utero* cohort and other birth cohorts, this second approach uses geographic variation to identify within-cohort differences in fetal exposure to the Pandemic.

These two approaches generate large estimates of the impact of health endowment changes across a range of outcomes. For males whose mothers were infected during pregnancy, disability rates are 20 percent higher (where a chronic physical disability prevents working). For workers, wages are 5 to 6 percent lower due to maternal influenza infection. Adult mortality is dramatically accelerated as well. For those who survive to age 60, death occurs approximately three years earlier. Moreover, the estimated effects are qualitatively similar across the two estimation approaches.

For differences in health investment to bias these estimates, they would have to behave in two specific ways. First, they would have to decrease discontinuously for the cohort born in the beginning of 1919 (and therefore *in utero* during the Pandemic), and then improve discontinuously for the cohort born in 1920. Second, investments would not merely have to be greater for people born in New York relative to people born in Pennsylvania, where the Pandemic was more severe. The deterioration in health investments for people born in 1919 versus 1918 in Pennsylvania would have to be larger than the corresponding change for people born in New York State. While such a pattern is possible, the geographic variation of the Influenza Pandemic requires an arguably idiosyncratic pattern of subsequent health investments in order to account for the observed outcome changes.

Two key factors exert downward bias on estimates of the effect of fetal health in this study. First, compensatory health investments by those in poor health are likely to reduce the estimated effects below the true structural parameters. The preponderance of medical spending goes to those in poor health, which would tend to counteract the effect of endowment differences. These investment differences are not accounted for in the current study. The first-order effect of cohort attrition would tend to reduce estimated effects as well. Fetal mortality increased sharply during the Pandemic. If weaker cohort members died as a result, then damage is being estimated in a positively-selected sample. The unavailability of fetal death data for the analysis period precludes adjustment for such selective attrition, thereby biasing downward the estimated effects.

Analysis is conducted using the most comprehensive data sources available on the incidence of the Influenza Pandemic and subsequent economic outcomes. The sharp timing of the Pandemic requires

precise information on the timing of birth, which the quarter of birth detail in the 1960-1980 Census microdata provides. These data can be linked to maternal and infant health conditions provided by the U.S. Vital Statistics data, using information on the state of each census respondent's birth. This is the first paper that uses this feature of the census data to link outcomes to early-life health conditions.

The remainder of the paper looks at previous research on the long-term impacts of early-life health conditions and at the challenges posed by such studies (Section I). The 1918 Pandemic is then described in greater detail, and how data can be applied to its analysis (Sections II and III). Section IV describes a conceptual framework for understanding the competing effects of changes in the newborn survival threshold versus deterioration in the distribution of early-life health. Section V outlines the empirical framework and Sections VI-VIII present the econometric results. Section IX investigates several alternative hypothesis, in particular potential bias toward finding an effect of fetal health that selective attrition in non-*in utero* cohorts could generate. Sections X presents the results, Section XI looks at the implications of these results in light of the rising likelihood of a new avian-flu pandemic, Section XII concludes.

I. Background and Previous Research

It is well known that environmental conditions affect health and mortality. This effect is believed to be strongest during the earliest periods of life, when growth is most rapid. Rather than being temporary effects that dissipate over time, it has been hypothesized that early environmental conditions have permanent effects on health. Particularly during the critical period of fetal development, the body may be "programmed" for susceptibility to disease later in life (Barker, 1998). When the fetal environment is unfavorable, a triage in the oxygen and nutrient supply is thought to occur in which the brain is given priority over other organs, such as the heart, which can suffer permanent damage as a result. These injuries may manifest themselves later in life with increased morbidity and accelerated mortality. Over recent years, the view that physiologic pathways exist between *in utero* conditions and adult health has been gaining acceptance. Discussions of fetal origins have been added to recent editions of medical textbooks (see Rudolph and Rudolph, 2003 and Winn et al., 2003) and have received increasing attention from epidemiologists and economists.

Epidemiological studies have found that low birth weight infants are at increased risk for Type 2 Diabetes, hypertension, and coronary artery disease, among other conditions. Leading work in this vein

is by Barker, who studied the association between health conditions in British localities between 1901-1910 and adult mortality rates in the same regions between 1968 and 1978. Finding a strong negative correlation between fetal nutrition proxies and subsequent mortality from heart disease, he argued for a causal link between fetal nutrition and adult health. In addition, several epidemiological studies have been made of the survivors of the "Dutch Hunger": a famine prompted by the blockade of food shipments by Nazi troops in part of the Netherlands at the end of World War II. Early analysis of this episode found no effects of fetal exposure to the famine on subsequent health (Stein, et al. 1975). Work by epidemiologists and physicians since then looked at various ages of exposure, including the fetal period, and often found effects on subsequent health outcomes, including coronary heart disease, glucose tolerance, and obesity. A major point of contention among researchers of the Dutch Famine is the appropriate measurement of early-life health, as results are highly sensitive to the particular measure used.

Economists have generally been cautious in their interpretation of statistical associations between measures of early life and adult health. Dora Costa found strong relationships between early life and adult health in both the 19th and 20th Century United States (see Costa 2000, Costa 2003, and Costa and Lahey 2003). These papers argued that the early childhood environment has an effect on life spans, and that increased longevity is in part related to improved early childhood environmental conditions. In recent years, health economists have analyzed the relationship between birth weight and subsequent adult health. For example, Case, Fertig and Paxson (2005) look at various measures of early life health for the 1946 Birth Cohort in Britain and subsequent health and socioeconomic outcomes. Currie and Moretti (2005) link natality records of siblings and children in California and analyze intergenerational correlations in birth weight.

While laboratory experiments on animals have established the causality of fetal-origins linkages, this method is of limited use for human populations for obvious ethical reasons. In addition, studies of humans invariably suffer from the potential that confounding factors, such as unobserved dimensions of family background, bias relationships estimated between early-life health measures and adult outcomes. Given the persistence of a nearly limitless set of individual-level factors, traditional regression analyses even with extensive sets of covariates are likely to find statistical associations between early- and later-life measures absent any structural relationships. Barker's raw geographic correlations almost certainly suffer from such omitted variables bias (e.g. from regional differences in average income) and existing research (even by economists) generally does not use variation in health measures over time, or variation

in early-life health where the source of identifying variation is known.[1] As Rasmussen (2001) noted, establishing causality remains a principal challenge for the "fetal origins" hypothesis. *The Lancet* echoed this call in a 2001 editorial: "An Overstretched Hypothesis?" -- and advocates the use of research designs that will present severe tests of the "fetal origins" hypothesis.

The current study attempts to heed this call by evaluating the sudden, unexpected shock to fetal health caused by the 1918 Influenza Pandemic. In addition, this paper also uses the peculiar geographic incidence of the Pandemic to evaluate long-term effects on subsequent socioeconomic outcomes, finding large long-term effects on socioeconomic outcomes from both estimation strategies. This paper is not, however, the first to look at the long-term effects of the Influenza Pandemic. Two economists, Brainerd and Siegel (2003) used the 1918 Influenza Pandemic as a shock to the size of the labor force, and look at the effects of Pandemic on subsequent economic growth. Epidemiological studies have studied the relationship between influenza exposure (not necessarily from Pandemic influenza) and the development of adult schizophrenia. A consensus on this subject has yet to emerge. Others have looked at whether the exposure of adolescents and younger adults to the 1918 Pandemic accelerated subsequent mortality (see Malemund (2003) and Reinert (2003)). The previous study that bears the greatest similarity to the present work is by Fritz Heider, was published in 1934, but since apparently overlooked.[2] Heider noted a striking pattern in the number of students enrolled in sixteen American schools for the deaf in 1933 and concluded that the effect on hearing "occurred only with children who were less than four months old at the time of the Pandemic."

II. The 1918 Influenza Pandemic

The 1918 Influenza Pandemic was an unprecedented global calamity. The Pandemic killed between 20 and 100 million persons, more people than either World War I or the Black Death of 1347-1351 (Kolata 1999: 5). It killed more Americans than all combat deaths of the 20th Century. (Corsby 1989: 207). In the United States, approximately 550,000 people were killed, causing (cross-sectional) life expectancy to drop by 12 years in 1918 (Noymer and Garenne 2000: 568). The onset of the Pandemic was very sudden: Figure 1a shows the incidence of influenza in the United States between 1911 and 1920 and the precipitous mortality spike in 1918. Moreover, this increase in mortality in 1918 was generated

[1] Two exceptions are Almond and Chay (2005), which looks at the health improvements among African-American infants during the late 1960s and Almond and Mazumder (2005), which looks at the effect of fetal exposures to the 1918 Influenza Pandemic on health outcomes using SIPP data.

[2] Public health and physician audiences of this paper have been unaware of Heider's work. It is also interesting to note that influenza is not one of the "TORCH" infections routinely screened for in newborns.

entirely by the last three months of 1918 (Figure 1b). The virus was transmitted when an infected person coughed, sneezed or spoke and thereby sent the virus into the air. When others inhaled the virus they could be infected (NCID 2003: 2). The 1918 influenza was especially contagious and spread quickly. The Pandemic diffused nationwide in about one month (Kolata 1999: 62).

A distinguishing feature of the 1918 Pandemic was also the age of those affected. While previous influenza outbreaks were most deadly for the relatively weak, the 1918 Pandemic had its largest proportionate effect on those in the prime ages of 25 to 35. This resulted in an unusual "W-shaped" age-distribution of influenza deaths, where the very young, those around age 30, and the elderly were most likely to die (Noymer and Garenne 2000: 567).

While mortality from the 1918 virus was unprecedented, the vast majority of people who became infected with the influenza virus survived. The best information on influenza infection rates comes from a house-to-house survey conducted for the U.S. Public Health Service shortly after the Pandemic. 130,248 people were canvassed in fifteen urban and rural communities; 28 percent reported being infected during the Pandemic (Jordan 1927: 189). The virulence of the 1918 strain suggests a large and negative health shock to the U.S. population. Pyle refers to the "temporary flattening or indisposition and mandatory bedrest" of one-quarter of the U.S. population, with "repeated instances of lethargy" often following bouts of influenza (Pyle 1986: 52, 41). Kolata notes that while some of influenza's victims had "a mild disease and recovered without incident," the majority subsequently developed pneumonia, requiring a "long period of convalescence" for survivors (Kolata 1999: 12). Influenza patients admitted to the University of Missouri hospital manifested "weight loss over time, afternoon fever, night sweating, and sputum." (Pyle 1986: 51).

The 1918 Pandemic appears to have had a disproportionate effect on pregnant women. For women aged 20 to 35, the infection rate was approximately 33 percent (Jordan 1927: p. 202). Figure 3a plots the influenza infection rate among females in Maryland, by age group.[3] Information on infection rates among pregnant women is more difficult to obtain. Obstetrics texts note that pregnant persons are among the most affected by influenza outbreaks (Lee et al. 2000: 745); similarly, Winn and Hobins note that influenza outbreaks have been associated with higher "morbidity and mortality in the pregnant patient than in the non-pregnant population." Crosby notes: "the lives of no group in a population afflicted by influenza are in greater jeopardy than those of pregnant women." (Crosby 1918: 207).

[3] The maternal mortality rate increased 40% in Maryland in 1918, versus 39% for all 19 states in the 1917 Birth Registration Area. Maryland infection rates are from Jordan 1927: 201.

The large deterioration in maternal health during the Pandemic led to a corresponding decline in fetal health. Figure 3b shows the trend in average stillbirth rates by month during 1918. The regular trend is interrupted in October of 1918, when stillbirth rates increased by 60 percent, or approximately 40 percent for October-December of 1918.

III. Data

This paper combines data from two sources. 1960-1980 Decennial Census Data are used to evaluate the adult outcomes of those born in the United States near the time of the 1918 Influenza Pandemic. The early-life health of these adults is gauged using annual Vital Statistics data for the United States. Mapping of adult outcomes to health conditions at birth is made possible by the reporting of state (and nation) of birth for census respondents.

Census data are useful in evaluating the long-term effects of the Influenza Pandemic for several reasons. First, the Census Bureau collects information on health-related measures, including whether the respondent had a physical disability that prevents the respondent from working.[4] Second, the census microdata provide precise information on when respondents were born. This information is important because of sharp month-of-birth discontinuities in the incidence of the Pandemic. Information on the quarter of birth of each adult respondent is recorded in three of the decennial census surveys following the 1918 Pandemic: 1960, 1970, and 1980. For this reason, analysis of the Pandemic is restricted to these three census years.[5] Thirdly, the large sample size allows comparisons within a narrowly-defined birth interval: a year before and a year after the Influenza Pandemic, and therefore among those who would tend to share relatively similar life-course experiences. In 1960, a 1 percent sample is available. Combining the state, metro, and neighborhood samples generates a 3 percent sample for 1970. The 5 percent sample available in 1980 enables comparisons across quarters of birth in addition to birth years.

Information on early-life health conditions is provided by annual volumes of the Census Bureau's *Mortality Statistics* and *Birth Statistics*. Information on the incidence of influenza infection is not available; influenza was not made a reportable disease in the United States until the Pandemic was

[4] While the Census also records whether a physical disability limits the amount of work that can be performed, this measure is excluded from the analysis due to discrepancies in this measure between the University of Minnesota's IPUMS datasets (used in this paper) and that available through the Econometrics Laboratory Software Archive at the University of California, Berkeley.

[5] Quarter of birth is required to calculate year of birth as well. For those born between April and December, year of birth is calculated as: survey year-age-1. For respondents born in the first quarter, year of birth is: survey year-age.

underway (Crosby 1989: 56). Thus, influenza infection rates by state are unavailable for this period, let alone information on whether a particular individual was impacted by maternal influenza infection. Instead, the number of deaths from various causes is available. Mortality data include the number of infant deaths (deaths in the first year of life) and maternal deaths (deaths related to pregnancy and delivery) for each state and year. For a subset of mortality outcomes, the state-level data are provided by month of death. Information available by month includes the total number of influenza and pneumonia deaths and information on the number of stillbirths (1918 only). Finally, the annual *Birth Statistics* volumes provide the number of births by state, year, and gender.

The timing of the Pandemic is fortunate from a data perspective. The collection of birth statistics by the federal government began in 1915, three and a half years before the Pandemic. But as collection was at its early stages, not all states provided data. For the 1917-1920 period, only data for nineteen states and the District of Columbia are available. Table 1 lists these states. Slightly over half of the U.S.-born population in 1960 was born in one of these twenty "states" with vital statistics data.

IV. Conceptual Framework

This analysis explores the long-term effects of changes in early-life health. Unfortunately, the early-life health of individuals is not observable, nor is it observed for an individual's birth cohort at large. Instead, only information on the early-life mortality rates to which a birth cohort was exposed is available. It is therefore useful to consider how early-life mortality rates are related to cohort health and how this might affect the subsequent empirical work. A framework is developed below that distinguishes between two factors that can determine early-life mortality.[6]

Infant mortality rates for a given birth cohort reflect two distinct pieces of information a) the unobserved distribution of initial cohort health and b) the health threshold which must be exceeded in order for newborns to survive infancy. During the Influenza Pandemic, it is likely that both of these factors changed. In particular, it is presumed that the unobserved distribution of health deteriorated and the infant death threshold became more selective.

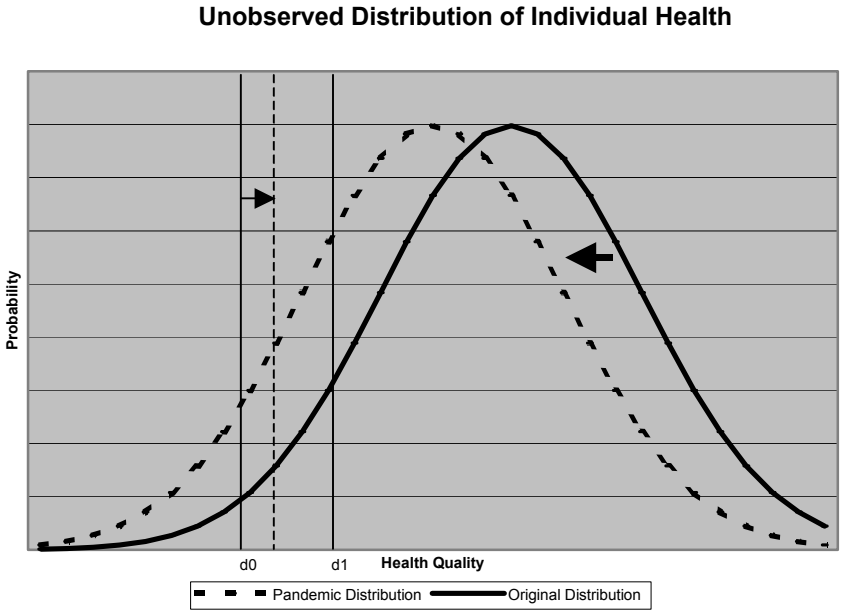
[6] The framework described below can be applied to various measures of early-life mortality, including the infant mortality rate, fetal death rate, and the maternal mortality rate. For exposition, these rates are referred to collectively as the infant mortality rate (defined as the number of deaths within the first year of life per 1,000 live births).

The primary hypothesized effect of the Pandemic is to shift the health distribution. The influenza infection of pregnant mothers may have caused the health of the cohort *in utero* to deteriorate, if, for example, the oxygen supply to the fetus was restricted when the mother contracted influenza or a secondary pneumonia infection. Such a shift in the unobserved distribution of initial cohort health would also generate changes in early-life mortality rates. More infants would fall below the threshold at which infant death occurs, and infant mortality rates would increase.

On the other hand, infant mortality rates may increase when the infant death threshold becomes more restrictive. For infants in "marginal" health, the Pandemic caused their death without altering their unobserved health index. This is possible if, for example, access to medical care deteriorated; if physicians were busy treating influenza patients, an infant in marginal health might have gone without medical care and died as a result.

While both of these factors will cause the infant mortality rate to increase, their implications for cohort health are polar. If infant mortality rates increased because the initial health distribution deteriorated, and if this distributional shift was persistent (as the "fetal origins" hypothesis predicts), then this cohort will be observed to be in worse health later in life. Albeit implicit, changes in the underlying health distribution are generally the focus of empirical work on long-term health linkages. Changes in the health threshold, by definition, have permanent effects. If the infant mortality rate is high because more infants of marginal initial health are dying, infants that survive infancy will be especially healthy. To the extent that this health threshold effect is at play, we would expect that cohorts exposed to high infant mortality rates to be more positively selected and therefore in better subsequent health.

The tension that exists between selective attrition and changes in underlying health can be considered more formally in a stylized latent variable model of initial health. Let h^*_i be the unobserved health of individual i which is fixed from birth. In the figure below, the probability distribution of h^*_i is given by the solid black line, with individuals in poor initial health being on the left and healthier individuals on the right. If h^*_i falls below a survival threshold d_0 (depicted in the figure by the leftmost vertical line), then the individual will die within the first year of life. Individuals with $h^*_i \geq d_0$ survive to adulthood. These adults will be physically disabled during the follow-up period if $d_0 \leq h^*_i < d_1$, that is if their initial health falls between the two vertical black lines in the figure. Individuals suffer neither death nor disability if h^*_i exceeds d_1.



Given these health thresholds, the infant mortality rate (IMR) may be defined using the cumulative distribution function $F(h^*_i)$ as:

$$IMR \equiv F(d_0)$$

That is, the infant mortality rate is given by the share of the health distribution to the left of d_0. The adult disability rate (ADR) is given by the share of persons surviving infancy that have initial health below d_1:

$$ADR \equiv (F(d_1) - F(d_0))/(1 - F(d_0))$$

Deterioration in the probability distribution for health at birth, $f(h^*_i)$ (depicted in the figure above as an decrease in the mean μ of the solid black distribution to the new dotted distribution) generates increases in both the early-life mortality rate and the adult disability rate. Therefore:

$$\frac{\frac{\partial ADR}{\partial \mu}}{\frac{\partial IMR}{\partial \mu}} > 0$$

The adult disability and infant mortality rates will move in the same direction when shifts in the probability distribution of unobserved health occur.

If influenza infection causes those whose initial health just exceeded the infant survival threshold d_0 to die, then the increase in the infant mortality rate may be affected by rightward shifts in d_0. As:

$$\frac{\frac{\partial ADR}{\partial d_0}}{\frac{\partial IMR}{\partial d_0}} < 0,$$

rightward shifts in d_0 (for a fixed d_1) will cause the expected value of health for adults at the follow-up period to increase when the infant mortality rate rises. To the extent that infant death during the Pandemic occurred among those in the weakest initial health, adult disability should *decrease* as a result.

Infant mortality, along with other early-life mortality rates, increased substantially during the Influenza Pandemic. These increases were presumably caused by some combination of rightward shifts in the mortality threshold and leftward shifts in the health distribution. For the effect of distributional shifts (the focus of the present study) to be apparent in analyses that use early-life mortality rates, they must overwhelm the effects of changes in the survival threshold. In this respect, the estimates of long-term damage caused by the Pandemic are biased downward.

V. Empirical Methodology

The data permit two distinct approaches for estimating the "fetal origins" effect of the 1918 Influenza Pandemic. The first approach compares the adult outcomes of cohorts *in utero* during the Pandemic with cohorts born shortly before and shortly after the Pandemic. The second approach uses variation in influenza severity by state to estimate long-term effects within birth cohort.

Cohort Analysis

The cohort analysis utilizes the discontinuous change in health that occurred at the end of 1918. Eighty-five percent of deaths during the Influenza Pandemic occurred between October of 1918 and January of 1919. The majority of those *in utero* during the Pandemic would have been born in 1919 and all of those born in the first six months of 1919 (and carried to term) would have been *in utero* during the peak of the Pandemic. The 1919 birth cohort, and in particular those born in the first two quarters of 1919, should have experienced the greatest influenza-induced changes to fetal health.

In the simplest specification, the cohort analysis estimates the departure in outcomes for the 1919 birth cohort from a linear trend in birth year:

$$outcome_i = \beta_0 + \beta_1(yob) + \beta_2(dummy_{yob = 1919}) + \varepsilon_i$$

where *yob* is the birth year. Certain cohort outcomes are non-linear in birth year, particularly as the influenza cohorts near retirement. Quadratic and cubic polynomials trends in birth year are fit for these outcomes. The analysis sample is restricted to respondents born in the United States, and therefore exposed to the Pandemic at nearly the same time. The 5% 1980 Census sample permits comparisons in outcomes across quarters of birth.

Fixed-Effects Analysis

The fixed effects analysis includes dummy variable for both state of birth and birth year, thereby identifying effects by geographic variation in the severity of the Pandemic. An enduring mystery of the 1918 Pandemic is why certain regions and towns were more affected than others. Jordan's extensive survey of the Pandemic notes that: "no consistent relation between geographic location and influenza attack rate appeared to exist" in the United States (Jordan 1927: 212). Pennsylvania experienced a markedly worse Pandemic than Maryland or New York and Kansas a much worse Pandemic than Wisconsin. Brainerd and Siegel (2003) note that "influenza and pneumonia deaths are nearly orthogonal to all of the other explanatory variables…." Variation within states also appeared idiosyncratic: St. Paul had a ten-week death rate fifty percent higher than neighboring Minneapolis. The regional variation in the Pandemic did not display any particular geographic pattern.[7]

Table 1 indicates that Kansas had ten times the increase in maternal mortality[8] as Wisconsin. Therefore, the maternal mortality rate is used as the proxy for changes in fetal health induced by the Pandemic. The apparent arbitrariness of the Pandemic's incidence reduces concern about correlation of within-cohort measures of influenza exposure with unobserved factors that are also correlated with long-term outcomes.

Because the effect of maternal health on *in utero* health is thought to be most important, the maternal mortality zero to nine months prior to birth is most relevant. As only annual measures of maternal mortality by state are available, the rate in the year prior to birth is applied. As nearly all of the 1917-1919 variation in the maternal mortality rate is generated by the Influenza Pandemic, the fact that

[7] One possibility is that the 1918 flu virus traced the path of the 1890 flu pandemic. Suggestive evidence of this possibility comes from a) the unusual age-distribution of mortality among young adults (those 28 and older in 1918 being unexposed to the 1890 flu) and b) statistically-significant correlations of mortality rates across Massachusetts counties between 1890 and 1918 (results available from author).

[8] As noted above, direct measures of fetal exposure to influenza were not available for the analysis period and fetal death data are available only for 1918.Reporting of stillbirth data began in 1918 but did not resume until 1922.

the Pandemic occurred at the end of 1918 reduces the effective lagging in the fixed effects analysis below one year.

The primary fixed-effects regression specification is:

$$outcome_i = \beta_0 + \beta_1(MMR_{yob_i - 1, s_i}) + \beta_2(Dummy_{yob_i}) + \beta_3(Dummy_{s_i}) + \beta_3(Dummy_{notwhite_i}) + \varepsilon_i$$

where *MMR* is maternal mortality rate, *S* is state of birth, *YOB* is year of birth, and *notwhite* is a binary categorization for race. Inclusion of dummies for each state of birth absorbs permanent differences across states of birth, including, for example, the quality of reporting of maternal deaths. Standard errors are clustered at the state-of- birth/year-of-birth level.

Several shortcomings of the proxy for fetal health within cohort are apparent. First, maternal mortality was relatively rare, even in 1918. Less than 1 percent of mothers died due to childbearing. It is possible that changes in the tail of the maternal-health distribution do not reflect well the changes in fetal health for those whose mothers contracted influenza. Second, maternal death would tend to occur near the end of pregnancy. In contrast, injury to fetal health caused by maternal influenza infection could occur throughout the term of pregnancy. Both of these factors reduce correspondence of the proxy measure with maternal influenza infection during pregnancy.

As noted above in Section I, Heider found that the infection of *infants* with pandemic influenza caused deafness, as observed in 1933 among teenagers. The possibility that changes in infant in addition to fetal health might have permanent effects on subsequent outcomes can be evaluated by including the infant mortality rate (IMR)[9] in the fixed-effects regression. This will also eliminate bias in the estimated β_1 coefficient resulting from correlation between the infant and maternal mortality rates. The following specification is estimated:

$$outcome_i = \beta_0 + \beta_1(MMR_{yob_i - 1, s_i}) + \beta_2(IMR_{yob_i, s_i}) + \beta_3(Dummy_{yob_i}) + \beta_4(Dummy_{s_i}) + \beta_5(Dummy_{notwhite_i}) + \varepsilon_i$$

VI. 1960-1980 Census Outcomes by Year and Quarter of Birth

[9] The infant mortality rate available in the Vital Statistics is the *period* mortality rate rather than the *cohort* mortality rate. That is, all infant deaths in a given year, regardless of whether that infant was born in the current year or the previous year are included in the infant mortality calculated on a period basis. In this respect, it is fortunate that the Pandemic occurred toward the end of 1918, so that the slippage between the cohort and period mortality rates is minimized. However, the infant mortality rate to which these birth cohorts were exposed is measured with error.

The average Census outcomes of the 1919 birth cohort cannot be predicted by the outcomes of adjacent birth cohorts. This section demonstrates this deviation both graphically and statistically. In three separate decennial Census surveys, average educational attainment, income, socioeconomic status, and disability status of the birth cohort most likely to be *in utero* during the Pandemic are substantially compromised.

Educational attainment measured in the 1960, 1970, and 1980 Census microdata provide consistent evidence of compromised outcomes for the 1919 birth cohort. Figure 4a shows that cohorts born five years after the Pandemic received approximately one more year of schooling than cohorts born five years before the Pandemic. This regular age trend is interrupted for the 1919 year-of-birth cohort, which actually had lower average education than the 1918 birth cohort. This would imply a $1/10^{th}$ year decrease in schooling for the 1919 birth cohort due to the Pandemic, or a 1% decrease. As the house-to-house survey data indicated that approximately one third of women of childbearing age contracted influenza, no effect is presumed to exist for approximately two-thirds of the sample. This would imply that for those whose mothers were infected, average education fell 3%. (Deviation of the 1919 birth cohort mean education from a linear age trend is significant at the 1 percent level, and is not affected by dummies for quarter of birth.)

Figure 4b, which plots schooling levels in 1970, highlights two additional points. First, the similarity in patterns for both males and females born around the time of the Pandemic imply that these patterns are not a result of military service in World War II.[10] Second, a key margin along which education was lowered for the 1919 birth was completion of high school. For men, the time-series pattern would imply a 3 percent decrease in the likelihood of high school graduation. For those men whose mother contracted influenza, the likelihood of high school graduation fell approximately 9 percent.

Figures 5a and 5b plot average education and average income in 1980 by year of birth. Again, average educational attainment of the 1919 birth cohort is off trend. Total personal income displays a more curved pattern in birth year around age 60, and a more muted decrease for the 1919 birth cohort. With a quadratic in year of birth, income is estimated to be about $250 lower on average for the 1919 birth cohort, or about 1.4% below the age trend.[11] The average Duncan Socioeconomic Index, which

[10] Moreover, it is apparent from veteran status measured in 1960 that service in World War II was substantially more likely for cohorts born in the early 1920s. Regression adjustment for veteran status does not substantially affect the results of this section.

[11] An upper-bound estimate of the return to education from these numbers would thus be 14%. The income change for the 1919 birth cohort is also significant at the 1% level.

reflects occupation choice, has a similarly smooth pattern in year of birth, with the exception of the 1919 birth cohort (figure not shown). The index decreases 2.4% for those born in 1919 (relative to a quadratic in year of birth), which is significant at the 1% level.

The potential bias of selective attrition across birth cohorts can also be assessed through cohort trends in educational attainment, income, and socioeconomic status. Were comparisons between the 1919 and 1918 birth cohort biased by higher early-life attrition among weaker infants born in 1918 and killed by influenza, we would expect the surviving members of the 1918 birth cohort to have higher educational attainment than the steady secular trend would predict. The education, income, and socioeconomic trends do not support this hypothesis. 1918 appears to be on trend with previous birth years. Similarly, if the 1920 birth cohort had better subsequent outcomes because the set of potential mothers and fathers had been culled by the Influenza Pandemic, outcomes for the 1920 birth cohort should also deviate from the age trend, which is not observed. (These potential biases are explored in greater detail in Section IX.)

In addition to "updating" outcome differences among the influenza cohorts later in the life cycle, the 1980 Census data permit finer comparisons within birth year due to the larger sample size. The precise predictions made by the abrupt timing of the Influenza Pandemic can be evaluated. If indeed cohorts *in utero* during the Pandemic suffered the greatest injury, then outcomes for the first two quarters of 1919 should be impacted.

This prediction is substantiated in the differences in high school graduation rates by quarter of birth. Figure 6a plots the unadjusted graduation rates by birth cohort. Seasonality in graduation rates implies that those born between January and March are less likely to complete high school.[12] Adjustment for seasonal variation with quarter of birth dummies and a linear age effect (estimated using the 1912-1915 and 1921-1923 years of birth) yields Figure 6b, where a pronounced drop in the likelihood of high school graduation occurs for those born in the first quarter of 1919. Departure from trend for each of the first two quarters of 1919 is significant at the 5% level.

The same discontinuous pattern in quarter of birth is observed for the presence of a chronic physical disability that prevents the respondent from working at a job.[13] Disability rates are significantly

[12] See Angrist and Krueger, 1991, who use quarter of birth as an instrument for educational attainment.
[13] Disability rates are plotted beginning with the third quarter of 1917 in Figure 1. Disability rates before this interval appear to depart from the age trend in late 1916 and early 1917. This pattern is of (separate) interest because it occurs around the Social Security "Notch" – i.e. cohorts born in the fourth quarter of 1916 versus those born in the first quarter of 1917. For a detailed description of the Notch see, for example, Evans and Snyder, 2003.

lower for younger men but this trend is interrupted for the first three quarters of birth of 1919 (Figure 7).[14] Rates for these cohorts are roughly one percentage point higher than a linear age trend would predict (and significant at the 1% threshold). As approximately 15 percent of men born in these years had such a disability, the increase in disability for these cohorts is estimated at 6.5 percent. For men whose mothers contracted influenza, the estimated increase in disability risk is therefore twenty percent. The greater magnitude of this estimated effect suggests that as finer comparisons are made – isolating precise birth cohorts most likely to be affected and focusing on more health-related measures where the long-term health linkage should be more direct – the magnitude of estimated linkages may increase.

VII. Fixed Effect Regression Results in 1960 Census Microdata

The fixed-effects regression results provide three new pieces of information. First, they demonstrate that a distinct component of the variation in health caused by the Pandemic – the arguably arbitrary geographic variation – operates in the same direction as the secular effect; cohorts with greater estimated exposure to the Pandemic *in utero* display markedly worse subsequent outcomes. Secondly, the regressions provide an elasticity – the estimated relationship between a change in the maternal mortality rate in the year preceding the birth year and various long-term outcomes. The elasticities estimated are large and generally statistically significant. Finally, the regressions permit separation of the effect of changes in infant health from changes in fetal health. Results indicate that while both fetal and infant health changes have long-term effects, the effects of fetal health are stronger. This is consistent with the results of Section VI in that the 1919 birth cohort experienced that largest long-term effects of the Pandemic.

Results

The 1960 regression sample includes men born in 1918, 1919 or 1920 in the nineteen states with complete vital statistics data for 1917 forward.[15] Approximately 2.5 percent of records had age allocated by the Census. These records were dropped; the resulting 1960 analysis sample has 16,566 records. As Table 2 indicates, 1960 Census respondents born during this interval are around 40 years old. Slightly over half completed high school, 88 percent were married, and nearly four percent were not in the labor force. The states of birth tended to be in the northeast, the mid-atlantic, and east north central regions and

[14] The incidence of physical disabilities preventing use of public transportation is collected, but occurs less frequently (among approximately 5% of 1980 respondents) and displays noisy patterns in quarter of birth. Rates for the second quarter of birth of 1919, however, are the highest for the 1918-1923 interval.
[15] 1917 data is required for the measure of fetal health conditions experienced by the 1918 birth cohort.

comprise approximately half of all births in the United States. These areas were almost exclusively white around the time of the Pandemic – only 5 percent of persons born in these nineteen states (and sampled in 1960) were non-whites. The infant mortality rate experienced by this cohort was quite high by modern standards: 9.1 percent died in the first year of life. Approximately 6.4 percent died between age one and follow-up in 1960.[16]

The first column of Table 3 presents regression results for total personal income among men in 1960 who were born between 1918 and 1920. Income is regressed on the maternal mortality rate in the state and year of birth preceding the respondents' year of birth and a constant. Log income is observed to be lower in 1960 where the maternal mortality rate was higher in the "*in utero*" year.

The second column of Table 3 adds a dummy variable for the 1919 and 1920 years of birth. As income is increasing in age for those around age 40, the year of birth dummies register the age effect. In the column two specification, maternal mortality is free to register the fact that 1919 was the "bad" birth year – the point estimate for maternal mortality doubles and is now significant at the 5% threshold. Higher maternal mortality rates of the 1919 birth cohorts are negatively associated with income forty years later. The 0.26 unit increase in the maternal mortality rate during the Pandemic is estimated to have caused income to fall nearly 20 percent for the 1919 birth cohort.

However, key variables are omitted from the regression. The maternal mortality rate will also be registering the effect of being born in different states. For example, males born in North Carolina will tend to have experienced higher maternal mortality rates – North Carolina had the highest average maternal mortality from 1917 to 1919 among states in the Birth Registration Area. Males born in North Carolina are also likely to have lower subsequent income for reasons not directly related to health – for example persons born in North Carolina were likely to have parents that received less schooling, and therefore received less schooling themselves. Even absent an effect of maternal health on subsequent income, we would expect to estimate a negative coefficient on the maternal mortality rate due to the omitted characteristics, such as those of respondents born in North Carolina. This specification demonstrates the predisposition of many analyses to find large and significant relationships between measures of early childhood health and subsequent outcomes.

The first column of Table 4 adds dummies for each state of birth to the Table 3 specifications and therefore using purely the idiosyncratic variation in the Pandemic – that is, deviation from the average

[16] 15.5% (total attrition) minus 9.1% (infant mortality rate).

maternal mortality rate spike. For total income in 1960, each one-unit change in the maternal mortality rate leads to as 21 percent change in income. This estimate is approximately one third as large as the estimate from Table 3, which is likely biased upward by omitted state-level variables correlated with the "treatment" variable.

During the Pandemic, average maternal mortality increased from 0.66 deaths per 100 live births to 0.92 deaths in the 19 states with vital statistics data. This implies that income fell six percent for the 1919 birth cohort due to the Pandemic. However, as it is likely that most respondents did not have mothers who were infected with influenza in 1918, the 33 percent estimated infection rate would imply that a third of the children are generating the 6 percent decline in income, or that the Pandemic caused a 18 percent income decline for those whose fetal development was impacted. This is obviously a large effect, especially if it is one that persists over the life course, as the time-series graphs would indicate. For example, avoiding Pandemic infection *in utero* would have been worth more than a year of schooling in terms of its effect on income.

Column 2 of Table 4 adds the infant mortality rate to the regression. The infant mortality rate enters negatively: higher infant mortality rates are associated with decreases in subsequent income (significant at the 5% level). The magnitude of the effect of infant mortality is substantially smaller than that estimated for the maternal mortality rate. As infant mortality increased approximately 1 percentage point between 1917 and 1918, this would imply a 2 percent decrease in income due to the Pandemic (less than a third of the estimated effect of the change in fetal health, as measured by the maternal mortality rate). Note however, this effect is estimated for the 1918 birth cohort that experienced the brunt of the infant mortality increase. Meanwhile, the magnitude of the estimated fetal health effect has increased in the Column 2 specification, and is now significant at the 1% threshold.

Column 3 of Table 4 adds a set of dummies for the state in which the respondent lived in 1960. State of residence is strongly associated with the state of birth – two-thirds of the analysis sample lived in their state of birth. 1960 income differences across states could be contributing to the estimated relationship between maternal mortality and subsequent income. However, the pattern of income differences would have to mimic idiosyncratic variation in the Pandemic to bias estimates of the effect of fetal health. On the other hand, the moving decision may well be endogenous to health, in which case it is not obvious that 1960 residence should be included in the estimation. Column 3 adds dummy variables for the state of residence of each respondent in 1960. If this is done, the estimated coefficient on maternal mortality rates falls slightly. Income differences associated with state of residence in 1960 do not account

for estimated negative relationship between maternal mortality and subsequent income. Inclusion of additional indicators for where the respondent lived in 1960 – a center city, metropolitan area, urban area, or rural area – reduces the estimated coefficient for the maternal mortality rate slightly more, but still does not account for the estimated importance of the fetal health measure.[17]

Table 5 looks at educational attainment in 1960 and its relationship to fetal health. After controlling for state of birth and the inferior outcomes of the 1919 birth cohort, years of schooling are negatively related to the idiosyncratic variation in the maternal mortality rate. For those whose mothers contracted influenza, educational attainment is estimated to be approximately two-thirds of a year less as a result (six percent). The infant mortality rate (Column 2) again enters negatively, consistent with a negative long-term effect of infant influenza infection. The estimated effect of the Pandemic on fetal health is now estimated to have decreased educational attainment for those whose mothers were infected by three-quarters of a year (or 7 percent). Inclusion the indicators for state of residence reduces this coefficient slightly.

Table 6 looks at the likelihood of high school graduation. This is of interest in helping to understand what points in the educational attainment distribution were affected by the Pandemic and in view of results from the cohort analysis finding larger effects for high school graduation. Results in column 1 indicate that the likelihood of graduation fell 13 percentage points for each 1 percent rise in the maternal mortality rate. On a 53 percent graduation rate for these cohorts, the Pandemic in aggregate lowered the likelihood of high school graduation by 6 percent in the 1919 birth cohort. For those whose mother contracted influenza, the chances of graduating high school fell approximately 19 percent. Inclusion of the infant mortality rate increases the estimated effect slightly, though the infant mortality rate itself does not enter significantly.

Results in Table 7 indicate that men were less likely to be in the labor force in 1960 if they experienced a higher maternal mortality rate in their state of birth.[18] The estimated coefficient implies that participation in the labor force fell just over one percentage point due to the Pandemic. On the overall rate of 3.6 percent for men born 1918-1920, this implies a 35% change in the likelihood of being out of the labor force. For those whose mother was infected, this would imply labor force participation

[17] Results presented in Columns 4 and 5 will be discussed in Section IV.
[18] Questions on physical disability and its effect on ability to work were not asked in 1960.

was determined by the Influenza Pandemic.[19] This estimate does not change in the column 2 and column 3 specifications.

Table 8 looks at the likelihood of being in poverty in 1960 (measured at 150% of the poverty line). In column 1, the maternal mortality rate enters positively but is not significant at conventional levels. Inclusion of the infant mortality rate increases the point estimate and reduces the sampling error. Both it and the infant mortality rate are associated with increased likelihood of being in poverty in 1960, with the maternal mortality rate coefficient significant at the 10% level. On the 24 percent poverty rate, changes in fetal health are estimated to have increased the likelihood of poverty by 7 percent, or 21 percent for those impacted *in utero*. A similar pattern is observed in Table 9, which looks at the determinants of socioeconomic status, as measured by the Duncan Socioeconomic Index. Declines in fetal health are associated with decreases in socioeconomic position in 1960, but this effect is not statistically significant. Regression results for the 1960 poverty rate and socioeconomic index underscore the fact that the parameter of interest was not benefiting from correlation with changes in infant health.

The estimates reported above indicate that maternal health during pregnancy exerts an effect on a range of subsequent outcomes for offspring. The strength of the effect estimated in the preceding tables, as well as the fact that it is an intergenerational one, raises the question of whether the second generation after the Pandemic might also have been impacted, i.e. children of parents whose mothers were infected with influenza. The most direct means of estimating effects for the second generation involves fertility. The 1960 Census asks women the number of children they have ever given birth to. In 1960 women exposed to the Influenza Pandemic would be around age 40, and therefore their childbearing generally complete.

Table 10 estimates whether women's fertility may have responded to *in utero* influenza exposure (using the same specifications as above). Results indicate that fewer children were born to mothers exposed to the Pandemic. For the 1919 birth cohort, the Pandemic is estimated to have reduced fertility by 0.6 percent, or 2 percent in those infected (significant at the 5% level). Inclusion of the infant mortality rate reduces this estimate slightly.

[19] Results estimated using a logit model are quite similar (with lower p-values).

VIII. Fixed Effect Regression Results in 1980[20] Census Microdata

By age 60, approximately thirty percent of the Pandemic birth cohorts have passed away. This attrition, to the extent that it was concentrated among those exposed to the Pandemic, could cause the effect on survivors in 1980 to be reduced or even eliminated. But as seen in the cohort analysis, effects of iearly-life exposure to the Influenza Pandemic persist past age 60.

In 1980, approximately one-quarter of men born between 1918 and 1920 were no longer in the labor force. Table 12 present results from the 1980 Census data for the likelihood of being out of the labor force. Results indicate that greater exposure to the Influenza Pandemic makes it substantially more likely that one is no longer in the labor force at age 60, after accounting for birth year and birth state. Labor force participation falls 1.5 percentage points due to the Pandemic (significant at the 1% threshold), when the maternal mortality rate increased 40%, or 0.26 units. For those whose mothers contracted influenza, the estimated effect is around 4.5 percentage points, or about 17 percent. Adjustment for infant health changes strengthens the estimated relationship slightly, as does inclusion of 1980 state of residence.

The 1980 Census is appealing for analysis of long-term health effects because it includes a measure more directly related to health – whether the respondent is prevented from working at a job because of a physical disability lasting six months or longer. Approximately fourteen percent of male respondents said a physical disability prevented them from working. Again, deterioration of maternal health at the time of the Pandemic is associated with her children's subsequent disability status (Table 13). The 0.26 percentage point increase in maternal mortality during the Pandemic is estimated to increase the likelihood of a physical disability 1 percentage point in the 1919 birth cohort (6 percent), and 2.5 percentage points (or 17 percent) for those whose mother was infected. Adjustment for infant health and state of residence strengthens this estimated relationship marginally, but neither regressor enters significantly in either Column 2 or Column 3 specifications.

Total Personal Income in 1980 also appears to be affected by *in utero* health conditions (Table 13). With state and year of birth dummies, income is estimated to be 2.5 percent lower for the 1919 birth cohort due to the 1918 Pandemic. For those whose mother was infected with influenza, income is

[20] The 1970 Census fixed effect results are not consistent actor the state, city, and metro samples. As the fixed-effects specification is identified using year of birth interacted with state of birth, the fact the 1970 Census alone requires combination of samplings at different strata of geography could somehow vitiate proper estimation. In contrast, the year-of-birth results using the 1970 Census data are consistent with the 1960 and 1980 Census data.

estimated to be 7.5% lower. Infant mortality does not enter significantly and inclusion of controls for 1980 state of residence increases the point estimate for the maternal mortality rate coefficient, which is significant at the 10% threshold.

Table 15 looks at socioeconomic status in 1980, using the Duncan Socioeconomic Index of occupation categories. Column 1 finds a negative coefficient on the maternal mortality rate in the state and year preceding birth, indicating a negative effect on subsequent occupation. Inclusion of the infant mortality rate in column 2 increases the point estimate in absolute value to –2.7. For sons of influenza-infected mothers the occupation index is estimated to have fallen 2 points or approximately 6 percent. State-level differences in infant mortality rate changes indicate a negative effect of infant health on the socioeconomic index as well. The 1 percentage point increase in the infant mortality rate is estimated to decrease socioeconomic status by half the amount as indicated by the change in maternal mortality. Inclusion of state of residence information strengthens the estimated effect of both early-life factors, the maternal mortality rate in particular.

Table 16 shows the relationship between poverty status in 1980 and maternal mortality in the state and year preceding birth, but this is not significantly different from no effect. Inclusion of the infant mortality rate increases the magnitude of the maternal mortality rate point estimate and decreases the sampling error and the estimate is not significant at conventional levels. A similar pattern is apparent in Table 17, which looks at years of education as reported in 1980. The parameter of interest increases with additional control, but the null hypothesis of zero effect cannot be rejected at conventional levels until 1980 state of residence dummies are included. Infant mortality, in contrast, enters negatively and is precisely estimated (significant at the 1% threshold). However, the effect is not large – the infant mortality variation induced by the Pandemic is estimated to have decreased educational attainment by one-fifth of a year for those who were actually infected, comparable to the effect estimated for maternal infection. Adjustment for state of residence increases the maternal mortality point estimate; it is now significant at the 10 percent threshold. Point estimates scaled by the Pandemic changes imply that fetal health exerts a larger effect on educational attainment than infant health changes.

IX. Alternative Explanations

This section looks at five alternative explanations to the finding of large long-term effects of *in utero* influenza exposure. The primary threat concerns the selective attrition of birth cohorts, i.e. that deaths caused by the Influenza Pandemic account for the outcome differences. This alternative

explanation is investigated in detail: the effect of selective mortality during the course of the Pandemic is considered, followed by the effect of selective mortality occurring after the Pandemic but prior to follow-up. It is demonstrated that neither of these attrition-related explanations, nor three less vexing alternative explanations, account for the observed findings.

Alternative Explanation #1: Selective Attrition During the Pandemic

The sharp increase in the death rate in late 1918 generates three distinct types of attrition that could potentially bias estimates of the long-term effect of influenza infection on people born around the time of the Pandemic. The three potential sources of attrition bias result from increases in deaths in the fall of 1918 among three age groups: 1) infants 2) the *in utero* cohort (i.e. stillbirths) and 3) the set of mothers and fathers who would potentially conceive a child after the Pandemic. Deaths in these age groups could affect subsequent outcomes in the 1918, 1919, and 1920 birth cohorts, respectively. For each birth cohort, these biases all work in the same direction, causing subsequent outcomes to appear stronger. For three distinct reasons, it is argued that the potential biases caused by selective cohort attrition do not undermine the primary conclusion of this research.

How Selective Attrition Could Bias Comparisons Across Birth Cohorts

The increase in infant mortality caused by the Influenza Pandemic was largest for the 1918 birth cohort. Nevertheless, subsequent observations on each cohort born near the time of the Influenza Pandemic are potentially subject to the attrition bias described in the preceding section. The means by which each cohort may have experienced attrition bias will be described in turn, beginning with the 1918 birth cohort.

Approximately 13,000 infants, or slightly over 1 percent of infants under age one, died from influenza in 1918.[21] Since the Pandemic arrived in October of 1918 in the United States, this implies a disproportionate share of the 1918 birth cohort died in the first year of life due to influenza. In 1919, the number of influenza deaths fell 40 percent to 8,000. The greater likelihood of infants born in 1918 to die from the Pandemic implies that surviving members of the 1918 birth cohort may have experienced greater selective attrition due to the Pandemic than the 1919 birth cohort. If the model outlined above applies – that is if infant death due to the Pandemic was more likely to occur among infants in very poor health

[21] In the 1918 death registration area. Some death results from the Influneza Pandemic were also likely coded as pneumonia, which increased by approximately 3,000 deaths over 1917.

prior to infection, then the 1918 birth cohort could display better subsequent health due to the selective effect of the Pandemic, i.e. the absence of weaker cohort members.

Similarly, the 1919 birth cohort may have been subject to selective attrition that makes the cohort appear relatively strong in subsequent observation. Most of the 1919 birth cohort was *in utero* at the end of 1918 when the Influenza Pandemic struck. Data indicate that stillbirths increased approximately 40 percent in the last three months of 1918 due to the Pandemic. If these additional stillbirths were concentrated among fetuses in marginal initial health, the Pandemic generated selection into live births in 1919. If this marginal initial health was to have been positively correlated with subsequent health, then the 1919 birth cohort also experienced selective attrition causing subsequent observations to appear more favorable than had this attrition not occurred.

Finally, the 1920 birth cohort may have experienced selective attrition that causes subsequent outcomes to appear stronger. While this cohort was both conceived and born after the Pandemic had largely concluded, the parents of those born in 1920 were exposed to the Pandemic. Attrition among the set of potential parents in late 1918 could have been concentrated among those in marginal health. If these marginal parents would have had children of similarly marginal initial health, subsequent outcomes of the 1920 birth cohort could be biased upward by attrition of potential parents during the Pandemic.

The effect of selective attrition on cohorts born around the time of the Pandemic is relevant because we would like to compare the long-term outcomes of these three groups. In particular, we would like to compare the long-term outcomes of the *in utero* cohort with outcomes of persons of similar age whose fetal health was not directly impacted. The natural comparison groups are those born immediately before (1918) and after the conclusion of the Pandemic (1920). If these two birth cohorts manifest improved subsequent outcomes due to attrition to a greater degree than the 1919 birth cohort, we could be led to conclude that the *in utero* cohort is in worse subsequent health absent any effect of a deterioration of *in utero* health caused by the Pandemic.

Insufficient Magnitude of Attrition

While over half a million people died because of the Influenza Pandemic, selective attrition is too small to account for substantial changes in the long-term outcomes of these birth cohorts. Across all age groups, influenza killed approximately half of 1 percent of the U.S. population. If we assume that this one-half percent had the lowest possible educational attainment, income, highest disability rates, etc., the

average outcomes for the surviving population would have improved by at most one-half of 1 percent immediately after the Pandemic as a result. Attrition is simply too small to account for substantial differences in subsequent outcomes for the surviving population.

The number of infant deaths in the 1918 birth cohort is too small to account for the inferior health of the surviving members of the 1919 birth cohort compared to the 1918 birth cohort. Infants were less likely to die from the Pandemic than those of older ages. As Figure 8b indicates, the share of total deaths in 1918 accounted for by those under age one fell substantially during 1918. Figure 8a indicates that infant mortality – deaths under age one – increased approximately ten percent during 1918, or about one percentage point.[22]

To account for even a 1 percent deterioration in outcomes (e.g. increased disability rates) for the 1919 birth cohort relative to 1918 would require that *all* the infants who died would have received been physically disabled. Implausible as this degree of selection is, if it indeed occurred, it would account for only a small fraction of the effects of fetal health estimated below for most outcomes. For example, the estimated effect of the Pandemic on adult disability rates in the 1919 birth cohort is between 5 and 6 percent. The absolute "worst-case" scenario of infant attrition could account for at most one fifth of this effect. Estimates for the effect of the Pandemic on income range from 1.5% to 6.5%, implying that "worst-case" attrition could not account for even the lower-bound estimate. More plausible degrees of selection would imply an inconsequential effect on observed outcomes.[23]

Deaths among persons of childbearing age from influenza could also skew the set of parents giving birth in 1920. This could in turn imply that the 1920 birth cohort was positively selected. Records from the nearly eighteen million policies of the Metropolitan Life Insurance Company indicate that around 0.35 percent of men and women of childbearing age died from influenza and pneumonia (Jordan 1927: 235). Among 20-29 year olds, 60,000 people died of influenza in 1918 in the 30 death registration states. Using the 1920 Census counts for these 30 states as the denominator: around 0.4 percent of the 20-

[22] Comparable to the number of deaths where influenza was listed as the cause.
[23] Another way to assess the magnitude of attrition differences across birth cohorts uses population counts from the Census microdata. As the Census identifies both year and place of birth, attrition can be calculated as the share of a state of birth cohort that is "missing" from the decennial census. Comparison of attrition shares between the 1918 and 1919 birth years using the 1960 Census data yields the same conclusion as the infant mortality data – attrition differences are too small to account for the observed outcome differences. The 1918 and 1919 birth cohorts have 1960 attrition shares within half of a percent of each other, even smaller than the infant mortality differences noted above.

29 year old population died from influenza in 1918. When the 38,000 pneumonia deaths in this age group are included, approximately 0.7 percent of people of childbearing age are estimated to have been killed.

The magnitude of the attrition as it affects conception is likely to be larger than the average mortality rate in the childbearing age group. If either member of a potential couple died from influenza, it could be the case that birth in 1920 was prevented. The "worst-case" assumptions that will maximize attrition as it affects the 1920 cohort are to assume 1) that none of the deaths in this age group occurred in both partners and 2) that where one of the two partners died, the surviving partner did not conceive a child born in 1920. These assumptions would imply that the mortality figure in this age group should be doubled. Using the largest mortality figure above, this would imply 1.4% percent of 1920 births could have been prevented by selective attrition of parents. However, even if these children would have had the worst possible outcomes had they been born, we could still generate a fraction of the relative deterioration in most outcomes for the 1919 birth cohort. For example, less than one third of the changes in disability rates could be explained by attrition, even if deaths occurred *only* among potential parents who would have given birth to physically disabled children. In summary, even under the most vitiating assumptions of bias magnitude, attrition is not large enough to account for the large deterioration in outcomes observed for the 1919 birth cohort.

Weakness of Selective Effect

In order for conclusions regarding effects of *in utero* health to be biased, attrition must also select on dimensions related to subsequent outcomes. The available evidence indicates that this selective effect was not especially strong. Indeed, the Pandemic was notorious for the apparent arbitrariness with which it struck. This section first summarizes evidence on the Pandemic's incidence from contemporaneous sources. Second, this section looks at the 1960 Decennial Census data to see if regions where influenza-induced attrition was greatest manifest better subsequent outcomes, as the model outlined above would predict. Finally, infant attrition is reconsidered in the context of the fixed-effects regression specifications. Insofar as infant and parental attrition is concerned, this prediction of the selection model is not borne out.

Descriptions of the Pandemic invariably refer to the fact that many in excellent health died. The lack of mortality microdata that include demographic and socioeconomic information means that the degree of selection is difficult to assess. To the extent that there was selection, death during the Pandemic

was apparently more likely among people of lower socioeconomic status, and therefore presumably the tendency was for people in poorer initial health to die of influenza.

The geographic incidence of the Pandemic provides little evidence of a selective effect. Mortality in urban and rural areas exhibited similar patterns. For example, the maternal mortality rate increased 37 percent in cities and 40 percent in rural areas. Variation in the Pandemic across states appears idiosyncratic; as Crosby notes, "The states with the highest excess mortality rates – Pennsylvania, Montana, Maryland, and Colorado – had little indeed in common economically or demographically, climatically or geographically." (Crosby 1989: 66). The pattern of maternal mortality changes across states appears arbitrary (Table 1, Figure 2). Within states as well, the Pandemic seemed to vary arbitrarily. For example, Dayton, Ohio's death rate was 80 percent higher than that in Columbus (Huntington 1933: 29). A 1933 study by the National Research Council found that factors such as latitude, longitude, percent black, and percent foreign born did not help explain the severity of the Pandemic by city.[24] In summary, the geographic variation in the Pandemic provides few clues to the relationship between influenza infection and socioeconomic status.

With regard to socioeconomic status, to the extent that there is a relationship, the Pandemic appeared to affect the poor more. According to Jordan, "It has been the general belief that the disease has prevailed more extensively in the poorer quarters of the large cities than in those inhabited by the well-to-do." (Jordan 1927: 208). Surveys of infection rates by "social status" in Little Rock and San Antonio found slightly elevated infection rates in poorer groups than among the wealthy (Jordan 1927: 208). Mortality rates paint a similar picture. Crosby notes that in Philadelphia, the Pandemic had a higher death rate in the "immigrant slums" than in other parts of the city (Crosby 1989: 87). While a survey in Chicago found that income was not associated with the likelihood of influenza death (Crosby 1989: 87), records from the Metropolitan Life Insurance Company stratified on three "industrial classes" lead to a different conclusion. Mortality rates were negatively correlated with the economic status of these three classes, which was true within each age grouping as well (Jordan 1927: 249). The available information on influenza infection and mortality rates indicate that wealth provided some protection.

Another approach for assessing the selectivity of death during the Pandemic uses the decennial census microdata. If selective attrition through infant mortality were indeed biasing results, we would expect that states suffering higher infant mortality in 1918 would have better outcomes for survivors

[24] The authors find that weather, the absolute humidity in particular, as measured near the time of the Pandemic's spike was the most important factor associated with Pandemic severity.

observed in the decennial census, because those with poorer subsequent outcomes had died. State-level differences in infant mortality rates do not indicate that selective early-life attrition is accounting for long-term outcome differences between the 1918 and 1919 birth cohorts. Appendix Figure 1 panels A, B, and C plot the 1919-1918 infant mortality changes by state against subsequent outcomes for these birth cohorts measured in 1960. For educational attainment, income, and labor force participation, despite substantial state variation in 1918-1919 infant mortality rate changes, no discernible pattern is observed in subsequent outcomes. The decennial census data do not indicate that infant deaths at the end of 1918 cause the 1918 birth cohort to appear stronger subsequently, and therefore bias comparisons toward finding the 1919 birth cohort in relatively poor health.

Fixed-effect regressions results (presented in Sections VII and VIII) that include the infant mortality rate provide a statistical test of the effect of infant attrition on subsequent outcomes. To the extent that states experiencing an idiosyncratically high maternal mortality rate in 1918 tended also to experience a high infant mortality rate in 1918, we would expect people born in such states to display worse outcomes if born in 1919 than if born in 1918. But this effect could exist for two distinct reasons -- selective attrition of the 1918 birth cohort or the deterioration in fetal health in 1918 exerting a persistent effect on outcomes. The fact that in some states maternal mortality was more impacted than infant mortality permits identification of whether infant attrition is confounding the hypothesized role of fetal health.

The fixed-effect regression results reject such a role for infant attrition. For each census outcome among men, inclusion of the infant mortality rate (in column (2)) strengthens the estimate of long-term *in utero* damage. Moreover, the infant mortality rate enters in the same direction as the maternal mortality rate. Thus, any effect of selective attrition through the first year of life is overwhelmed by the effect of damage to the health of surviving infants caused by influenza infection.

Nor does parental attrition appear to operate in a way that would make the 1920 birth cohort appear strong relative to the 1919 birth cohort. As noted above, the death of either a woman or man of childbearing age during the Pandemic could cause selection into conceptions and thereby births occurring after the Pandemic. Were selective attrition in the set of potential parents biasing results, we would expect that states with the highest 1918 death rates among those of child-bearing ages would tend to be states with the best subsequent outcomes among those born in 1920. Appendix Figure 2 panels A, B, and

C plot the 1918 death rate among 20-29 year olds[25] against the 1920-1919 difference in census outcomes. If parental attrition were accounting for the relative inferiority of the 1919 birth cohort, we should observe a positive relationship between 1918 attrition and the outcomes for the 1920 birth cohort. As with infant attrition, this pattern is not observed. If anything, the reverse is observed. States with higher 1918 attrition tend to have slightly worse subsequent outcomes.

Bias Against Finding an Effect of *In Utero* Health in the 1919 Birth Cohort

Attrition bias is a concern in this analysis to the extent that it affects the 1918 and 1920 birth cohorts more than the 1919 birth cohort. But the cohorts *in utero* during the last three months of 1918 experienced a 1¼ percentage point increase in the rate of stillbirths. If half of these infants would have been born in 1919, this implies that the 1919 cohort experienced a 0.6 percent increase in attrition due to the Pandemic. Moreover, as stillbirths are notoriously underreported even today, and given 1918 was the first year in which stillbirth data was reported in the United States, we would expect that the actual increase in stillbirths affecting the 1919 cohort was above 0.6 percent. Thus, insofar as magnitude is concerned, the level of attrition affecting the 1919 birth cohort is expected to be of a similar magnitude to that affecting the adjacent birth cohorts.

Moreover, the estimation is unable to account for the effect of increased attrition in the 1919 birth cohort through fetal death. For the 1918-1920 interval, data on stillbirths are available for 1918 alone. Therefore, the effect of increased attrition in the *in utero* cohort cannot be included in the estimation as it is for infant deaths. Thus, estimates of the impact of changes in fetal health (measured by state-level changes in the maternal mortality rate) cannot be adjusted for the fact that attrition is likely to have increased at the same time. In this important respect, the measure of the fetal health effect is biased toward zero.

Alternative Explanation #2: Selective Attrition After the Pandemic

Deaths occurring after the Influenza Pandemic among cohorts born near the time of the Influenza Pandemic may also alter the aggregate outcomes of survivors. If, for example, infant exposure to the Influenza Pandemic accelerated post-Pandemic death, the 1918 birth cohort would have experienced greater attrition between ages 1 and adult follow-up. If individuals who would have had worse average

[25] Measured as the increase from the 1917 death rate among 20-29 year olds in the state.

census outcomes died, then survivors in the 1918 birth cohort would appear stronger relative to other cohorts. Comparisons with the 1919 birth cohort would be biased toward finding damage in this cohort. Alternatively, correlation between the maternal mortality rate and post-Pandemic attrition could bias the fixed-effects estimates of the Pandemic's long-term effect.

This section looks at post-Pandemic attrition by year of birth and quarter of birth and finds no evidence that such attrition is operating in a way that would bias estimates of *in utero* influenza exposure. In addition, when post-Pandemic attrition is included in the fixed-effects analysis, it generally does not enter significantly. Moreover, its inclusion does not substantially affect the estimated effect of the *in utero* health proxy variable (the maternal mortality rate).

No longitudinal data on mortality are available prior to 1980 to analyze post-Pandemic attrition. In lieu of such data, the 1960 and 1980 Census estimated population counts are used to estimate post-Pandemic attrition. This approach is possible given 1) information on when and where each census respondent was born 2) information on the number of births by state and quarter of birth. Post-Pandemic attrition (PPA) is estimated from the 1 percent 1960 Census sample as:

$$\hat{PPA} = \frac{Births_{qob,s} - \hat{N}^{1960}}{Births_{qob,s}}$$

While the cohort born in 1918 experienced the highest infant mortality rate, the estimated attrition rate is not substantially higher than other birth years. 15.3 percent of the 1918 birth cohort had attrited by 1960 versus 15.1 percent of the 1919 birth cohort. The weakness of the age effect on attrition over these three birth years relates in part to the fact that two-thirds of attrition over the first forty years of life is accounted for by death in the first year of life. However, it is somewhat surprising that the approximately 1 percentage point increase in the infant mortality rate in 1918 generates such a small attrition difference when compared with the 1919 birth cohort. This is consistent with higher infant death rates occurring among those who were more likely to die in the subsequent 39 years of life, and implicitly negative selection of the Influenza Pandemic. No less surprising is that the highest attrition is estimated for the 1920 birth cohort at 16.1 percent. One possible explanation concerns service in World War II. The 1920 birth cohort had the highest service rates in the sample, at 78%.

In any event, PPA rates calculated from the 1960 Census are inconsistent with selective attrition accounting for inferior outcomes of the 1919 birth cohort. Similarly, PPA calculated using the 1980

Census data, while generating the more expected year-of birth-mortality profile, does not suggest a confounding role for post-Pandemic attrition. The 1919 birth cohort does not appear to have low post-Pandemic attrition relative to the adjacent cohorts.

The timing of the deterioration in adult outcomes by quarter of birth provides additional evidence against the role of selective post-Pandemic attrition. Figures presented in Section VI indicate a sharp departure from trend in the outcomes of the first birth quarter of 1919. Educational attainment, average income, and disability status all move discontinuously between the last three months of 1918 and the first three months of 1919. If attrition were accounting for this change, then it would have to follow the same pattern of changing discontinuously for those born at the end of 1918 versus those born at the beginning of 1919. This is not the case. The attrition rate for the last quarter of 1918 is relatively similar to the attrition rate for the first quarter of 1919. And in fact, the attrition rates for the first two quarters of 1919 are *higher* than at the end of 1918. This pattern is inconsistent with selective attrition accounting for the deterioration in outcomes in the 1919 birth cohort.

Finally, the effect of PPA may be gauged by including this variable in the fixed effects regression:

$$Log(income1960_i) =$$
$$\beta_0 + \beta_1(MMR_{yob_i-1,s_i}) + \beta_2(IMR_{yob_i,s_i}) + \beta_3(\hat{PPA}_{qob_i,s_i}) + \beta_4(Dummy_{yob_i}) + \beta_5(Dummy_{s_i}) + \beta_6(Dummy_{notwhite_i}) + \varepsilon_i$$

With infant mortality also included in the estimation, the coefficient for PPA, β_3, should reflect primarily the effect of post-infancy sample attrition. However, as PPA is measured at the quarter-of-birth level, it may also reflect within-year differences in infant attrition, and therefore potentially the effect of influenza infection in surviving infants as well.

Results from regression specifications that include PPA are reported in column (4) and (5) of Table 4 through Table 17. The estimated results are consistent with the existence of substantial measurement error, or the competing effects of infant damage and selective post-Pandemic attrition, but not of bias toward finding a "fetal origins" effect.

The sign of PPA is ambiguous – in approximately half of the regressions it enters in a direction consistent with selective attrition and in the balance the direction is consistent with lingering damage to infant health. For only one of the male outcome measures – the likelihood of high school graduation as

measured in 1960 – is β_3 estimated to be significantly different from zero. The effect of including PPA on the parameter of interest, β_1, is small and ambiguous in direction. The fixed-effects regression results indicate that attrition after the Pandemic is neither accounting for nor contributing to the finding of large effects of *in utero* health.

Alternative Explanation #3: Selective Fertility Caused by Mobilization for World War I

Fertility decisions may have been altered by the mobilization for U.S. entry into World War I. Troops began to leave for Europe in large numbers during the spring and summer of 1918. If these troops tended to be negatively selected and their imminent departure increased or accelerated fertility, inferior outcomes for the 1919 birth cohort could result from selective fertility.

However, if this effect were to exist, we would expect births to have increased in 1919. Births in each of the first three quarters of 1919 were lower than births in the corresponding quarters of 1918. While births in the fourth quarter of 1919 were higher than the fourth quarter of 1918, this difference is likely explained by the increase in stillbirths during the Pandemic. Moreover, births occurring in the fourth quarter of 1919 would have been conceived after World War I ended.

Alternative Explanation #4: Developmental Damage Caused by Parental Death

In the fixed effects regressions, the maternal mortality rate is used to proxy for the damage to fetal health caused by the Influenza Pandemic. In some cases of maternal death, newborn infants survived. The estimated negative effect of maternal mortality might represent the strictly "mechanical" effect of maternal death.

There are several reasons such an explanation is implausible. First, maternal deaths increased in the last quarter of 1918. The largest effect on infants whose mothers died in childbirth would therefore exist for children born in the last quarter of 1918. Results of Section VI indicate that outcomes deteriorated for the first birth quarter of 1919. Moreover, the magnitude of the increase in maternal mortality is insufficient to account for the outcome differences by birth cohort. The maternal mortality rate increased approximately one-quarter of one percentage point in 1918 while the estimated effects generally exceed one percent.

Many more influenza-related deaths occurred in the parent population generally than among mothers giving birth during the Pandemic. Death of a parent at an early age could cause inferior subsequent outcomes. However, the pattern of long-term effects is inconsistent with such an explanation. It would imply that for infants born in 1918, the death of a parent during the Pandemic exerted a relatively small effect, while having a parent die prior to birth exerted a much larger effect on development. Moreover, this effect would need to be discontinuous between those born when the Pandemic arrived, and those born three months later.

Alternative Explanation #5: "Age Heaping" In Census Data

The clustering of reported age at round numbers in some census years could generate inferior outcomes for the 1919 birth cohort. This is possible given the timing of the Influenza Pandemic vis-à-vis the decennial census. If "age heaping" is caused at least in part by older individuals reporting lower round-number ages, those reporting a round-number could include those with worse average outcomes.[26] This effect would be pernicious in the present context as those reporting an age of 60 in the 1980 Census, for example, would be assigned to the 1919 birth cohort, but actually tend to be older than 60.

The similarity of outcomes for the first and second quarters of birth of 1919 indicates that "age heaping" is not a factor in these results. As the decennial census is conducted as of April 1, those assigned to the first birth quarter of 1919 actually report an age of 61 in the 1980 Census data. Moreover, the incidence of age heaping in the 1960 and 1980 census is relatively uncommon. Finally, the incidence of age heaping would have to mimic the idiosyncratic geographic variation of the Influenza Pandemic to account for the fixed-effect results.

X. Discussion

With income, as with other long-term measures, the estimated coefficients on the measures of fetal health (the maternal mortality rate in the year preceding the birth year) are smaller in absolute value in the 1980 Census data than in the 1960 Census data. One possible explanation for this difference is that the relationship between early-life health conditions and subsequent outcomes is weakened by age 60 due to attrition. As attrition in 1980 is around 30 percent, and if death prior to 1980 was more likely among those exposed to influenza *in utero*, it would be the case that the share of persons exposed to the Pandemic and observed in 1980 has decreased. While the inclusion of the attrition components in both

[26] For most outcomes in the 1980 Census and some outcomes in the 1960 Census.

1960 and 1980 would imply that the maternal mortality rate coefficient incorporates the effect of attrition, estimations of the effect of the Pandemic on those whose mothers were actually infected are biased downward if a disproportionate share of those infected died before follow-up. In this case, a smaller share of people are then accounting for aggregate estimated effect, implying a larger "treatment on the treated" effect than was reported above for 1980.

For such a factor to be at play, the timing of adult mortality would also have been affected by exposure to the Influenza Pandemic. Figure 9 plots data from the National Longitudinal Mortality Study,[27] a random sample of the U.S population in 1980 that followed these respondents for the next nine years and observed whether the respondent died over this interval. The risk of dying in the 1980s falls steadily as birth year increases, reflecting that those who were younger at the 1980 baseline were naturally more likely to be alive at the end of the follow-up period. However, a departure from this steady trend is observed for the 1919 birth cohort. In terms of 1980s mortality risk, those born in 1919 are roughly one year "older" than their chronological age would suggest.[28] To the extent that adult attrition between 1960 and 1980 was also more likely for cohorts *in utero* during the Influenza Pandemic, the estimates of the effect of *in utero* health on income, disability, etc. in 1980 should be correspondingly scaled upward. This would increase the 1980 estimates of the effect of the Pandemic on those whose mothers were infected in the direction of the 1960 estimates.

The likelihood of death during the 1980s for cohorts born between 1918 and 1920 can be modeled using the fixed effects specification described in Section V. Within-cohort analysis indicates that the increase adult death rate for the 1919 cohort was greatest where influenza increased maternal mortality rates the most (results in Appendix Table 1). The .26 unit increase in the maternal mortality rate is estimated to have increased the probability of death during the 1980s by caused by 1.6 percentage points, or ten percent. This is approximately equal to one year of the estimated age trend in adult mortality risk. These results provide additional evidence that the 1980 Census estimates may be biased downward by selective attrition among adults.

Second, deterioration in the quality of data on birth information as respondents age may reduce fixed effect estimates in 1980. Information provided by the 1980 Census "data quality" flags for age, birth quarter, and place of birth indicate greater allocation of values for older respondents. As

[27] NLMS Public Use File, Release 2.
[28] This finding of increased mortality risk for the 1919 cohort can also be seen using the *universe* of death certificates from the 1980s.

respondents in the 1980 Census are twenty years older than the 1960 respondents, the quality of the information for records in which no allocation was made may deteriorate with age as well. This deterioration would be key in the regression analysis because identification of the maternal mortality rate coefficient comes from the differences in state and year of birth. Under classical measurement error, the result would be attenuation bias in the direction of the estimated 1960-1980 point-estimate differences.

Third, it is possible that the effect of early-life conditions exerts less of an impact as cohorts age. To the extent that post-birth health and economic shocks affect health independent of fetal-induced damage, we would expect early injuries to become a less important determinant of subsequent outcomes as cohorts age. Alternatively, compensatory heath investments could mitigate the impact of early-life health conditions over time. Nevertheless, even if these effects diminish, it appears that early-life health conditions exert an important effect on outcomes throughout life. As noted above, the estimated coefficients indicate that income was eight percent lower at age sixty for those whose mothers contracted influenza and disability rates were 17 percent higher.

Estimates of the effect of the Influenza Pandemic were calculated using two distinct approaches. Results of the cohort differences approach are similar to the results obtained from the 1980 fixed effects regressions. For example, the fixed-effect approach gives a slightly larger estimate for the effect on disability status (6% increase for the 1919 birth cohort) than the cohort differences approach (5%). The effect for socioeconomic status is estimated to be lower using the fixed-effects approach (1.6%) than that estimated using cohort differences (2.4%). Larger differences exist for the 1960 fixed effects estimates, indicating larger long-term effects of the Pandemic than estimates based on cohort differences. To the extent that attrition in the 1919 cohort (e.g. through stillbirths, which cannot be included in the analysis because data are unavailable) biases downward the estimates of damage caused by fetal health in survivors using cohort differences, it could be the case that the estimates of the effect of maternal mortality exhibit less downward bias if such selective attrition is weakly correlated with the maternal mortality rate differences. Maternal mortality in the fixed effects regression is then registering the relatively unadulterated effect of fetal damage on survivors.[29]

Including infant mortality in the fixed-effects regressions accounts for differences in early-life attrition that might be correlated with differences in fetal health exposures. Such attrition has the

[29] That the maternal mortality rate coefficient increases only slightly when attrition components are included in the estimation could suggest that the relationship between maternal morality rates and selective attrition is indeed weak. However, conceptually it is difficult to imagine that fetal damage on survivors would operate very differently from fetal damage leading to selective attrition.

potential for biasing estimates toward finding an effect of changes in fetal health (to the extent that comparisons are made between 1918 and 1919 aggregate outcomes, or to the extent that infant mortality rate changes are correlated with changes in the maternal mortality rate in the preceding year). But under this conception of infant mortality as reflecting selective attrition, the infant mortality rate would have entered in the opposite direction from maternal mortality, which is not observed.

Instead, fixed-effects regression results indicate that exposure to the Influenza Pandemic during infancy caused long-term damage in survivors. Such a finding is plausible given a) the small magnitude of the expected attrition effect, as described in Section IX, b) the fact that approximately twenty times as many infants were infected with influenza and survived as died in infancy due to influenza, and c) results from other research finding long-term effects of infant health. While in aggregate the magnitude of the estimated effect of changes in infant health in the regression results is small, this is consistent with the fact that in aggregate, 1918 (when infant mortality rates increased) was not the "bad" birth year. The primary long-term effect of the Pandemic operates through changes in fetal health, and therefore exerts the largest effect on the 1919 birth cohort.

The costs borne by those born just after the Influenza Pandemic were large. Over the life course, income losses for survivors of the Pandemic exceed $14 billion dollars. Results from the National Longitudinal Mortality Study indicate accelerated adult mortality as well. If the estimates of mortality risk during the 1980s can be extrapolated to the life-course, those exposed to influenza *in utero* died three years earlier due to the Pandemic. For the approximately 2.2 million people born in the United States in 1919 (and surviving infancy), one-third of whom were likely to have been exposed to influenza *in utero*, this implies 2.2 million adult years of life lost.

XI. Can It Happen Again?

Using a sample of preserved lung tissue from 1918, it was reported in *Science* in 2004, that the 1918 influenza virus was a strain of the avian flu (Gamblin et al., 2004). Avian flu strains can circulate in poultry, but are far less common in humans. Therefore, humans have little resistance if infected. The last major outbreak was in 1968, killing 700,000 people. While humans have contracted avian flu directly from poultry in recent years, no Pandemic has resulted since, so far, the virus has not transmitted easily among humans.

This may change. In May of 2005, the World Health Organization declared that avian influenza is "the most serious known health threat the world is facing today."[30] By March 2005, a strain of H5N1 bird flu had spread to ten countries, prompting the slaughter of 50 million chickens.[31] Two months later, the strain had infected pigs in Java (Cyranoski, 2005). This development is particularly alarming, according to Cyranoski, because human influenza viruses also circulate in pigs. Were the two viruses to mix, a far more contagious version of the avian flu may result. Moreover, according to the Michael Osterholm, Associate Director of the Department of Homeland Security's National Center for Food Protection and Defense: "Recent clinical, epidemiological, and laboratory evidence suggests that the impact of a pandemic caused by the current H5N1 strain would be similar to that of the 1918-19 pandemic" (Osterholm, 2005). The May 25, 2005 editorial in the journal *Nature* stated:

> The maths of epidemiology says that pandemics are like fault lines: they inevitably give. But unlike earthquakes, pandemics tend to give warning signs, and all the alerts from Asia are now flashing red.

Still, the world is woefully unprepared for another flu pandemic. In January 2005, a *Nature* editorial stated: "governments are still not doing enough to monitor and prepare for the next viral pandemic. The inaction is scandalous." While certain international surveillance activities have improved, a major concern is the low level of production of flu vaccines. Currently, word-wide vaccine production capacity is only 450 million doses per year (*Nature,* 2005) and antivirals are available for only 40 million persons (Osterholm, 2005). Were a pandemic to begin in the next few years, it would be impossible to increase production before the bulk of the pandemic's destruction had unfolded. The shortage of flu vaccines in fall of 2004, caused by problems in the production of one of the two existing producers for the U.S. market, highlighted the inability of supply to respond in the short run to shortages.

An obstacle for producers is that vaccines and antivirals often go unsold in ordinary flu seasons. More to the point, were a Pandemic to begin, producers of vaccines and antiviral drugs would probably not be allowed to sell at the profit-maximizing price. Rather, producers expect price controls and requisitions of existing supplies.*[32] In order to restore proper incentives for higher production in the event of a Pandemic, the government would need to commit to additional purchases.

---

[30] Statement of Dr. Lee Jong-Wook, Director-General of the World Health Organization to the 58th World Health Assembly, May 18, 2005.

[31] "Web Focus – Avian flu timeline." Available at: http://www.nature.com/nature/focus/avianflu/timeline.html (downloaded 6/22/05).

[32] In the July 2005 Conference on Pandemic Preparedness, organized by the Royal Institution World Science Assembly, it was concluded that vaccine producers would probably be nationalized were a Pandemic to unfold.

The obvious question is whether the benefit of a policy where most of the output may go unused exceeds the cost. Were we to consider a consider a policy of ensuring a vaccine or antiviral supply for all pregnant women in the United States so as to protect the health of their babies, the majority of benefits may not be reaped for a generation to come, when these children begin to work. In contrast, the costs of the purchasing program would begin immediately. Nevertheless, it appears that such a policy would be socially desirable.

This in part is because the wholesale cost of flu vaccines and antiviral medications is low. The Centers for Disease Control (CDC) in the United States contracted to purchase flu vaccines in 2005 for nine dollars per dose (although the vaccine can wholesale for as little as two dollars per dose in some countries,[33] the lower price presumably reflecting its marginal cost). Antiviral medications, such as Oseltamivir (Tamiflu), are available for eight dollars per dose.[34] Assuming that the influenza pandemic that *Nature* and the World Health Organization believe to be looming does not arrive for another twenty years, approximately four million vaccines will need to be purchased for twenty years since the virus mutates, a new vaccine is required every year, until the Pandemic arrives.[35] (In estimating expected costs of government purchases, the CDC vaccine purchase price of $9 for each year will be assumed.)

The benefits of being treated as a Pandemic begins would accrue to both the mother and child. If we ignore the benefit to the mother (which is commonly put forward as the primary reason for vaccinating pregnant women during "ordinary" influenza seasons), as well as the benefit to the child for the first twenty-five years of life, then we are left with the benefits to the *in-utero* child as he or she enters the workforce. Moreover, if we ignore the increase in disability rates that prevented many from working after the 1918 Pandemic, and instead limit the potential benefit to unaltered productivity rates among workers, we can use the 4.2% decrease in wages caused among those infected by the 1919 Pandemic as a conservative starting point for calculating benefits of inoculation. If the Pandemic does not occur until twenty years from today (as assumed with the costs above), then the benefits of a federal purchasing program will not begin to be realized until *forty-five years* from today.

Nevertheless, at a five percent discount rate, the present-day expected benefits of such a program exceed the costs by a factor of five, despite conservative assumptions that overestimate costs and underestimate benefits. Therefore, a U.S. policy of stockpiling flu vaccines or antiviral drugs for

[33] July 7, 2005 email from Dr. Klaus Sohr, World Health Organization Global Influenza Program.
[34] Ibid.
[35] Antiviral drugs, as they are not strain-specific, could presumably be purchased with less frequency, depending on shelf life.

pregnant women seems warranted.[36,37] Moreover, the time for implementing such a policy is opportune. On July 6, 2005, the Royal Institution World Science Assembly convened representatives from the World Health Organization, the World Bank, the National Institutes of Health, and experts in pandemic influenza to discuss ways to improve pandemic preparedness, including the development of action plans to be presented to policy makers. It would seem unfortunate if long-term health and socioeconomic costs were ignored in that plan.

XII. Conclusion

Early population-based studies of the fetal origins hypothesis have been faulted for being "stacked" by omitted variables toward finding an effect. Consequently, *The Lancet*, among others have advocated the use of research designs that would provide more of a challenge for finding evidence of fetal origins. The unexpected and sudden fetal health shock caused by the 1918 Pandemic, together with is apparently arbitrary geographic incidence, make it a fitting natural experiment. Moreover, maternal influenza infection has not even been put forward by the "fetal origins" school as potential route for this effect. Nevertheless, results from the 1960, 1970, and 1980 U.S. Census surveys indicate a broad spectrum of persistent effects, including reduced educational attainment, socioeconomic status, labor force participation, and increased disability.

This paper therefore concludes that fetal origins linkages indeed exist, and that the magnitude of the effect on adult economic outcomes can be large. Moreover, the implications of the existence of such linkages are squarely in the purview of economics. It implies, of course, that resources are not being allocated optimally across the life cycle: individual investments and public policies that benefit maternal and fetal health have been under-funded if fetal origins effects have not been accounted for in expenditure decisions, as they presumably have not. A means of improving education, productivity, and socioeconomic status has been overlooked. Therefore, social welfare can be substantially improved.

The practical challenge, of course, remains to identify which improvements to fetal health exert positive effects on subsequent outcomes, and the cost-effective public policies that achieve these improvements. Preventing fetal damage during a new Pandemic appears to be one such policy.

[36] Once a Pandemic arrives, flu vaccines or antivirals could be made available for short interval (say three days, or the length of time less than that required to generate a positive pregnancy test) to women who passed a pregnancy test.

[37] The United Kingdom has already begun stockpiling antiviral drugs.

Identifying additional policies and their expected returns is clearly a vital and demanding area for future economic research.

Sources

Almond, Douglas and Kenneth Y. Chay. "The Long-Run and Intergenerational Impact of Poor Infant Health: Evidence from Cohorts Born During the Civil Rights Era." Mimeo, UC Berkeley Department of Economics, May 2005.

Almond, Douglas and Bhashkar Mazumder. "The Long-term Health Effects of the 1918 Influenza Pandemic: An Analysis of SIPP Data." *American Economic Review Papers and Proceedings*, forthcoming.

Brainerd, Elizabeth, and Mark V. Siegel. "The Economic Effects of the 1918 Influenza Epidemic." CEPR Discussion Paper No. 3791, February 2003.

Angrist, Joshua and Alan Krueger. "Does Compulsory School Attendance Affect Schooling and Earnings?" Quarterly Journal of Economics, November 1991, 106: 979-1014.

Barker, D.J.P. Mothers, Babies and Health in Later Life. Edinburgh: Churchill Livingstone: 1998.

Barker, D.J.P. Fetal Origins of Cardiovascular and Lung Disease. New York: Marcel Dekker: 2001.

Case, Anne, Angela Fertig, and Christina Paxson. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics*, 2005, 24: 365-389.

Case, Anne, Darren Lubotsky, and Christina Paxson. "Socioeconomic Status and Health in Childhood: The Origins of the Gradient." American Economic Review, December 2002, 92(5): 1308-1334.

Costa, Dora L. "Understanding Mid-Life and Older Age Mortality Declines: Evidence from Union Army Veterans." Journal of Econometrics, January 2003, 112(1): 175-192.

Costa, Dora L. "Understanding the Twentieth-Century Decline in Chronic Conditions Among Older Men." Demography, February 2000, 37(1): 53-72.

Costa, Dora L and Joanna Lahey. "Becoming the Oldest-Old: Evidence from Historical U.S. Data." NBER Working Paper No. 9933: 2003.

Couzin, Jennifer. "Quirks of Fetal Environment Felt Decades Later." Science, June 2002, 296: 2167-2169.

Crosby, Alfred W. American's Forgotten Pandemic: The Influenza of 1918. New York: Cambridge University Press: 1989.

Currie, Janet and Enrico Moretti. "The intergenerational Transmission of Health: Evidence from Grandmothers, Mothers, and Daughters." Mimeo, UCLA Department of Economics, May 2005.

Cyranoski, David. "Bird Flu Spreads Among Java's Pigs." Nature, May 26, 2005, 435 (1038): 390-391.

Elo, Irma T. and Samuel H. Preston. "Effects of Early-Life Conditions on Adult Mortality: A Review." Population Index, Summer 1992, 58(2): 186-211.

Evans, William N. and Stephen E. Snyder. "The Impact of Income on Mortality: Evidence from the Social Security Notch." NBER Working Paper No. 9197: 2003.

Fogel, Robert W. "Economic Growth, Population, Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." American Economic Review, June 1994, 84(3): 369-395.

Fogel, Robert W. and Dora L. Costa. "A Theory of Technophysio Evolution, With Some Implications For Forecasting Population, Health Care Cost, and Pension Costs." Demography, February 1997, 34(1): 49-66.

Gamblin, S. J., et al. "The Structure and Receptor Binding Properties of the 1918 Hemagglutin." Science, March 19, 2004, 303: 1838-1842.

Harris, John W. "Influenza Occurring in Pregnant Women." The Journal of the American Medical Association, April 5, 1919, 72(14): 978-980.

Heider, Fritz. "The Influence of the Epidemic of 1918 on Deafness." American Journal of Hygiene, 1934, 19: 756-766.

Huntington, Ellsworth, et al. "Causes of Geographical Variations in the Influenza Epidemic of 1918 in the Cities of the United States." Committee on the Atmosphere and Man, National Research Council, July 1924, 6.3(34): 1-36.

Jordan, Edwin O. "Epidemic Influenza: A Survey." American Medical Association, Chicago: 1927.

Kolata, Gina. Flu: The Story of the Great Influenza Pandemic of 1918 and the Search for the Virus That Caused It. Touchstone, New York: 1999.

The Lancet. "An Overstretched Hypothesis?" February 10, 2001, 357 (9254): 405.

Langford, Christopher. "The Age Pattern of Mortality in the 1918-19 Influenza Pandemic: An Attempted Explanation Based on Data for England and Wales." Medical History, 2002, 46: 1-20.

Lee, Richard V., et al. "Medical Care of the Pregnant Patient." American College of Physicians, Philadelphia: 2000.

MacDorman, Marion F. "From Data to Action: CDC's Public Health Surveillance for Women, Infants, and Children." Monograph, Centers for Disease Control and Prevention: 1994.

National Center for Infectious Diseases. Influenza: The Disease. Atlanta: Centers for Disease Control and Prevention: 2003.

National Longitudinal Mortality Study Bibliography, March 21, 2001.

Nature. "Dangerous State of Denial." January 13, 2005, 433 (7022): 91.

Nature. "On a Wing and a Prayer." May 26, 2005, 435 (1038): 385-386.

Noymer, Andrew and Michael Garenne. "The 1918 Influenza Epidemic's Effect on Sex Differentials in Mortality in the United States." Population and Development Review, September 2000, 26(3) : 565-581.

Osterholm, Michael T. "Preparing for the Next Pandemic." *The New York Times*, June 21, 2005.

Rasmussen, Kathleen M. "The 'Fetal Origins' Hypothesis: Challenges and Opportunities for Maternal and Child Nutrition." Annual Review of Nutrition, 2001, 21: 73-95.

Robinson, Roger. "The Fetal Origins of Adult Disease." British Medical Journal, February 2001, 322: 375.

Rudolph, Colin D. and Abraham M. Rudolph. Rudolph's Pediatrics McGraw-Hill, 2003.

Ruggles, Steven and Matthew Sobek et al. "Integrated Public Use Microdata Series: Version 3.0." Historical Census Projects, University of Minnesota, Minnespolis, 2003.

Smith, James P. "Healthy Bodies and Thick Wallets: The Dual relation Between Health and Economic Status." Journal of Economic Perspectives, 1999, 13(2): 145-167.

Stein, Zena, et al. Famine and Human Development: The Dutch Hunger Winter, 1944-1945. Oxford University Press, New York, 1975.

Vital statistics of the United States (various annual volumes). United States Bureau of the Census, Washington DC.

Warrell, David A.., Timothy M. Cox, John D. Firth, and Edward J. Benz Jr. Oxford Textbook of Medicine Oxford University Press, 2003.

Winn, Hung N. and John C. Hobbins (editors). *Clinical Maternal-Fetal Medicine*. The Parthenon Publishing Group, New York: 2000.

Figure 1a: U.S. Influenza Deaths By Year

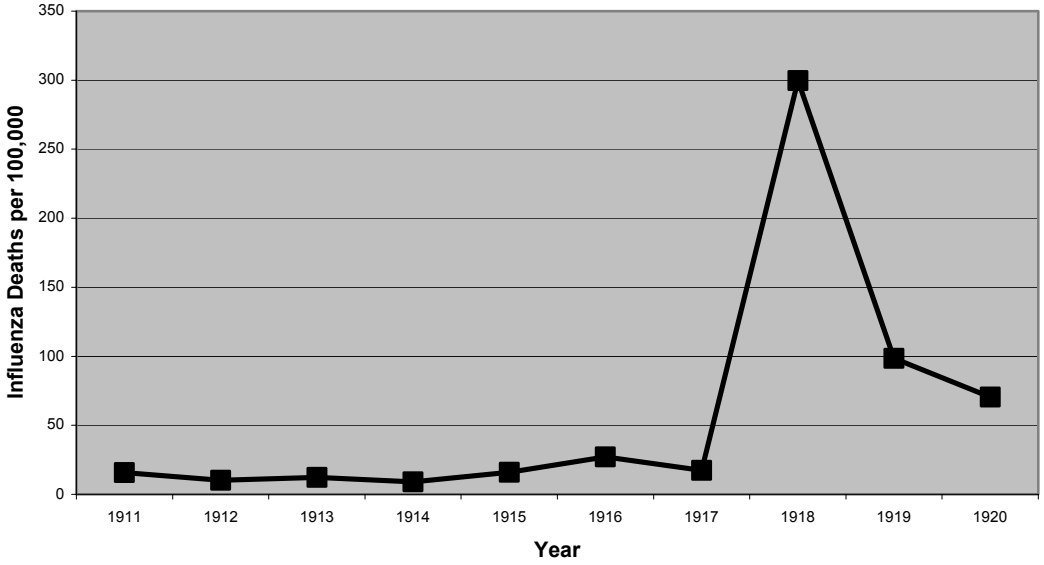


Figure 1b: U.S. Influenza Deaths By Month

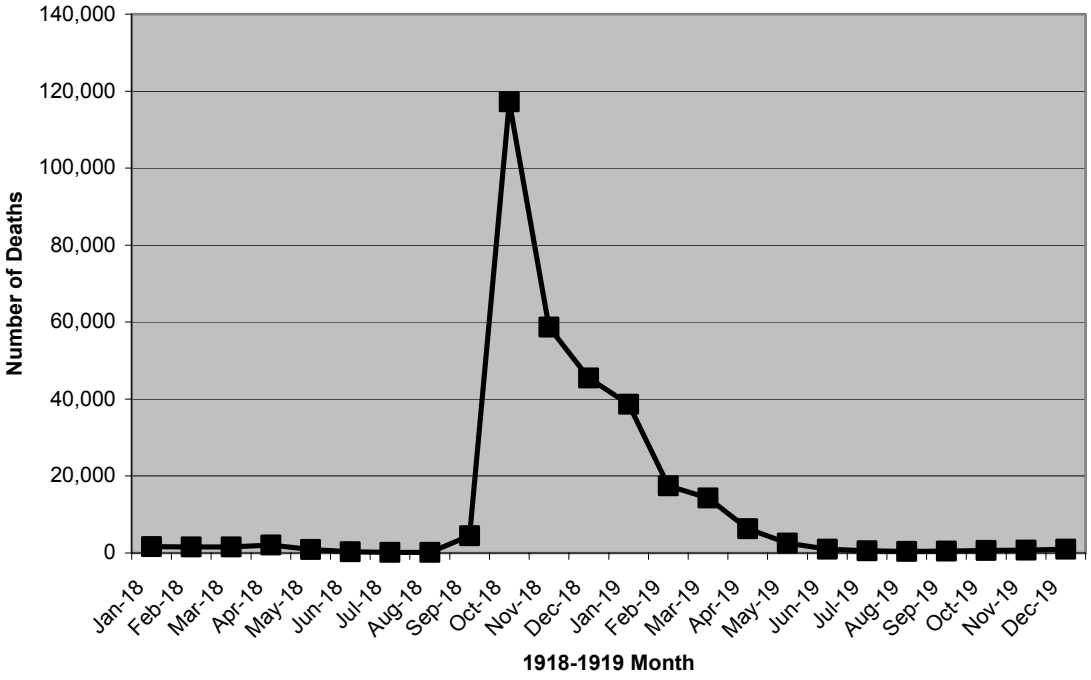


Figure 2: Maternal Mortality Rates During 1918 Influenza Pandemic by State

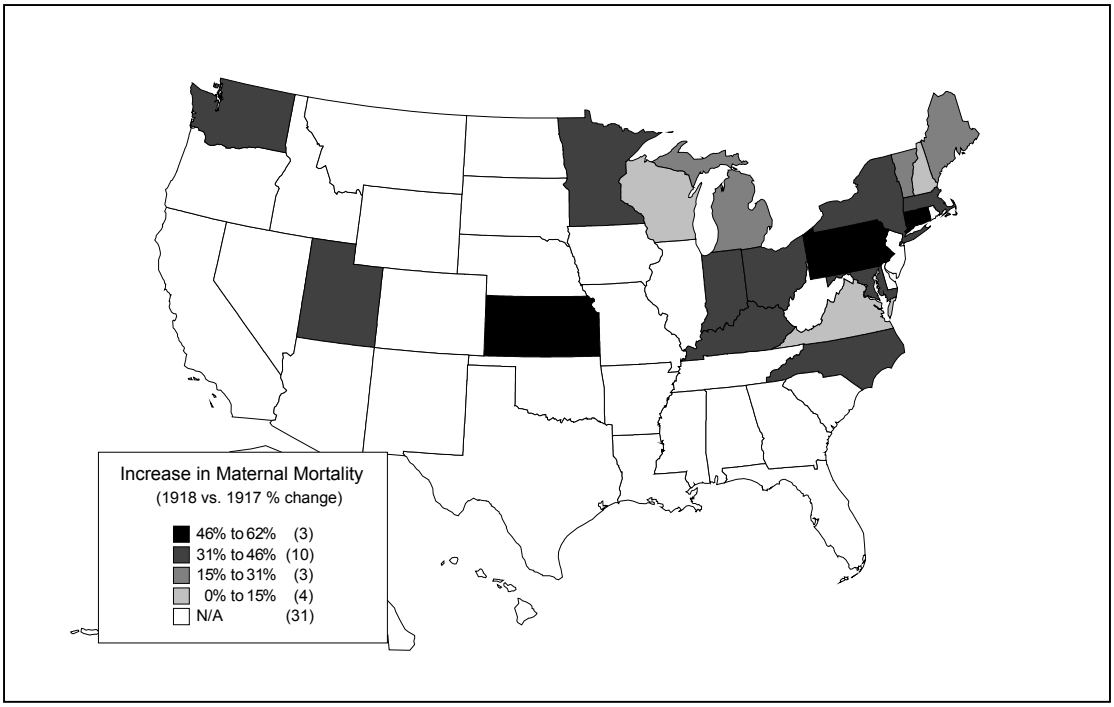


Figure 3a: Fall 1918 Influenza and Pneumonia Infection Rates in Maryland Females by Age

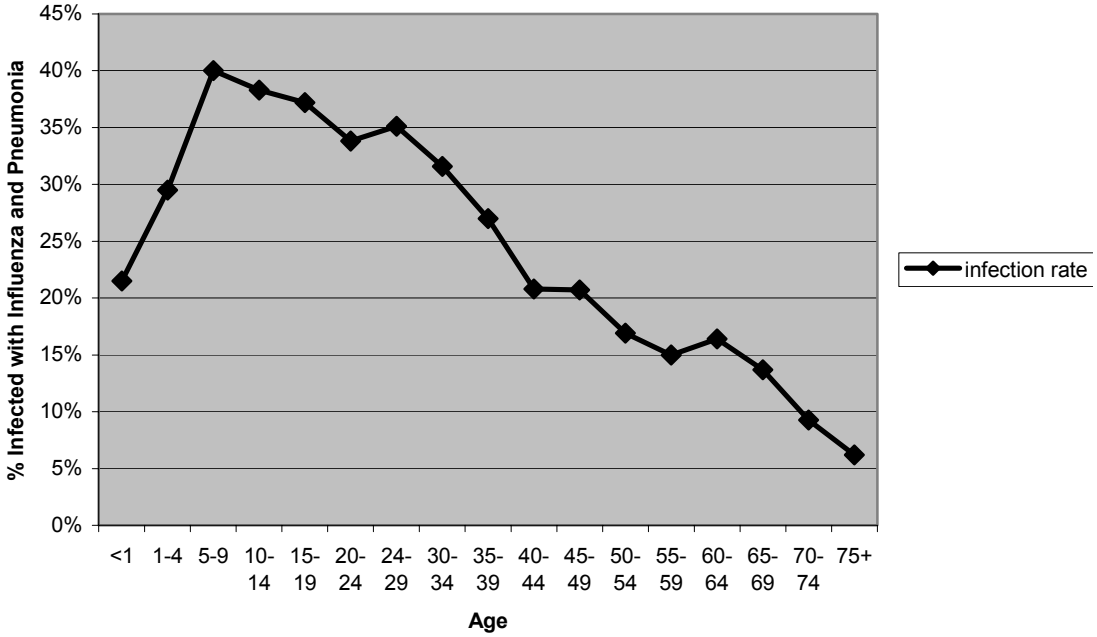


Figure 3b: Average Stillbirth Rate in Seventeen U.S. States By Month of 1918

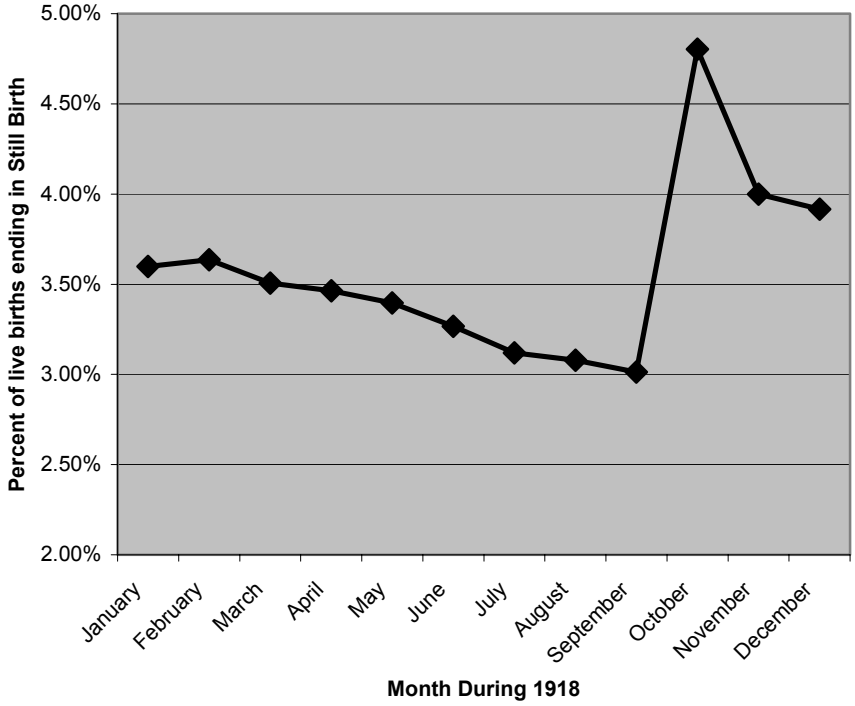


Figure 4a: 1960 Average Years of Schooling
(Men and Women Born in the United States)

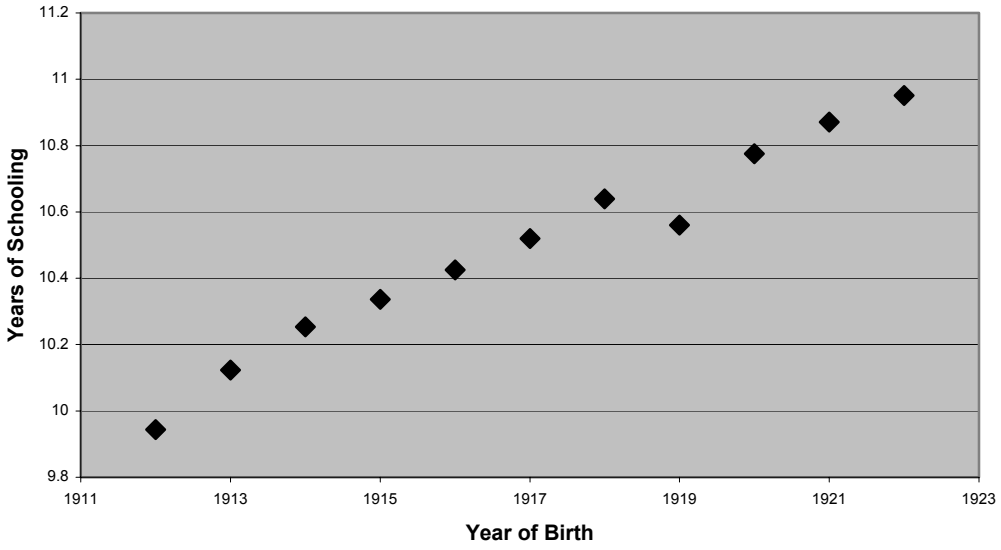


Figure 4b: 1970 High School Graduation
By Birth Cohort and Gender

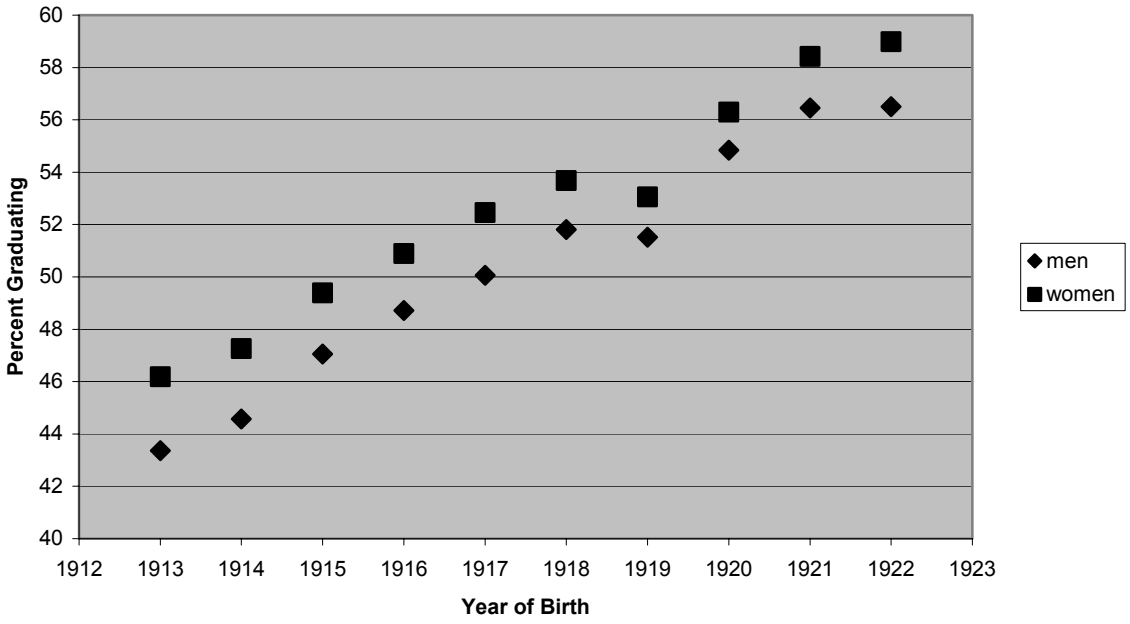


Figure 5a: 1980 Average Years of Schooling Among Men

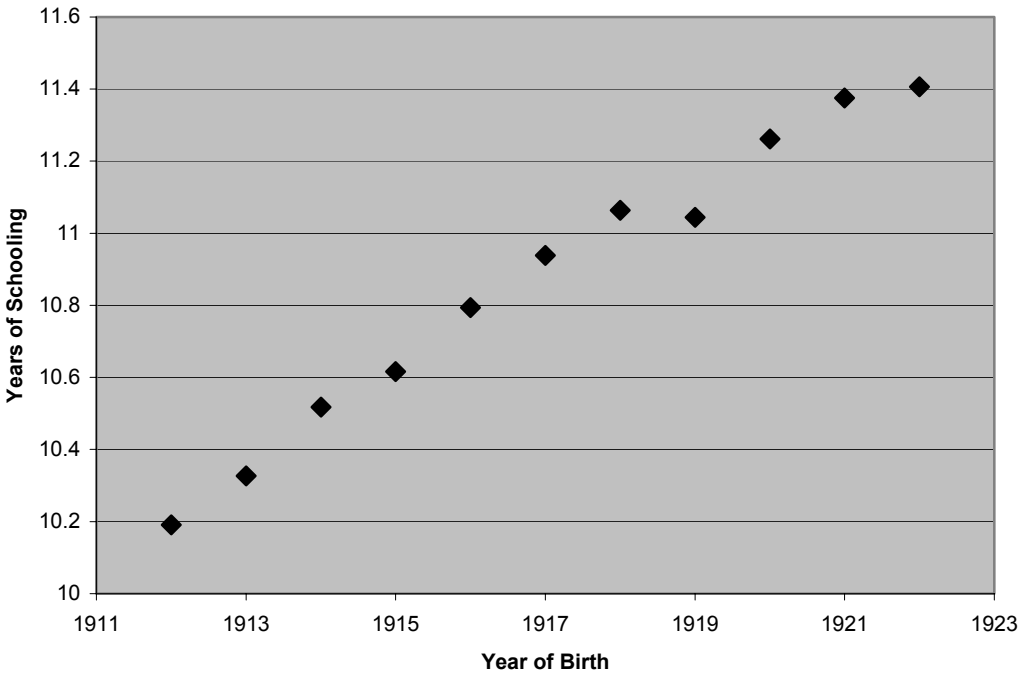


Figure 5b: 1980 Average Income Among Men

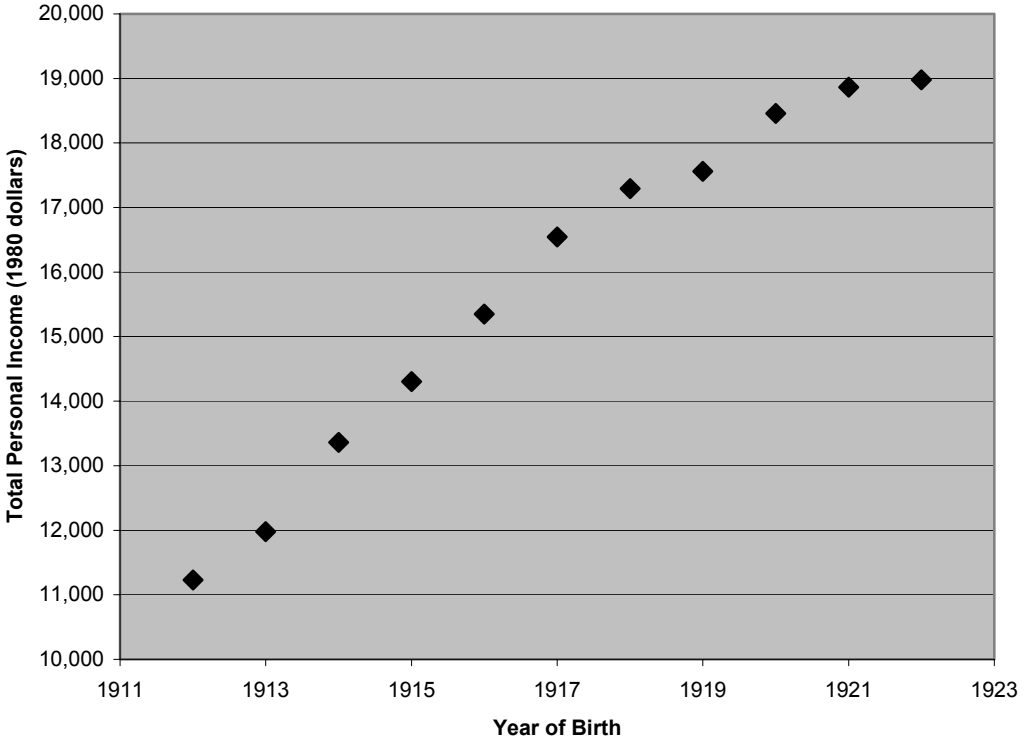


Figure 6a: 1980 High School Graduation Rate By Quarter of Birth

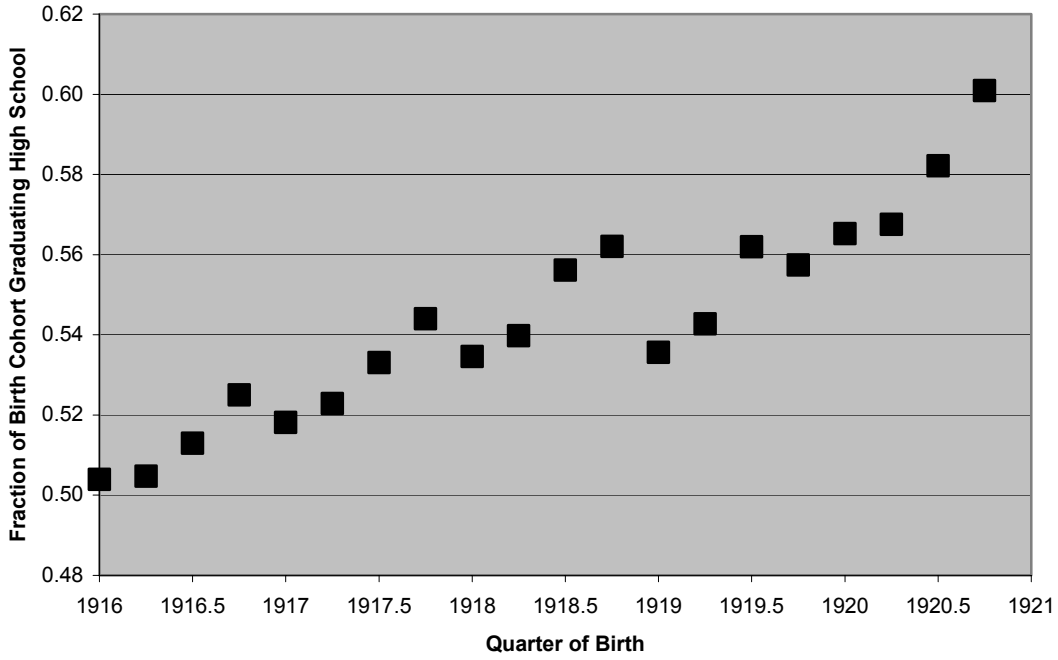


Figure 6b: Regression-Adjusted 1980 High School Graduation Rate By Quarter of Birth

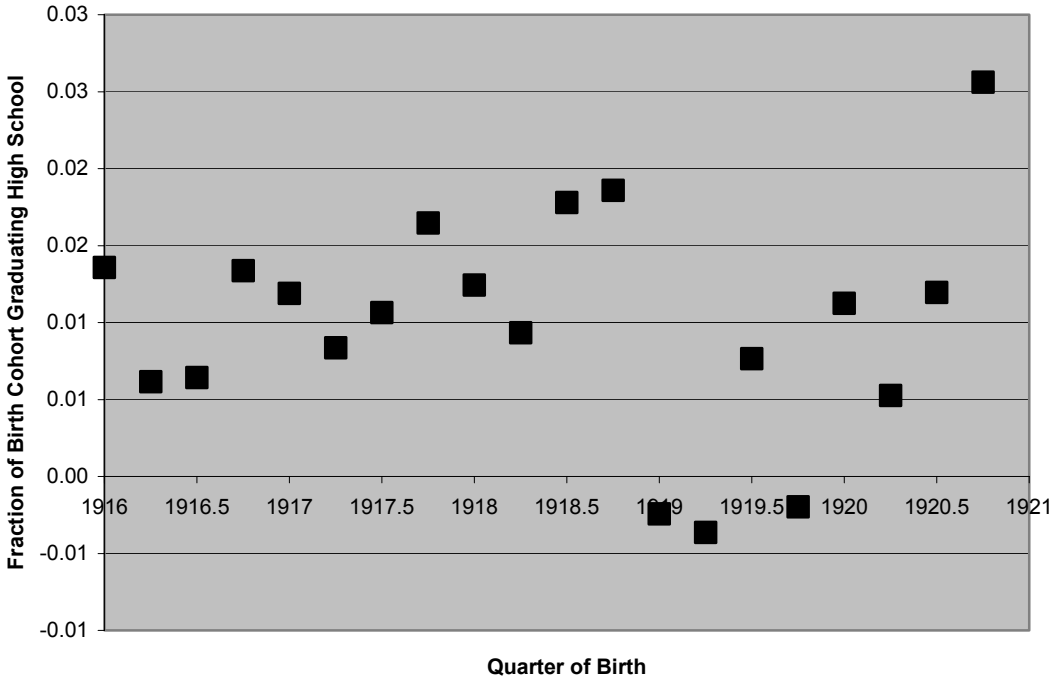


Figure 7: 1980 Male Disability Rate By Quarter of Birth

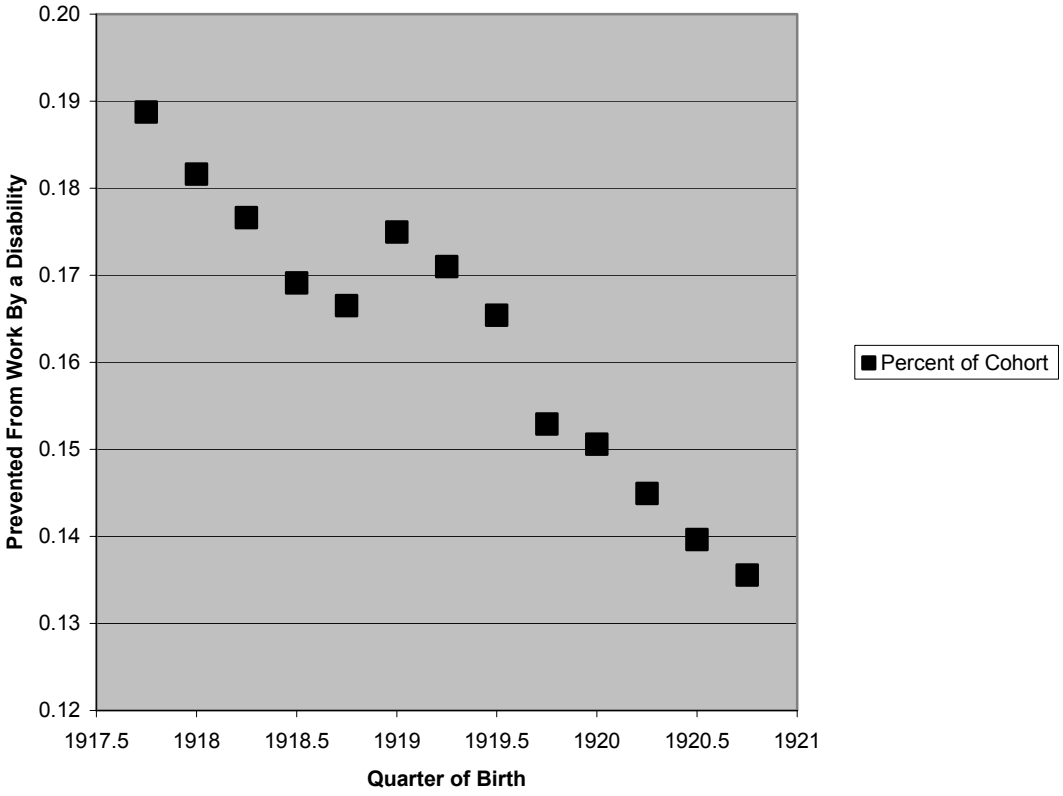


Figure 8a: Infant Mortality in United States
(1917 Birth Registration Area)

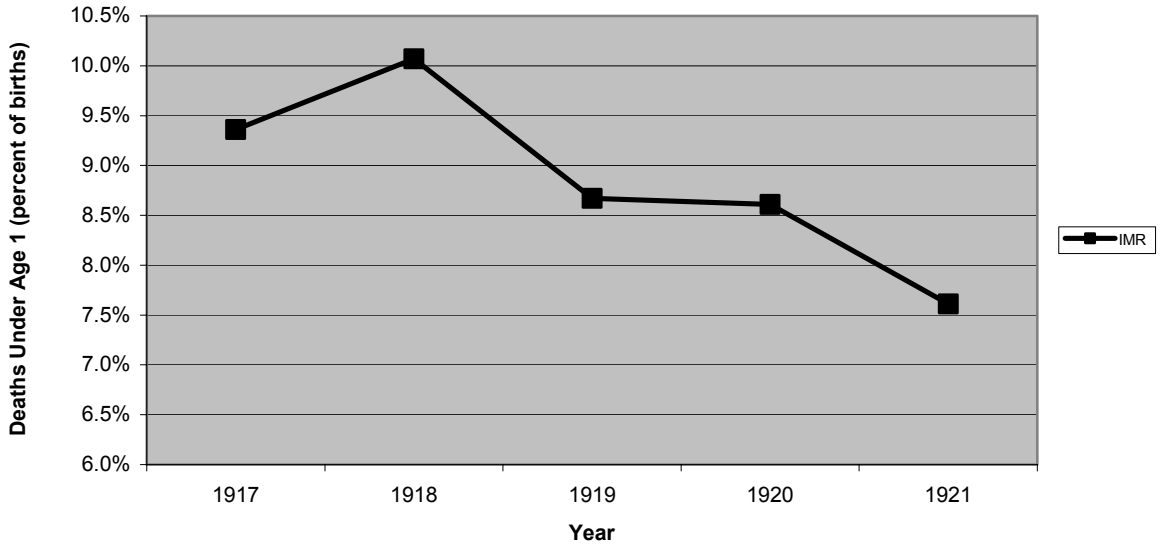


Figure 8b: Infant Share of Total Deaths

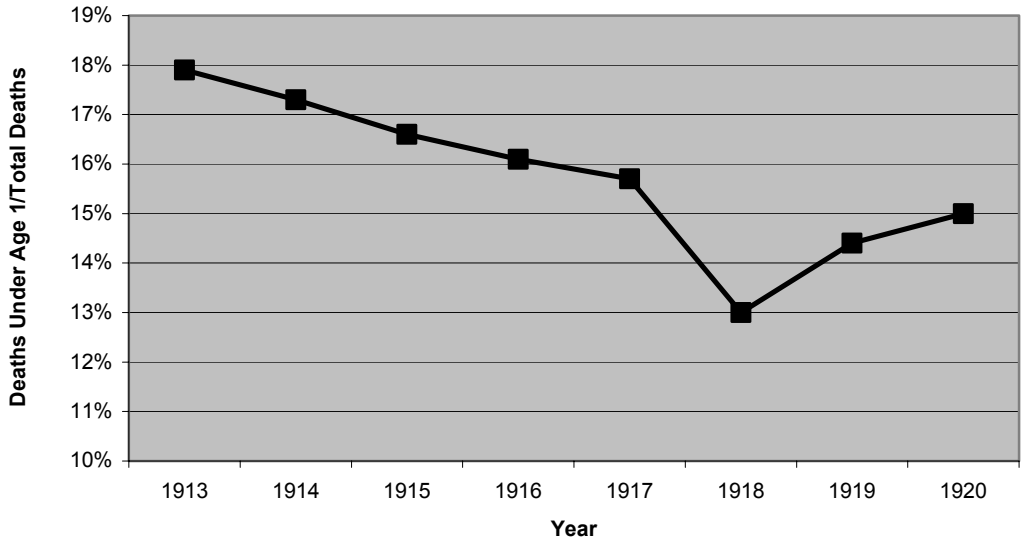


Figure 9: Percent of Adults Dying During 1980s
By Year of Birth (NLMS data)

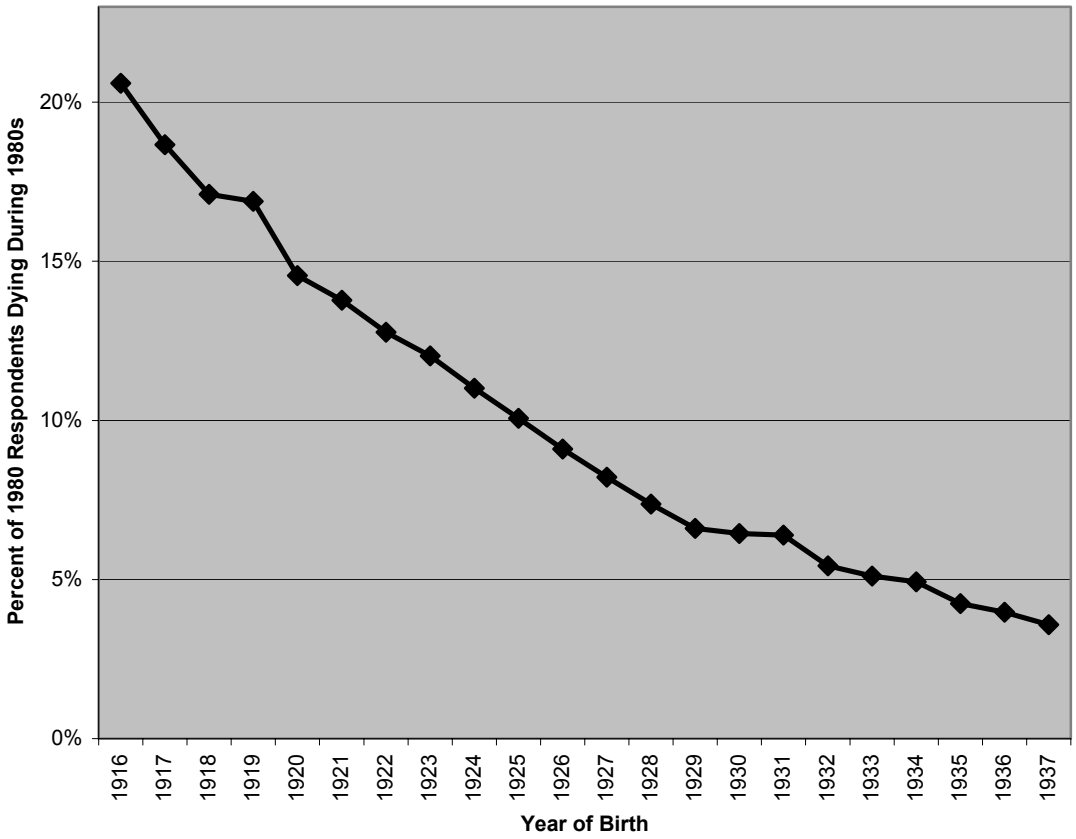


Table 1
Maternal Mortality Rate By State
(States in the 1917 U.S. Birth Registration Area)

Maternal Mortality Rate By State
(Deaths per 1,000 live births)

State	1917	1918	1919	**Spike measure** (1918-1917 % change)
Pennsylvania	6.5	10.5	6.8	62%
Kansas	7.6	11.4	8.2	50%
Connecticut	5.1	7.5	6.2	47%
Utah	5.9	8.6	8.4	46%
Indiana	7.3	10.4	8.4	42%
Massachusetts	6.5	9.2	7.1	42%
New York	5.7	8	6.2	40%
Maryland	6.8	9.5	8.4	40%
Minnesota	5.6	7.8	6.7	39%
Ohio	7.1	9.7	7.4	37%
Washington	7.4	9.9	8.6	34%
Kentucky	6	8	6.3	33%
North Carolina	8.2	10.8	9.3	32%
Virginia	8.2	10.7	6.3	30%
Maine	6.7	8.6	8.6	28%
Vermont	6.4	8	8	25%
Michigan	7.4	8.6	7.7	16%
New Hampshire	7	7.8	8	11%
District of Columbia	8.6	9.1	8.6	6%
Wisconsin	5.7	6	4.8	5%
Average	6.6	9.2	7.0	39%

Table 2
Summary Statistics of 1918-1920 Birth Cohorts:
Men Born in the Nineteen United States

		Mean	Standard deviation
1960 Census Measure			
	Respondent's Age	40.3	0.93
	Total Income (1960 dollars)	$6,200	$4,106
	Years of schooling	11.0	3.3
	Graduated from high school	0.53	0.50
	Not in labor force	.036	0.19
	Below 150% Poverty Line	.24	.42
	Duncan Socioeconomic Index of Occupational Category	36.9	24.1
	World War II Veteran	.71	.45
	Married	.87	0.34
	Race is White	.95	.22
Pre-1960 Measure (% of live births)			
	Maternal mortality rate in state f birth[38]	.76	.16
	Infant Mortality Rate	9.1	1.5
	Estimated attrition between birth and 1960 Census follow-up	15.5	9.3
Observations		16,566	N/A

Table 3
Log Personal Income in 1960
Among Men Born 1918-1920

	(1)	(2)
	LINCTOT	LINCTOT
Maternal Mortality Rate	-0.333*	-0.669**
	(0.181)	(0.265)
Birth year =1919		0.168*
		(0.085)
Birth year =1920		0.026
		(0.058)
Observations	16240	16240

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

[38] For the year preceding the birth year (i.e. 1917-1919).

Table 4
Log Personal Income in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	LINCTOT	LINCTOT	LINCTOT	LINCTOT	LINCTOT
Maternal Mortality	-0.228*	-0.254***	-0.243**	-0.247**	-0.236**
	(0.118)	(0.095)	(0.094)	(0.098)	(0.097)
Infant Mortality Rate		-0.019**	-0.014*	-0.019**	-0.015*
		(0.009)	(0.008)	(0.009)	(0.008)
Total Attrition				0.061	0.069
				(0.056)	(0.060)
State of Birth					
Indiana	-0.007	-0.024	0.104	-0.023	0.106
	(0.044)	(0.046)	(0.068)	(0.047)	(0.069)
Kansas	-0.108**	-0.140***	0.021	-0.138***	0.023
	(0.047)	(0.050)	(0.069)	(0.050)	(0.070)
Kentucky	-0.366***	-0.389***	-0.085	-0.384***	-0.080
	(0.036)	(0.043)	(0.067)	(0.043)	(0.067)
Maine	-0.214***	-0.205***	0.043	-0.201***	0.047
	(0.062)	(0.056)	(0.064)	(0.056)	(0.065)
Maryland	-0.047	-0.002	0.047	-0.002	0.047
	(0.055)	(0.061)	(0.083)	(0.063)	(0.085)
Massachusetts	-0.015	-0.007	0.134**	-0.011	0.129**
	(0.032)	(0.033)	(0.057)	(0.035)	(0.058)
Michigan	0.019	0.013	0.122*	0.012	0.120*
	(0.039)	(0.040)	(0.063)	(0.041)	(0.065)
Minnesota	-0.178***	-0.227***	-0.010	-0.228***	-0.011
	(0.030)	(0.043)	(0.067)	(0.044)	(0.068)
New Hampshire	-0.132	-0.124	0.118	-0.120	0.122
	(0.125)	(0.128)	(0.142)	(0.128)	(0.142)
New York	0.050	0.039	0.129**	0.039	0.129**
	(0.030)	(0.035)	(0.057)	(0.036)	(0.059)
North Carolina	-0.368***	-0.368***	0.021	-0.365***	0.025
	(0.053)	(0.051)	(0.078)	(0.051)	(0.078)
Ohio	0.035	0.027	0.121*	0.030	0.124*
	(0.036)	(0.039)	(0.067)	(0.039)	(0.067)
Pennsylvania	-0.067*	-0.038	0.117*	-0.034	0.122*
	(0.038)	(0.034)	(0.060)	(0.034)	(0.062)
Utah	0.114***	0.067	0.166**	0.069	0.169**
	(0.043)	(0.049)	(0.071)	(0.050)	(0.073)
Vermont	-0.331***	-0.335***	-0.080	-0.335***	-0.081
	(0.072)	(0.070)	(0.069)	(0.069)	(0.070)
Virginia	-0.203***	-0.203***	-0.043	-0.206***	-0.046
	(0.046)	(0.043)	(0.086)	(0.045)	(0.087)
Washington	0.116**	0.067	0.110	0.070	0.113
	(0.049)	(0.058)	(0.077)	(0.058)	(0.079)
Wisconsin	-0.105**	-0.139***	0.001	-0.135***	0.006
	(0.046)	(0.045)	(0.066)	(0.046)	(0.068)
Non-white	-0.574***	-0.574***	-0.609***	-0.574***	-0.609***
	(0.041)	(0.041)	(0.051)	(0.041)	(0.051)
Birth year =1919	0.059*	0.039	0.038	0.037	0.036
	(0.033)	(0.032)	(0.031)	(0.033)	(0.032)
Birth year =1920	0.012	-0.014	-0.009	-0.015	-0.011
	(0.013)	(0.017)	(0.017)	(0.017)	(0.017)
State of Res. Dummies	No	No	Yes	No	Yes
Observations	16240	16240	16240	16240	16240

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5
Years of Education in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	EDYRS	EDYRS	EDYRS	EDYRS	EDYRS
Maternal Mortality	-0.863**	-0.947**	-0.877**	-0.993**	-0.919**
	(0.423)	(0.378)	(0.381)	(0.382)	(0.386)
Infant Mortality Rate		-0.061*	-0.037	-0.058*	-0.034
		(0.031)	(0.032)	(0.031)	(0.032)
Total Attrition				-0.426	-0.393
				(0.319)	(0.321)
State of Birth					
Indiana	0.396***	0.339**	0.571	0.329**	0.562
	(0.143)	(0.136)	(0.369)	(0.132)	(0.365)
Kansas	0.408***	0.301**	0.012	0.288**	-0.000
	(0.152)	(0.148)	(0.348)	(0.142)	(0.344)
Kentucky	-2.108***	-2.181***	-1.676***	-2.210***	-1.701***
	(0.143)	(0.136)	(0.354)	(0.134)	(0.349)
Maine	-0.061	-0.032	0.849*	-0.055	0.822*
	(0.310)	(0.291)	(0.444)	(0.297)	(0.448)
Maryland	-0.901***	-0.756***	-0.958**	-0.756***	-0.959**
	(0.207)	(0.231)	(0.417)	(0.219)	(0.408)
Massachusetts	0.278**	0.302**	0.709*	0.329**	0.735**
	(0.134)	(0.131)	(0.369)	(0.126)	(0.366)
Michigan	-0.002	-0.020	0.351	-0.009	0.362
	(0.135)	(0.121)	(0.321)	(0.117)	(0.315)
Minnesota	-0.272**	-0.435***	-0.204	-0.426***	-0.196
	(0.119)	(0.149)	(0.396)	(0.146)	(0.393)
New Hampshire	-0.005	0.022	0.503	0.003	0.482
	(0.419)	(0.431)	(0.588)	(0.438)	(0.593)
New York	0.578***	0.542***	0.767**	0.542***	0.767**
	(0.097)	(0.092)	(0.323)	(0.084)	(0.317)
North Carolina	-1.701***	-1.704***	-1.139**	-1.726***	-1.159**
	(0.220)	(0.212)	(0.454)	(0.204)	(0.449)
Ohio	0.435***	0.410***	0.601*	0.390***	0.583*
	(0.118)	(0.113)	(0.342)	(0.106)	(0.337)
Pennsylvania	-0.151	-0.056	0.202	-0.086	0.175
	(0.120)	(0.112)	(0.289)	(0.102)	(0.282)
Utah	1.405***	1.251***	0.678*	1.236***	0.661
	(0.139)	(0.170)	(0.403)	(0.165)	(0.398)
Vermont	-0.076	-0.089	0.891*	-0.087	0.894*
	(0.340)	(0.310)	(0.521)	(0.311)	(0.520)
Virginia	-1.569***	-1.569***	-1.409***	-1.548***	-1.392***
	(0.183)	(0.164)	(0.452)	(0.160)	(0.447)
Washington	1.165***	1.005***	0.570	0.987***	0.555
	(0.167)	(0.181)	(0.407)	(0.173)	(0.407)
Wisconsin	-0.410***	-0.521***	0.001	-0.552***	-0.028
	(0.140)	(0.145)	(0.407)	(0.150)	(0.409)
Non-white	-1.690***	-1.692***	-1.721***	-1.692***	-1.720***
	(0.226)	(0.226)	(0.240)	(0.226)	(0.240)
Birth year =1919	0.276**	0.212*	0.219*	0.228**	0.233**
	(0.116)	(0.112)	(0.113)	(0.113)	(0.114)
Birth year =1920	0.217***	0.132**	0.147**	0.143***	0.157***
	(0.045)	(0.055)	(0.056)	(0.053)	(0.054)
State of Res. Dummies	No	No	Yes	No	Yes
Observations	16240	16240	16240	16240	16240

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6
High School Graduation Rate in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	HSGRAD	HSGRAD	HSGRAD	HSGRAD	HSGRAD
Maternal Mortality Rate	-0.131**	-0.136**	-0.133**	-0.146***	-0.143**
	(0.057)	(0.054)	(0.058)	(0.054)	(0.058)
Infant Mortality Rate		-0.004	-0.002	-0.003	-0.001
		(0.005)	(0.005)	(0.005)	(0.005)
Total Attrition				-0.092*	-0.092*
				(0.050)	(0.050)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	16566	16566	16566	16566	16566

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7
Not In the Labor Force in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	NILF	NILF	NILF	NILF	NILF
Maternal Mortality Rate	0.049***	0.050***	0.049***	0.050***	0.049***
	(0.017)	(0.016)	(0.016)	(0.016)	(0.017)
Infant Mortality Rate		0.001	0.001	0.001	0.001
		(0.002)	(0.002)	(0.002)	(0.002)
Total Attrition				-0.007	-0.005
				(0.016)	(0.016)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	16566	16566	16566	16566	16566

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8
Poverty Status (below 150% of poverty line) in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	POV15	POV15	POV15	POV15	POV15
Maternal Mortality Rate	0.055	0.064*	0.057	0.058*	0.051
	(0.039)	(0.033)	(0.036)	(0.034)	(0.037)
Infant Mortality Rate		0.006	0.004	0.007	0.004
		(0.004)	(0.004)	(0.004)	(0.004)
Total Attrition				-0.053	-0.053
				(0.040)	(0.040)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	16566	16566	16566	16566	16566

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9
Duncan Socioeconomic Index in 1960
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	SEI	SEI	SEI	SEI	SEI
Maternal Mortality Rate	-3.696	-3.924	-3.876	-4.030	-3.998
	(2.736)	(2.530)	(2.596)	(2.587)	(2.663)
Infant Mortality Rate		-0.166	-0.049	-0.159	-0.041
		(0.305)	(0.297)	(0.307)	(0.298)
Total Attrition				-0.971	-1.109
				(2.376)	(2.366)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	16566	16566	16566	16566	16566

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10
Number of Children Born
(to Women born 1918-1920)

	(1)	(2)	(3)	(4)	(5)
	CHBORN	CHBORN	CHBORN	CHBORN	CHBORN
Maternal Mortality Rate	-0.049**	-0.044**	-0.050**	-0.045**	-0.050**
	(0.020)	(0.021)	(0.022)	(0.021)	(0.022)
Infant Mortality Rate		0.003	0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Total Attrition				-0.443**	-0.427**
				(0.186)	(0.187)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	17058	17058	17058	17058	17058

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11
Summary Statistics of 1918-1920 Birth Cohorts:
Men Born in the United States

		Mean	Standard deviation
1980 Census Measure			
	Respondent's Age	60.2	0.93
	Total Income (1980 dollars)	$18,501	$14,274
	Years of schooling	11.5	3.4
	Graduated from high school	0.60	0.49
	Not in labor force	.27	0.44
	Below 150% Poverty Line	.11	.32
	Duncan Socioeconomic Index of Occupational Category	36.5	26.3
	Race is White	.95	.21
Pre-1980 Measure (% of live births)			
	Maternal mortality rate in state of birth	.76	.15
	Infant Mortality Rate	9.1	1.4
	Estimated attrition between birth and 1980 Census follow-up	28.6	5.4
Observations		68,580	N/A

Table 12
Not In Labor Force in 1980
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	NILF	NILF	NILF	NILF	NILF
Maternal Mortality Rate	0.058***	0.060***	0.063***	0.060***	0.063***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Infant Mortality Rate		0.002	0.003	0.002	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Total Attrition				-0.009	-0.005
				(0.049)	(0.047)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	68580	68580	68580	68580	68580

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 13
Have a Physical Disability That Prevents Work in 1980
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	DISPREV	DISPREV	DISPREV	DISPREV	DISPREV
Maternal Mortality Rate	0.032*	0.033**	0.034**	0.031*	0.032*
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)
Infant Mortality Rate		0.001	0.001	0.001	0.001
		(0.002)	(0.002)	(0.002)	(0.002)
Total Attrition				-0.023	-0.024
				(0.038)	(0.037)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	68580	68580	68580	68580	68580

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 14
Log Personal Income in 1980
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	LINCTOT	LINCTOT	LINCTOT	LINCTOT	LINCTOT
Maternal Mortality Rate	-0.095*	-0.096	-0.112*	-0.099	-0.114*
	(0.057)	(0.058)	(0.058)	(0.060)	(0.059)
Infant Mortality Rate		-0.001	-0.001	-0.001	-0.001
		(0.005)	(0.005)	(0.005)	(0.005)
Total Attrition				-0.031	-0.014
				(0.087)	(0.087)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	67118	67118	67118	67118	67118

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 15
Duncan Socioeconomic Index in 1980
Among Men Born 1918-1920

	(1)	(2)	(3)	(4)	(5)
	SEI	SEI	SEI	SEI	SEI
Maternal Mortality Rate	-2.234*	-2.658**	-3.206***	-2.487**	-3.024***
	(1.269)	(1.022)	(1.050)	(1.040)	(1.066)
Infant Mortality Rate		-0.350***	-0.379***	-0.338***	-0.366***
		(0.129)	(0.129)	(0.124)	(0.124)
Total Attrition				1.797	1.920
				(2.208)	(2.149)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	68580	68580	68580	68580	68580

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 16
Poverty Status in 1980 (Income <150% of Poverty Line)
Among Men 1918-1920

	(1)	(2)	(3)	(4)	(5)
	POV15	POV15	POV15	POV15	POV15
Maternal Mortality Rate	0.024	0.026	0.028	0.024	0.026
	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)
Infant Mortality Rate		0.002	0.002	0.001	0.002
		(0.002)	(0.002)	(0.002)	(0.002)
Total Attrition				-0.017	-0.021
				(0.029)	(0.029)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	68580	68580	68580	68580	68580

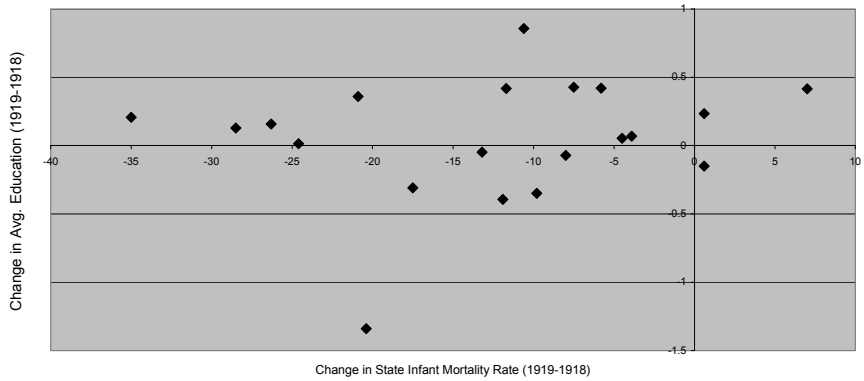
Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 17
Years of Education in 1980
Among Men Born 1918-1920

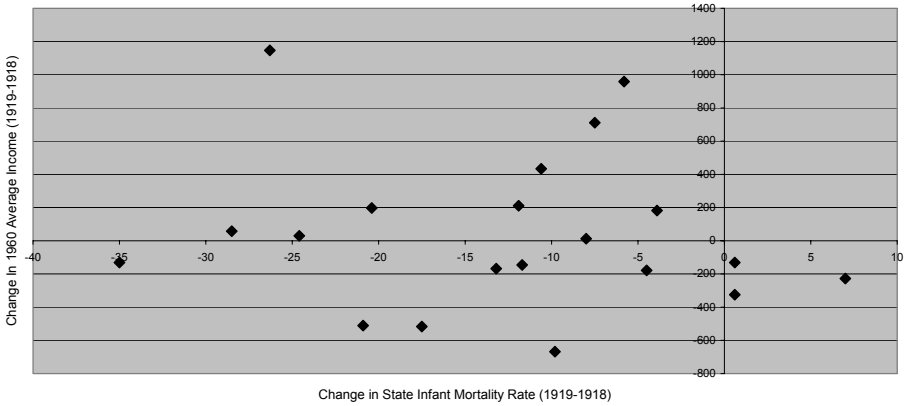
	(1)	(2)	(3)	(4)	(5)
	EDYRS	EDYRS	EDYRS	EDYRS	EDYRS
Maternal Mortality Rate	-0.127	-0.172	-0.237*	-0.187	-0.247*
	(0.146)	(0.126)	(0.134)	(0.129)	(0.134)
Infant Mortality Rate		-0.037***	-0.038***	-0.038***	-0.039***
		(0.012)	(0.014)	(0.013)	(0.014)
Total Attrition				-0.164	-0.106
				(0.293)	(0.295)
State of Residence Dummies	No	No	Yes	No	Yes
Year of Birth Dummies	Yes	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes	Yes
Observations	68580	68580	68580	68580	68580

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

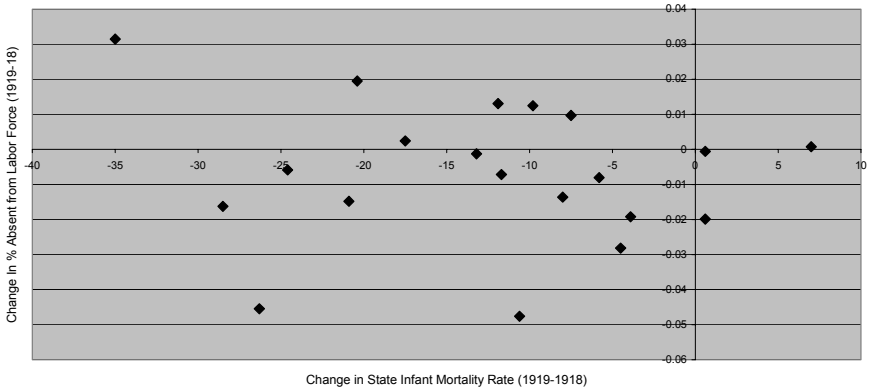
Appendix Figure 1a: Change In 1960 Average Educational Attainment
By Change in Infant Mortality Rate In State and Year of Birth
(1919-1918 difference)



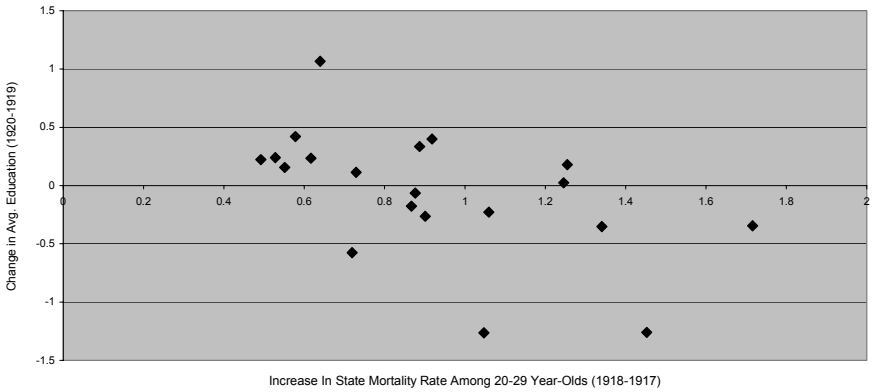
Appendix Figure 1b: Change In 1960 Average Income
By Change in Infant Mortality Rate In State and Year of Birth
(1919-1918 difference)



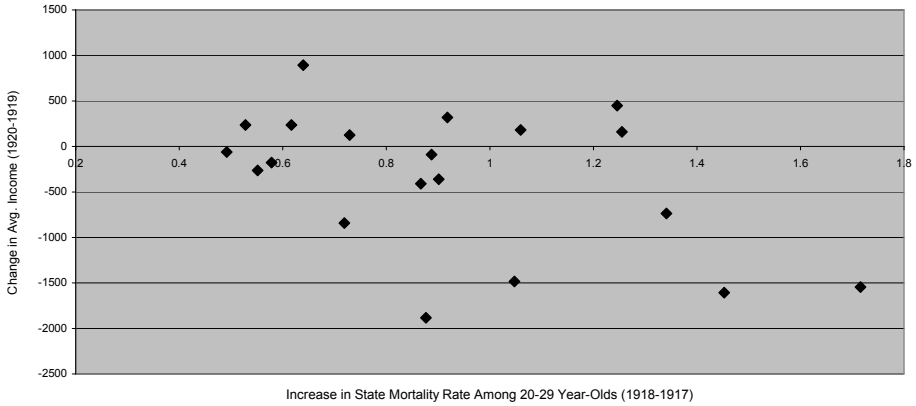
Appendix Figure 1c: Change In 1960 Share Absent From Labor Force
By Change in Infant Mortality Rate In State and Year of Birth
(1919-1918 difference)



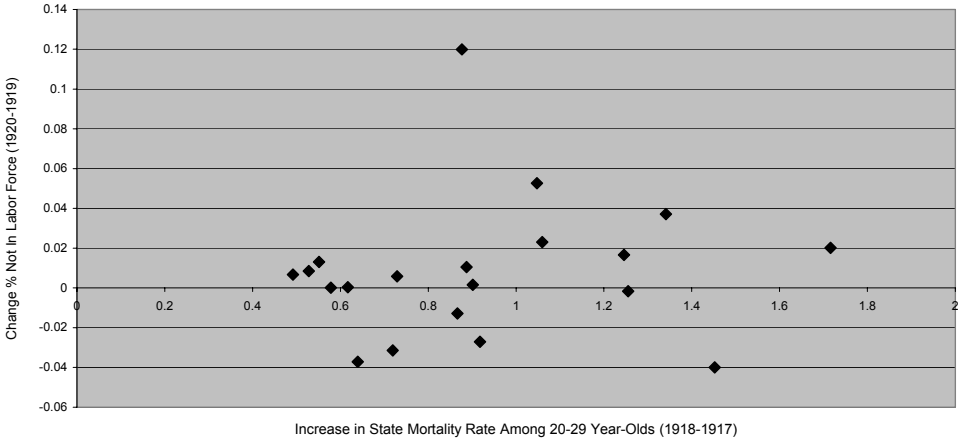
Appendix Figure 2a: 1960 Average Educational Attainment for 1920 vs.
1919 Birth Cohorts By 1918 Increase in Mortality Among 20-29 Year-Olds



Appendix Figure 2b: 1960 Average Income for 1920 vs. 1919 Birth Cohorts
By 1918 Increase in Mortality Among 20-29 Year-Olds



Appendix Figure 2c: 1960 Share Not In Labor Force for 1920 vs. 1919 Birth Cohorts
By 1918 Increase in Mortality Among 20-29 Year-Olds



Appendix Table 1
Dependent Variable: Died During 1980s
Persons Born 1918-1920 and alive in 1980

	(1)	(2)	(3)	(4)
	ALLDEAD	ALLDEAD	MENDEAD	MENDEAD
Maternal Mortality Rate	0.061	0.066*	0.106	0.113
	(0.040)	(0.037)	(0.089)	(0.087)
Female	-0.102***	-0.102***	NA	NA
	(0.010)	(0.010)		
Infant Mortality Rate		0.013**		0.021*
		(0.005)		(0.011)
Year of Birth Dummies	Yes	Yes	Yes	Yes
State of Birth Dummies	Yes	Yes	Yes	Yes
Race Dummy	Yes	Yes	Yes	Yes
Observations	6343	6343	2922	2922

Robust standard errors (clustered on state and year of birth) in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%