

Long-Term Mortality Impacts of Childhood Disease Exposure for Males: Evidence from the 1916 Polio Epidemic*

Hamid Noghanibehambari[†]

Jason Fletcher[‡]

Abstract

This paper examines the long-run impacts of exposure to the 1916 polio pandemic during early-life and childhood on later-life old-age mortality outcomes for males. We employ Social Security Administration death records linked with the 1940 census and explore the longevity differences of individuals across various ages at exposure to the 1916 pandemic in high versus low polio rate states. The results provide negative and significant impacts specifically for the first year of life and school-age children. We interpret these results as a combination of direct exposure to the disease, maternal stress, and mental pressure during prenatal and postnatal periods, limited access to health care, school closures, and lockdowns. We provide empirical evidence that childhood exposure to the pandemic is associated with reductions in education, socioeconomic scores, income, cognitive scores, and anthropometric outcomes later in life. Further, we find increases in the incidence of independent living difficulty and self-care difficulty during the late years of life. We discuss the policy implication of these findings in light of recent pandemics, specifically Covid-19.

Keywords: Mortality, Longevity, Pandemic, Early-Life Exposures

JEL Codes: I18, J13, N31, N32

* The authors claim that they have no conflict of interest to report. The authors acknowledge financial support from NIA grants (R01AG060109, R01AG076830) and valuable comments from participants in the Frontiers in Mortality, Risk, and Insurance seminar series at UW and the Southern Economic Association 2024 conference.

[†] College of Business, Austin Peay State University, Marion St, Clarksville, TN 37040, USA

Email: noghanih@apsu.edu, Phone: +1-806-620-1812

[‡] La Follette School of Public Affairs, University of Wisconsin-Madison, 1225 Observatory Drive, Madison, WI 53706-1211, USA

Email: jason.fletcher@wisc.edu

1. Introduction

It is projected that the number of Americans aged 65 and older will rise from 58 million (17% of the country's population) in 2023 to more than 82 million (23% of the projected US population) in 2050. This notable increase in the elderly demographic reflects the continuous trend of population aging that has persisted over the past several decades in the country (Mather & Scommegna, 2024). Therefore, understanding the processes and underlying factors that explain old-age health and longevity contains first-degree policy implications in various settings.

A recent and growing strand of research explores the role of early-life and childhood conditions in shaping later-life health and longevity patterns (Almond et al., 2018; Barker, 1990, 1992, 1994, 1995, 1997; Cunha & Heckman, 2007; Goodman-Bacon, 2021; Heckman et al., 2013; Heckman, 2007). This literature examines a wide range of influences during the early-life and childhood, including nutritional deprivation, environmental toxins, pollution, shocks to healthcare access, economic conditions, neighborhood violence, family adversities, and various categories of government policies (Currie et al., 2014; Fletcher, 2015; Hayward & Gorman, 2004; Lindeboom et al., 2010; Lleras-Muney et al., 2022; Van Den Berg et al., 2006).

A narrower strand of this research explores the role of early-life disease environment on later-life old-age health and mortality outcomes. For instance, Case & Paxson (2009) document that the variations in cognitive function outcomes of old people in the US can partly be explained by the variations in their birth-region-level mortality due to infectious diseases. Noghanibehambari & Fletcher (2023) show that birth state infant mortality rates are correlated with mortality rates past age 55, suggesting the role of early-life disease environment for old-age mortality. Several studies in this literature investigate the impact of pandemics experienced during in-utero and early-life on later-life disease and mortality outcomes and find mixed evidence. For instance, some studies that examine the long-term effects of in-utero exposure to the infamous 1918 influenza pandemic find deteriorations in health outcomes and increases in mortality (Almond, 2006; Almond & Mazumder, 2005b; Fletcher, 2018b; Mazumder et al., 2010) while other studies provide mixed and sometimes null effects (Fletcher, 2018c; Myrskylä et al., 2013).

The current study joins this stream of empirical research by examining the long-run impacts of early-life exposure to the 1916 polio pandemic on later-life old-age longevity. The year 1916 witnessed the largest polio pandemic in US history. The pandemic started in New York and New

Jersey and rapidly spread across the country. The pandemic claimed roughly 7,000 lives. Although about 26,000 cases were reported by authorities, the relatively small case-fatality rate of polio suggests considerably larger figures (Doshi et al., 2011; OWID, 2024). Despite the noninvasive nature of the polio virus in most cases, once the individual contracts the disease and it enters the bloodstream and nervous system, it can be fatal, especially in an environment with few available treatments. Rough estimates suggest that about 96% of children who were infected with the disease died within 2 weeks during the 1916 pandemic (BMJ, 1917). We use Social Security Administration death records (covering the years 1975-2005) linked with the full count 1940 census to examine differences in longevity of people born in high-polio versus low-polio case rate states across different ages at exposure to the 1916 pandemic. We observe significant reductions in longevity, specifically for in-utero exposure, the first year of life exposure, and exposure among school-age children. A series of heterogeneity analyses suggest larger impacts among nonwhites and children of lower socioeconomic status families. We extensively document the robustness of the results to alternative specifications, alternative functional forms, and alternative estimation models. We further provide empirical evidence to rule out that endogenous changes in sample sociodemographic and socioeconomic characteristics driven by selective survival into adulthood explain our findings.

The 1940 census suggests reductions in schooling outcomes and socioeconomic measures as a result of exposure to the pandemic. Given the growing evidence linking education and socioeconomic measures to mortality outcomes, the adverse influence of the pandemic on these outcomes may operate as mechanism channels (Fletcher, 2015; Galama et al., 2018; Lleras-Muney, 2005, 2022; Lleras-Muney et al., 2022). We further explore pathways using the World War II enlistment data and find reductions in height, cognitive score, and the probability of having a normal Body Mass Index (BMI). Moreover, using the 1980-2000 censuses and the 2001-2005 American Community Survey, we find significant reductions in household income and homeownership.

The contribution of this paper to the literature is threefold. First, this paper joins the recently growing literature on medium-run and long-run impacts of pandemics. While the immediate social and economic effects of the recent Covid-19 pandemic were unprecedented and shocking, there is a fair coverage of its short-term impacts across a variety of outcomes in the literature (Beach et al., 2022; Brodeur et al., 2021; Favara et al., 2022; Onyeaka et al., 2021).

However, studies of earlier pandemics suggest considerable adverse effects may materialize in the future for infants and children exposed to Covid-19, impacting long-term outcomes (Beach et al., 2022). Therefore, understanding the affected outcomes, affected subpopulations, and the pathways through which pandemics may influence those outcomes contain important policy implications, specifically for *ex-ante* policy designs. The 1916 polio pandemic and associated disease environment and school closure policies may provide an important case study with vital implications for the recent pandemic and inevitable future pandemics.

Second, our study adds to the ongoing and rapidly growing literature that examines the later-life impacts of early-life and childhood exposures. More specifically, we add to the narrow strand of research that examines the influence of early-life environment in shaping later-life old-age mortality outcomes (Aizer et al., 2016; Fletcher & Noghanibehambari, 2022; Fletcher & Noghanibehambari, 2024; Lindeboom et al., 2010; Noghanibehambari et al., 2022; Noghanibehambari & Fletcher, 2023a, 2023b, 2023c, 2024, 2023d; Van Den Berg et al., 2006; Yeung et al., 2014). The current paper also relates to the stream of research that evaluates the long-run mortality effects of pandemics and disease environment during early-life (Almond, 2006; Almond & Mazumder, 2005; Fletcher, 2018a, 2018b; Noghanibehambari & Fletcher, 2023e).

Third, this study also contributes to the small literature that evaluates the medium-run and long-run impacts of the 1916 polio pandemic. One notable example is the work of Meyers & Thomasson (2021) which documents that exposure during childhood and school-age periods to the pandemic is associated with reductions in educational outcomes during adulthood.

The rest of the paper is organized as follows. Section 2 provides a background on polio disease and the 1916 pandemic. Section 3 discusses data and the empirical method. Section 4 overviews the results. We conclude the paper in section 5.

2. Polio Disease

Polio, short for poliomyelitis, is a highly infectious viral disease caused by poliovirus. While its primary target is children under age five, people of all ages may contract polio. The transmission pathways include contaminated food, water, or contact with infected individuals. Poliovirus enters the body through the mouth, multiplies throughout the digestive tract, and may move to the nervous system. Most infected cases are asymptomatic or present with mild symptoms, such as sore throat, fever, nausea, headache, fatigue, and stomach pain. In some cases, more severe

symptoms appear, including paresthesia and paralysis. The disease might also infect the spinal cord and brain, resulting in meningitis and death.

While the origin of poliovirus is still an ongoing debate, studies suggest that increases in human population density may have facilitated the spread of the virus and its adaptation to humans. Before the 20th century, there were scattered historical cases and endemics that represent polio-like symptoms. The earliest dates back to an Egyptian stele (1403–1365 BC) that illustrates a priest with a withered leg. In the 19th century, there were reports of polio cases in geographically clustered areas, including the 1843 and 1841 outbreaks in Louisiana, the 1893 endemic in Boston, and the 1894 Vermont epidemic. The first polio epidemic in the 20th century occurred in New York City in 1907 with roughly 2,500 reported cases (NYNS, 1910). One plausible explanation for the observed spikes in polio pandemics during the 20th century is, ironically, that many cities at the beginning of the century embarked on ambitious projects to improve sanitation and water quality. Prior to this, individuals contracted the disease at much earlier ages especially during infancy due to exposure to contaminated water, milk, and food. As infants possessed maternal antibodies for several months after birth, the disease contraction was usually mild. The early contraction also helps the body to produce antibodies that protect individuals against later-age exposures to the disease. Improvements in sanitation and water quality, which became a necessity with rapid urbanization across US cities delayed the disease contraction to later childhood ages when maternal antibodies were no longer circulating in children's bodies. This fact contributed to the severity of the disease and also to the rise in age of infection.

All the previous and later endemics and epidemics of polio are eclipsed by the severity of the 1916 polio pandemic (Meyers & Thomasson, 2021; Trevelyan et al., 2005a). The pandemic resulted in reported polio cases of about 26,000 and claimed roughly 7,000 lives. The Top panel of Figure 1 shows the distribution of polio rates across states. New York, New Jersey, Connecticut, and Massachusetts had the highest polio per capita. The combination of widespread fear of polio and limited knowledge regarding its spread and transmission pathways resulted in various nonpharmaceutical interventions, including quarantines, lockdowns, and school closures. Due to the decentralized structure of the public health system, these interventions varied across states. For instance, Vermont and Pennsylvania postponed the starting date of the 1916-1917 school year while New Jersey left the decisions to local authorities and school boards. However, there is historical evidence that many cities, even in states with very few reported cases, implemented

school closures (Meyers & Thomasson, 2021). In highly afflicted areas, many parents forbid their children from attending school out of fear of the disease, infamously known as *infantile paralysis*.

3. Data and Method

3.1. Data

The primary source of data comes from Death Master Files (DMF) of the Social Security Administration death records. This data is extracted from the Censoc project (Goldstein et al., 2021). The DMF reports deaths that occurred to male individuals between the years 1975-2005.⁴ There are several advantages of this data to alternative ones that make it a unique source for the current study. First, the DMF data is linked to the full-count 1940 census at the individual level (extracted from Ruggles et al. (2020)). Therefore, we can observe a wide range of family-level and individual-level sociodemographic and socioeconomic characteristics that are useful for the main analysis and for the later analyses related to mechanism channels. Second, in comparison with other data that are linkable to the 1940 census and contain several thousand observations, the DMF data contained several million observations which allows us to examine an array of heterogeneity analyses. Moreover, the linked DMF-1940-census data contains the birth-state variable, which is essential in our research setting. We restrict the sample to birth cohorts of 1890 to 1930 have many cohorts with the full exposure during childhood at year 1916 and many cohorts with no exposure, i.e., born after 1916.

The prevalence of polio cases across states is extracted from Tycho (2021). The data reports weekly cases of polio for each state.^f We focus on reported cases in the year 1916 and aggregate the data at the annual level. We calculate the polio rate using the state-level population data for the year 1916, computed by linearly interpolating state-level population counts between the full count censuses of 1910 and 1920. We then merge this data with the DMF data based on state and year of birth. We also include several state-level characteristics in our analysis, which are taken from the 1880-1940 full count censuses and linearly interpolated for inter-decennial years. All full-count censuses are extracted from Ruggles et al. (2020).

⁴ In Appendix A, We use three alternative data sources (Censoc – Numident, Censoc – BUNMD, and the NCHS data) to examine the effects on females as well as exploring the concern regarding truncation of death window. Comparison of coefficients imply that the impacts are primarily concentrated among male individuals and that the truncation of death window may indeed bias the estimates downward in the true effects might be larger than those estimated using the DMF data.

Since a significant portion of these cohorts had already left their original household in 1940, we extract parental information from historical censuses 1900 – 1930. In so doing, we employ cross census linking rules provided by Price et al. (2021) in order to link individuals from 1940 census to earlier censuses. We extract parental literacy and socioeconomic index measures in the first census of observation. For those individuals that are not linked to any of previous censuses and who still reside in the original household, we use parental information in 1940. Using this method, we are able to extract parental information for about 84% of the observations in the final sample.

We restrict the sample to individuals who were born in urban areas for two reasons. First, the polio pandemic occurred mainly in larger cities with higher population density (Meyers & Thomasson, 2021; Noori et al., 2017; Paul, 1947; Trevelyan et al., 2005b). Second, polio was more likely to be endemic in rural areas, making the pandemic less severe in these communities (Groce et al., 2014; Nathanson & Kew, 2010; Trevelyan et al., 2005a).

Figure 1 illustrate the geographic distribution of polio rate in the final sample. Figure 2 shows the density distribution of age at death for the subsample of states with below-median and above-median 1916 polio rates. Descriptively, states that have above-median polio rates revealed higher longevity than states with below-median polio rates.

Summary statistics of the final sample are reported in Table 1. The sample covers birth cohorts of 1890 to 1930 who died between the years 1975 to 2005. The average age at death in the sample is 908.4 months (75.7 years). The average polio rate in the 1916 pandemic is 1.6 per 100K population. About 7.4% and 43.4% of individuals are aged $[-1,0]$ and $[1,16]$ in 1916. Roughly 96.3% and 3.5% of individuals are whites and Blacks, respectively. Among observations with non-missing values for parental characteristics, 77% have literate mothers. Using data from (Beach et al., 2022), we also calculate state-level rates of influenza and pneumonia for the years 1918-1919, covering the Spanish flu era. The data suggests 206.5 reported cases of influenza per 100K population in our final sample.

3.2. Method

Our identification strategy compares longevity outcomes of individuals across different ages at the onset of the 1916 pandemic in high polio rate states versus low polio rate states. We operationalize this comparison using the difference-in-difference equations of the following form:

$$y_{isrt} = \alpha_0 + \sum_{g=1}^{\bar{T}} \beta_g I(1916 - t = g) \times PR_s + \alpha_1 X_i + \alpha_2 Z_{st} + \xi_s + \zeta_{rt} + \varepsilon_{isrt} \quad (1)$$

The outcome is age at death (in months) of individual i born in state s and year t . The parameter g represents different groups of ages at exposure to the onset of the 1916 pandemic. The variable PR represents the state-level polio rate in 1916. To ease the interpretation, in all regressions, we normalize the polio rate with respect to the minimum and maximum values in the final sample, such that a one-unit change represents the transition from the lowest to the highest polio rate among states. The matrix X contains individual and family controls, including dummies for race and ethnicity, dummies for maternal education, and dummies for paternal socioeconomic index. The matrix Z includes state-level controls, including share of literate people, share of married people, and average socioeconomic index. Moreover, since many of the treated cohorts were also exposed to the 1918 influenza pandemic, in all regressions, we include the average state-level influenza rate for the years 1918-1919 flexibility interacted with birth year fixed effects. Birth state fixed effects are represented by ξ . We assume nonparallel trends in longevity across census regions and include region by cohort fixed effects, represented by ζ . Finally, ε is a disturbance term. We cluster standard errors on birth-state to control for serial correlations in error terms.

4. Results

4.1. Main Results

The top panel of Figure 3 shows the effects across various ages of exposure. We observe reductions in longevity for exposure around birth (i.e., age at exposure of $[-1,0]$). Although polio's main target is children under age 5, we do not observe meaningful and discernible changes in the longevity of children between ages 1-4 in states with higher reported polio rates. Nonetheless, the coefficients start to rise in magnitude for ages at exposure between 5-16, i.e., school-age children. This fact may partly reflect the negative effects of school closures, social isolation, and disruptions in the socioemotional, cognitive, and human capital development of children (Buchanan et al., 2023; Egan et al., 2021; Engzell et al., 2021). To examine the robustness of these estimates to the recent innovations in difference-in-difference estimations, we replicate the results using the method developed by Sun & Abraham (2021). These results are depicted in the bottom panel of Figure 3 and suggest a very similar pattern as those produced by ordinary least squares.

We group several ages at exposure from -1 to 16 into in-utero/early-life (i.e., age at exposure of $[-1,0]$) and childhood (i.e., age at exposure of $[1,16]$) and replicate regressions of equation 1 in Table 2. We start by reporting the results of a regression that includes only birth-state and birth-year fixed effects and individual controls in column 1. We then add additional covariates across consecutive columns. We observe fairly robust and comparable coefficients across columns. In the fully parametrized model of column 4, we find 2.2- and 1.8-months reductions in longevity of individuals with age at exposure of $[-1,0]$ and $[1,16]$ for a one-unit rise in the normalized polio rate (i.e., moving from the least to the most disease-exposed state.).

To understand the magnitudes of 1.8-2.2 months reductions in longevity, we can use the estimates in the literature from other early-life shocks. For instance, Noghanibehambari & Fletcher (2023e) document that individuals exposed between ages 0-11 to the 1918 influenza residing in high versus low influenza cities experience a reduction in longevity of about 1.8-2.7 months. Therefore, the later-life impacts of the polio pandemic are comparable to about 66-100% of the longevity impacts of the 1918 influenza, the largest pandemic in the 20th century. Noghanibehambari & Fletcher (2024) examine the effects of exposure to the Dust Bowl of the 1930s (and the associated declines in income and agricultural products) during childhood on later-life adulthood and old-age longevity. They document a reduction of about 1 month in age at death. Noghanibehambari et al. (2022) examine the influence of local labor market conditions during early-life on later-life mortality. They calculate that the drop in the US GDP between the years 1929-1933 (peak to trough of the Great Depression) was associated with a decrease in longevity of about 6.1 months. The fact that our longevity estimates lie within similar range as the effects of large and catastrophic shocks such as the Great Depression and the Dust Bowl point to the economic significance of the impacts of the 1916 polio pandemic. Further, Aizer et al. (2016) examine the long-run effects of a generous welfare program for single poor mothers that transferred cash payments equivalent to about 30-40% of their annual income. They show that children of selected mothers enjoy about one year of additional life. Therefore, the effect of polio is roughly 20% of a relatively generous cash transfer that lasted for three years.

Overall, we should note that our results reveal intent-to-treat effects across the whole population. It is possible that not all the exposed populations were affected by the pandemic, and it also could be the case that many individuals in low polio states were affected not by the direct exposure to the disease but by the resulting school closures, absenteeism, parental mental pressure,

and other potential pathways. Therefore, our results provide a floor for the true long-run effects of the pandemic.

4.2. Heterogeneity Analysis

Disease environment and the associated nonpharmaceutical interventions may affect long-run outcomes in various ways. These multifaceted mechanisms may have differential impacts across people of different subpopulations. For instance, there is evidence that black individuals are disproportionately exposed to pollution, less likely to reside in areas with better sanitation, and more likely to utilize contaminated water sources (Balazs & Ray, 2014; Tessum et al., 2019). Consequently, they face higher susceptibility to pandemics and a heightened risk of contracting diseases, as exemplified during the recent Covid-19 pandemic (Park, 2021).

In Table 3, we explore the heterogeneity of the results across different subpopulations. The coefficients of age at exposure $[-1,0]$ and $[1,16]$ in the subsample of nonwhites are 8.3 and 4.3 times that of whites (column 2 versus column 1), respectively, suggesting 15.7 and 6.5 months reductions in their longevity.

The negative effects of the pandemic may dynamically interact with other early-life adversities, such as poverty, to influence longevity. Similarly, families of higher socioeconomic status may mitigate the adverse impacts by providing additional resources to their affected children (Fletcher, 2011; Grätz & Torche, 2016). To examine this source of heterogeneity, in columns 3-4, we replicate the results for the subsample of high- and low-socioeconomic-status fathers. We observe a comparable point estimate for age at exposure of $[-1,0]$ among the two subsamples. However, we find that the effects of age at exposure of $[1,16]$ is almost exclusively concentrated in the low socioeconomic status subsample, suggesting a decrease in longevity of roughly 3 months.

Although we include controls for the Spanish flu by flexibly interacting birth year fixed effects with state level influenza rate during 1918-1919, there remains of concerns regarding the interaction of these two pandemics in determining long-run longevity. One possible argument is that the effects of the 1916 pandemics can largely reflect the effects of the 1918-1919 pandemics. In that case, we should observe much smaller coefficients in our regressions when focus on the subsample of states with lower 1918-1919 influenza rate and vice versa. In columns 5-6, we examine the heterogeneity of the results in the subsample of high and low influenza rate states.

Contrary to the line of argument regarding the confounding influence of the 1918-1919 influenza, we observe substantially larger coefficients in the subsample of low influenza rate states for age at exposure between 1-16 and comparable effects for age at exposure of $[-1,0]$.

4.3. Robustness Checks

In Table 4, we examine the robustness of the results to alternative specifications. In column 1, we replicate the main results of column 4 of Table 2. In column 2, we add birth division by birth year fixed effects to account for unobserved factors that influence outcomes and are common to specific census divisions within the same cohorts.⁶ In column 3, we include the birth state linear trend to account for secular evolution of state-level characteristics across cohorts. In both specifications, we observe coefficients comparable to the main results.

In column 4, we allow fixed effects of states to flexibly vary across sociodemographic groups. We observe identical coefficients to those of column 1. In column 5, we include birth month and death month fixed effects to control for influences of seasonality across both early and late months of life (Vaiserman, 2021; Xuan et al., 2014). We observe robust coefficients. In column 6, we replace birth state fixed effects with birth county fixed effects to account for time invariant fixed effects of place of birth deriving from below state factors that affect longevity. Although the point estimate for age at exposure of $[1,16]$ becomes statistically insignificant, the coefficient sizes are quite comparable to those of column 1.

In column 7, we replace the outcome with the log of age at death. We should note that the coefficients of this column are multiplied by 100 for ease of interpretation. We observe a very similar pattern compared to column 1. Moreover, the magnitudes of these coefficients are also similar to the implied percentage change in coefficients of column 1 with respect to the mean of the outcome. For instance, the implied percentage change of age at exposure of $[-1,0]$ in column 1 is 0.23% and its respective coefficient in column 7 suggests a 0.25% reduction in age at death. In column 8, we use an Accelerated Failure Time Model assuming an exponential distribution in the outcome. These models are suggested for studies that examine survival analyses (Aizer et al., 2016). We observe almost identical coefficients to those of column 4, which uses log of age at death. To further evaluate the robustness to functional forms, we replace the outcome with a series

⁶ Census region is used for the main regressions of equation 1. The census regions are: Northeast, Midwest, South, and West. Here we add census divisions. The census divisions are: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.

of binary variables that indicate longevity beyond ages 70-85. We then report the estimated effects for age at exposure of $[-1,0]$ and $[1,16]$ in the top left and top right panels of Figure 4. In the bottom panels, we show the implied percentage change with respect to the mean of each outcome. For the age at exposure of $[-1,0]$ group, we find comparable point estimates and implied percentage changes, although we slightly larger impacts at older ages. age at exposure of $[1,16]$, the implied percentage change estimates suggest that as we look at older ages the effects rise in magnitude.

A final concern arises from the fact that the DMF data covers a truncated death window, between 1975-2005. In Appendix A, We employ alternative data sources from the Censoc – Numident, the Berkeley Unified Numident Mortality Database, and the National Center for Health Statistics death records. We extensively provide empirical evidence to discuss that expanding the death window to cover earlier years (as early as 1970) and later years (as late as 2019) provides larger estimates and suggests that the truncation of the DMF data may underestimate the true effects (Breen et al., 2023; Breen & Goldstein, 2022).

4.4. Balancing Tests

One concern in interpreting the main results is the possible association between exposure and other unobserved determinants of longevity. If this is the case, we are also likely to observe associations between the exposure variables and observe determinants of longevity, including race, parental education, and parental socioeconomic status (Hayward & Gorman, 2004; Huebener, 2019). We empirically test this by regressing observable individual and family characteristics on exposure measures, conditional on birth state and birth region-year fixed effects. We report these results in Table 5. We observe very small and insignificant correlations between selected independent variables and the probability of being white or black (columns 1-2).

We further find small and insignificant correlations with father’s socioeconomic index, either the measure extracted from historical censuses (used in regressions) or the father’s socioeconomic index in 1940 for the observations that still reside in the original household in the 1940 census (columns 3-4). However, we observe positive correlations between age at exposure $[1,16]$ and father’s literacy status, implying a 2.5% change from the mean. Since children of higher educated fathers are healthier and live longer lives, the positive coefficient in column 5 implies downward bias in our estimates (Huebener, 2019). The correlation with mother’s literacy status provides a similar story. Interestingly, we do not observe a similar coefficient for age at exposure

[-1,0]. The correlations of columns 5 and 7 for this age group provide opposite signs, both substantially smaller in magnitude compared with the age group of [1,16]. Lack of a consistent pattern in these measures reassures us that compositional changes in parental sociodemographic characteristics are not likely a confounding source. The overall picture of this table rules out associations between the exposure measures and the observable characteristics and lends to our argument that there is little concern of any associations with unobserved characteristics (Altonji et al., 2005; Fletcher et al., 2021).

4.5. Endogenous Merging

Another concern arises from the fact that our final sample is based on death records that are linked to the 1940 census. If linked observations are different from other death records and this difference is correlated with their exposure to the 1916 pandemic as well as their longevity, the data linking becomes endogenous and all coefficients are biased. We empirically test this endogenous merging concern using the full count 1940 census. Specifically, we focus on individuals in the 1940 census born between 1890-1930 and states that are present in our final sample. We then merge this population with our final sample and generate a dummy variable indicating successful merging. We regress this successful merging indicator on our exposure measures, conditional on fixed effects and controls. The results are reported in Table 6. We do not observe significant and meaningful correlations between exposure measures and the outcome across specifications. The evidence based on this table is not strong, consistent, and convincing enough for the endogenous merging concern.

4.6. Mechanism Channels

Exposure to the pandemic may affect later-life longevity through a wide range of pathways. For instance, poliovirus may inflict pregnant mothers and be transmitted through the placental barrier to inflict the fetus. Studies show that maternal exposure during pregnancy, especially during the 3rd trimester, is associated with fetal damage, stillbirth, and even congenital polio (Ornoy & Tenenbaum, 2006). Other studies also point to the long-run impacts of polio exposure during in utero on schizophrenia during adulthood (Suvisaari et al., 1999). Further, the disease environment, public alarms and discussions of a fearsome disease, lockdowns and closures, and the very fact of contracting the disease (even without a direct influence on the fetus) might affect mental health of mothers an increase their stress level, with devastating impacts on their newborns (Berthelon et

al., 2021; Noghanibehambari, 2022; Torche, 2011). Moreover, sharp rises in sick people make hospitals and health clinics overcrowded with adverse consequences for prenatal and postnatal care of mothers and their infants, the fact that was also observed in the case of Covid-19 pandemic (Kotlar et al., 2021).

Another interesting aspect of our finding is that the impacts rise around the onset of schooling ages (i.e., age 5) and fall around the end of school ages (i.e., age 16). One possible explanation is the adverse impacts of school closures. Disruptions in schooling of children does not only affect educational outcome as documented by the literature on the long-run effects of absenteeism (Ansari & Gottfried, 2021; Ansari & Pianta, 2019; Liu et al., 2021), but also change the trajectory of children's developmental outcomes, including socioemotional and cognitive developmental outcomes (Villegas et al., 2021; Santibañez & Guarino, 2021).

The impacts of these adversities can be detected in several intermediary outcomes, including education and socioeconomic status during adulthood, which are documented to be associated with later-life mortality (Fletcher, 2015; Fletcher & Noghanibehambari, 2023; Halpern-Manners et al., 2020; Lleras-Muney, 2005; Lleras-Muney et al., 2022). We empirically explore the effects of exposure to the 1916 pandemic on these intermediary outcomes using reported variables in the 1940 census and regressions similar to equation 1. We restrict the analysis sample to individuals above age 22 to limit the number of observations with incomplete education. These results are reported in Table 7. We observe reductions in years of schooling (column 1), increases in share of individuals with less than high school education (column 2), increases in the share of people with less than 12 years of schooling (column 3), and reductions in college-educated individuals (column 4). We also observe significant reductions in socioeconomic scores (columns 5-6) and income rank based on their 1940 wages (column 7).

The magnitude of the estimated coefficients for school-age children in column 1 of table 7 implies reductions of about 0.12-0.23 years of schooling for a one-unit increase in the normalized polio rate. This is quite comparable to the estimated effects reported by Meyers & Thomasson (2021) that an increase of one standard deviation in polio morbidity rate for age at exposure of 14-17 is associated with about 0.07 years reduction in schooling.⁷

⁷ This calculation is based on the assumption that three standard deviations covers almost all the variations in polio rate.

We can understand the magnitude of these results in explaining the long-run effects on longevity by using studies that directly examine the association between these outcomes and longevity. For instance, Lleras-Muney et al. (2022) examine the effects of education on longevity using sibling fixed effect and document that each additional year of schooling is associated with roughly 5 months higher longevity. The estimated effects of column 1 for age at exposure of $[-1,0]$ and $[1,16]$ suggest 0.23 and 0.12 years lower schooling. Therefore, we can estimate reduction in longevity of about 1.2 and 0.6 months if the effects operate solely through reductions in schooling. These estimates are between 35-52 percent of the reduced form estimates of Table 2.

Chetty et al. (2016) employ population level data for the years 1999-2014 and explore the income-longevity relationship. They suggest that each additional income percentile is associated with a roughly 1.6-month rise in longevity. Based on column 7, for age at exposure of $[-1,0]$ and $[1,16]$, we observe 0.8 and 1.2 percentiles drop, equivalent to a reduction of 0.5 and 1.9 months in longevity (using estimates of Chetty et al. (2016)), respectively. Although these are back-of-an-envelope calculations, they suggest that reductions in income and education can largely explain the long-run associations.

One notable benefit of the DMF data is that it is linked to World War II enlistment data (extracted from Goldstein et al. (2023)). This data reports information on height and weight, useful to construct the Body Mass Index (BMI). The data also reports the results of a cognitive test that enrollees should take before being enlisted. We use this available information to examine additional pathways relating to the literature that documents associations between cognitive and anthropometric outcomes and longevity (Batty et al., 2007; Crimmins & Finch, 2006; Deaton, 2007; Petursson et al., 2011; Steckel, 2009).

First, we merge the DMF data with the enlistment data and generate a dummy variable that indicates successful merging and regress this variable on exposure measures. These results are reported in column 1 of Table 8. Although most coefficients suggest increases in probability of being in the enlistment data, estimates are small and statistically insignificant. In column 2, we find significant reductions in height of individuals exposed at age $[-1,0]$, suggesting a 2.5% of a standard deviation change from the mean of the outcome. In column 3, we find negative, significant, and economically meaningful reductions in the Army General Classification Test (AGCT) score, a measure of cognitive abilities and aptitudes used to assign enlistees to various

military tasks. For instance, for age at exposure of $[-1,0]$, we observe a 6-unit reduction, implying an 8% change from the mean. Finally, in column 4, we examine the effects on a dummy variable indicating the individual's BMI being in the normal range (i.e., BMI greater than 18.5 and less than 24.9). We find significant reductions in normal BMI with a significant coefficient for age at exposure of $[-1,0]$.

To further examine the mechanisms for the years post-1940, we use censuses 1980-2000 and the American Community Survey data 2001-2005 to examine the effects on different outcomes. These results are reported in Table 9. In column 1, we find significant and economically large reductions in household income. In column 2, we observe reductions in the probability of being a homeowner. The point estimate for age at exposure of $[1,16]$ suggests the change of roughly 3.7% from the mean of the outcome. In column 3, we examine the associations with average per capita income of the current state of residence, as a proxy for lifecycle neighborhood achievement. We find negative coefficients with larger effects for age at exposure of $[1,16]$.

5. Conclusion

While the 20th century witnessed eradication of many diseases and provided a seemingly disease-free environment, the recent Covid-19 pandemic, H1N1 influenza pandemic (2009-2010), and SARS outbreak (2002-2003) are examples that alarm the public and policymakers. Moreover, the research highlights a realistic scenario for emerging pandemics and unknown diseases in future (Nasajpour et al., 2020; Ristaino et al., 2021). Despite fears of pandemics and their costs to society, our knowledge is limited regarding their legacies. A strand of literature evaluates the medium run and long-run outcomes of exposed individuals, with a specific focus on exposure during critical stages of development from in utero throughout childhood (Almond, 2006; Almond et al., 2012; Almond & Mazumder, 2005a; Case & Paxson, 2009; Fletcher, 2018a, 2018b, 2018d; Noghanibehambari & Fletcher, 2023d). However, the literature that examines the long-run impacts of exposure to the pandemics on later-life old-age mortality outcomes is much more limited. This study aimed to fill this gap in the literature by investigating the effects of early-life and childhood exposure to the 1916 polio pandemic on longevity.

We used Social Security Administration death records linked with the 1940 census and implemented an identification strategy that compared longevity of individuals in high versus low polio rate states who had different ages at exposure to the 1916 pandemic. We found negative and

significant impacts for in utero and 1st year of life as well as for exposure during school ages. Comparing the least to the most disease-exposed state based on the 1916 polio rate, our estimates suggest 2.2 months reduction in longevity. We found considerably larger impacts among nonwhites and people from lower socioeconomic status families. Furthermore, in exploring potential mechanisms, we documented reductions in years of schooling, socioeconomic indicators, wage income, household income, cognitive scores, and anthropometric outcomes.

We should note that these estimates are intent-to-treat effects and measured across the whole population. The fact that the adversities of the pandemic did not affect all individuals make these estimates a baseline for the true effects. The life expectancy conditional on survival up to age 45 in the US increased from 24.69 in 1900 to 25.73 in 1930, a change of 1.04 years across these cohorts (Social Security, 2024).⁸ The observed effect of pandemic documented in this paper is roughly 17% of the overall rise in life expectancy conditional on survival up to age 45 of these cohorts. Another way to understand the magnitude of our findings is by using the Value of Statistical Life (VSL) estimates. Studies suggest a VSL of about \$10 million for US-born individuals (Kniesner & Viscusi, 2019). Using the average longevity in the final sample and the estimated VSL, we calculate the value of each additional life-month at \$11K. Therefore, based on the estimated effects of the paper, we can calculate the total loss of \$28.4 billion due to longevity reductions as a result of exposure to the pandemic.⁹

References

Aizer, A., Eli, S., Ferrie, J., & Muney, A. L. (2016). The Long-Run Impact of Cash Transfers to Poor Families. *American Economic Review*, 106(4), 935–971.
<https://doi.org/10.1257/AER.20140529>

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero

⁸ We use conditional on age 45 because this is the minimum age at death in our final sample. Further, we use cohorts of 1900 – 1930 instead of 1890 – 1930 due to data limitation for period life tables related to these cohorts.

⁹ In the final sample, 7.4% and 43.4% of observations belong to age groups of [-1,0] and [1,16], i.e., 207257 and 1215537 individuals, respectively. Using the estimated effects of Table two, we can calculate a total of 214,744 life years lost as follows: $207257 \times \frac{2.17}{12} + 1215537 \times \frac{1.75}{12} = 214,744$. We can monetize this value by assuming a VSL of \$10 million and an average longevity of 908.4, as follows: $\frac{10M}{908.4/12} \times 214744 = 28.367B$

- influenza exposure in the post-1940 US population. *Journal of Political Economy*, 114(4), 672–712. <https://doi.org/10.1086/507154>
- Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, 56(4), 1360–1446.
- Almond, D., Currie, J., & Herrmann, M. (2012). From infant to mother: Early disease environment and future maternal health. *Labour Economics*, 19(4), 475–483. <https://doi.org/10.1016/J.LABECO.2012.05.015>
- Almond, D., & Mazumder, B. (2005a). The 1918 influenza pandemic and subsequent health outcomes: {An} analysis of {SIPP} data. *American Economic Review*, 95(2), 258–262.
- Almond, D., & Mazumder, B. (2005b). The 1918 influenza pandemic and subsequent health outcomes: An analysis of SIPP data. *American Economic Review*, 95(2), 258–262.
- Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1), 151–184. <https://doi.org/10.1086/426036>
- Ansari, A., & Gottfried, M. A. (2021). The Grade-Level and Cumulative Outcomes of Absenteeism. *Child Development*, 92(4), e548–e564. <https://doi.org/10.1111/CDEV.13555>
- Ansari, A., & Pianta, R. C. (2019). School absenteeism in the first decade of education and outcomes in adolescence. *Journal of School Psychology*, 76, 48–61. <https://doi.org/10.1016/J.JSP.2019.07.010>
- Balazs, C. L., & Ray, I. (2014). The drinking water disparities framework: On the origins and persistence of inequities in exposure. *American Journal of Public Health*, 104(4), 603–611. <https://doi.org/10.2105/AJPH.2013.301664>
- Barker, D. J. P. (1990). The fetal and infant origins of adult disease. *BMJ: British Medical Journal*, 301(6761), 1111.
- Barker, D. J. P. (1992). Fetal growth and adult disease. *British Journal of Obstetrics and Gynaecology*, 99(4), 275–276. <https://doi.org/10.1111/J.1471-0528.1992.TB13719.X>
- Barker, D. J. P. (1994). *Mothers, babies, and disease in later life*. BMJ publishing group London.
- Barker, D. J. P. (1995). Fetal origins of coronary heart disease. *BMJ*, 311(6998), 171–174.

<https://doi.org/10.1136/BMJ.311.6998.171>

- Barker, D. J. P. (1997). Maternal nutrition, fetal nutrition, and disease in later life. *Nutrition*, 13(9), 807–813. [https://doi.org/10.1016/S0899-9007\(97\)00193-7](https://doi.org/10.1016/S0899-9007(97)00193-7)
- Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and Later Mortality Risk: Systematic Review. In *Annals of Epidemiology* (Vol. 17, Issue 4, pp. 278–288). Elsevier. <https://doi.org/10.1016/j.annepidem.2006.07.010>
- Beach, B., Clay, K., & Saavedra, M. (2022). The 1918 Influenza Pandemic and Its Lessons for COVID-19. *Journal of Economic Literature*, 60(1), 41–84. <https://doi.org/10.1257/JEL.20201641>
- Berthelon, M., Kruger, D., & Sanchez, R. (2021). Maternal stress during pregnancy and early childhood development. *Economics & Human Biology*, 43, 101047. <https://doi.org/10.1016/J.EHB.2021.101047>
- BMJ. (1917). The Epidemic Of Poliomyelitis In New York, 1916. *The British Medical Journal*, 2(2950).
- Breen, C. F., & Goldstein, J. R. (2022). Berkeley Unified Numident Mortality Database: Public administrative records for individual-level mortality research. *Demographic Research*, 47, 111–142. <https://doi.org/10.4054/DEMRES.2022.47.5>
- Breen, C. F., Osborne, M., & Goldstein, J. R. (2023). CenSoc: Public Linked Administrative Mortality Records for Individual-level Research. *Scientific Data* 2023 10:1, 10(1), 1–12. <https://doi.org/10.1038/s41597-023-02713-y>
- Brodeur, A., Gray, D., Islam, A., & Bhuiyan, S. (2021). A literature review of the economics of COVID-19. *Journal of Economic Surveys*, 35(4), 1007–1044. <https://doi.org/10.1111/JOES.12423>
- Buchanan, D., Hargreaves, E., & Quick, L. (2023). Schools closed during the pandemic: revelations about the well-being of ‘lower-attaining’ primary-school children. *Education* 3-13, 51(7), 1077–1090. <https://doi.org/10.1080/03004279.2022.2043405>
- Case, A., & Paxson, C. (2009). Early Life Health and Cognitive Function in Old Age. *American Economic Review*, 99(2), 104–109. <https://doi.org/10.1257/AER.99.2.104>

- Chavez Villegas, C., Peirola, S., Rocca, M., Ipince, A., & Bakrania, S. (2021). Impacts of health-related school closures on child protection outcomes: A review of evidence from past pandemics and epidemics and lessons learned for COVID-19. *International Journal of Educational Development*, 84, 102431. <https://doi.org/10.1016/J.IJEDUDEV.2021.102431>
- Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The Association Between Income and Life Expectancy in the United States, 2001–2014. *JAMA*, 315(16), 1750–1766. <https://doi.org/10.1001/JAMA.2016.4226>
- Crimmins, E. M., & Finch, C. E. (2006). Infection, inflammation, height, and longevity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), 498–503. https://doi.org/10.1073/PNAS.0501470103/SUPPL_FILE/01470FIG4.PDF
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2), 31–47. <https://doi.org/10.1257/AER.97.2.31>
- Currie, J., Zivin, J. G., Mullins, J., & Neidell, M. (2014). What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution? *Annual Reviews*, 6(1), 217–247. <https://doi.org/10.1146/ANNUREV-RESOURCE-100913-012610>
- Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences*, 104(33), 13232–13237. <https://doi.org/10.1073/PNAS.0611500104>
- Doshi, S. J., Sandhu, H. S., Venczel, L. V., Hymbaugh, K. J., Deshpande, J. M., Pallansch, M. A., Bahl, S., Wenger, J. D., & Cochi, S. L. (2011). Poliomyelitis-Related Case-Fatality Ratio in India, 2002–2006. *Clinical Infectious Diseases*, 53(1), 13–19. <https://doi.org/10.1093/CID/CIR332>
- Egan, S. M., Pope, J., Moloney, M., Hoyne, C., & Beatty, C. (2021). Missing Early Education and Care During the Pandemic: The Socio-Emotional Impact of the COVID-19 Crisis on Young Children. *Early Childhood Education Journal*, 49(5), 925–934. <https://doi.org/10.1007/S10643-021-01193-2/TABLES/1>
- Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences of the United States of America*, 118(17), e2022376118. https://doi.org/10.1073/PNAS.2022376118/SUPPL_FILE/PNAS.2022376118.SAPP.PDF

- Favara, M., Freund, R., Porter, C., Sanchez, A., & Scott, D. (2022). Young Lives, Interrupted: Short-Term Effects of the COVID-19 Pandemic on Adolescents in Low- and Middle-Income Countries. *The Journal of Development Studies*, 58(6), 1063–1080.
<https://doi.org/10.1080/00220388.2022.2029421>
- Fletcher, J., Kim, J., Nobles, J., Ross, S., & Shaorshadze, I. (2021). The effects of foreign-born peers in us high schools and middle schools. *Journal of Human Capital*, 15(3), 432–468.
https://doi.org/10.1086/715019/SUPPL_FILE/2001111APPENDIX.PDF
- Fletcher, J. M. (2011). Long-term effects of health investments and parental favoritism: the case of breastfeeding. *Health Economics*, 20(11), 1349–1361. <https://doi.org/10.1002/HEC.1675>
- Fletcher, J. M. (2015). New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. *Social Science & Medicine*, 127, 101–107.
<https://doi.org/10.1016/J.SOCSCIMED.2014.09.052>
- Fletcher, J. M. (2018a). Environmental bottlenecks in children’s genetic potential for adult socioeconomic attainments: Evidence from a health shock. *Population Studies*, 73(1), 139–148.
<https://doi.org/10.1080/00324728.2018.1498533>
- Fletcher, J. M. (2018b). Examining the long-term mortality effects of early health shocks. *Applied Economics Letters*, 26(11), 902–908.
<https://doi.org/10.1080/13504851.2018.1520960>
- Fletcher, J. M. (2018c). New evidence on the impacts of early exposure to the 1918 influenza pandemic on old-age mortality. *Biodemography and Social Biology*, 64(2), 123–126.
<https://doi.org/10.1080/19485565.2018.1501267>
- Fletcher, J. M. (2018d). The effects of in utero exposure to the 1918 influenza pandemic on family formation. *Economics & Human Biology*, 30, 59–68.
<https://doi.org/10.1016/J.EHB.2018.06.004>
- Fletcher, J. M., & Noghanibehambari, H. (2022). Toxified to the Bone: Early-Life and Childhood Exposure to Lead and Men’s Old-Age Mortality. *Working Paper*.
- Fletcher, J., & Noghanibehambari, H. (2023). The effects of education on mortality: Evidence using college expansions. *Health Economics*. <https://doi.org/10.1002/HEC.4787>
- Fletcher, J., & Noghanibehambari, H. (2024). The siren song of cicadas: Early-life pesticide

- exposure and later-life male mortality. *Journal of Environmental Economics and Management*, 123, 102903. <https://doi.org/10.1016/J.JEEM.2023.102903>
- Galama, T., Lleras-Muney, A., & Kippersluis, H. van. (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. *Oxford Research Encyclopedia of Economics and Finance*.
<https://doi.org/10.1093/ACREFORE/9780190625979.013.7>
- Goldstein, J. R., Alexander, M., Breen, C., Miranda González, A., Menares, F., Osborne, M., Snyder, M., & Yildirim, U. (2021). Censoc Project. In *CenSoc Mortality File: Version 2.0*. Berkeley: University of California. <https://censoc.berkeley.edu/data/>
- Goldstein, J. R., Breen, C., Alexander, M., Miranda González, A., Menares, F., Osborne, M., Snyder, M., Yildirim, U., & Wickle, A. (2023). *CenSoc Army Enlistment Records*.
<https://doi.org/10.7910/DVN/ZFVVNA>
- Goodman-Bacon, A. (2021). The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. *American Economic Review*, 111(8), 2550–2593. <https://doi.org/10.1257/AER.20171671>
- Grätz, M., & Torche, F. (2016). Compensation or Reinforcement? The Stratification of Parental Responses to Children’s Early Ability. *Demography*, 53(6), 1883–1904.
<https://doi.org/10.1007/s13524-016-0527-1>
- Groce, N. E., Banks, L. M., & Stein, M. A. (2014). Surviving polio in a post-polio world. *Social Science & Medicine*, 107, 171–178. <https://doi.org/10.1016/J.SOCSCIMED.2014.02.024>
- Halpern-Manners, A., Helgertz, J., Warren, J. R., & Roberts, E. (2020). The Effects of Education on Mortality: Evidence From Linked U.S. Census and Administrative Mortality Data. *Demography*, 57(4), 1513–1541. <https://doi.org/10.1007/S13524-020-00892-6>
- Hayward, M. D., & Gorman, B. K. (2004). The long arm of childhood: The influence of early-life social conditions on men’s mortality. *Demography* 2004 41:1, 41(1), 87–107.
<https://doi.org/10.1353/DEM.2004.0005>
- Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33), 13250–13255.
<https://doi.org/10.1073/PNAS.0701362104>

- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103(6), 2052–2086. <https://doi.org/10.1257/AER.103.6.2052>
- Huebener, M. (2019). Life expectancy and parental education. *Social Science & Medicine*, 232, 351–365. <https://doi.org/10.1016/J.SOCSCIMED.2019.04.034>
- Kniesner, T. J., & Viscusi, W. K. (2019). The Value of a Statistical Life. *Oxford Research Encyclopedia of Economics and Finance*. <https://doi.org/10.1093/ACREFORE/9780190625979.013.138>
- Kotlar, B., Gerson, E., Petrillo, S., Langer, A., & Tiemeier, H. (2021). The impact of the COVID-19 pandemic on maternal and perinatal health: a scoping review. *Reproductive Health*, 18(1), 1–39. <https://doi.org/10.1186/S12978-021-01070-6/TABLES/1>
- Lindeboom, M., Portrait, F., & Van Den Berg, G. J. (2010). Long-run effects on longevity of a nutritional shock early in life: The Dutch Potato famine of 1846–1847. *Journal of Health Economics*, 29(5), 617–629. <https://doi.org/10.1016/J.JHEALECO.2010.06.001>
- Liu, J., Lee, M., & Gershenson, S. (2021). The short- and long-run impacts of secondary school absences. *Journal of Public Economics*, 199, 104441. <https://doi.org/10.1016/J.JPUBECO.2021.104441>
- Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, 72(1), 189–221. <https://doi.org/10.1111/0034-6527.00329>
- Lleras-Muney, A. (2022). Education and income gradients in longevity: The role of policy. *Canadian Journal of Economics/Revue Canadienne d'économique*, 55(1), 5–37. <https://doi.org/10.1111/CAJE.12582>
- Lleras-Muney, A., Price, J., & Yue, D. (2022). The association between educational attainment and longevity using individual-level data from the 1940 census. *Journal of Health Economics*, 84, 102649. <https://doi.org/10.1016/J.JHEALECO.2022.102649>
- Mather, M., & Scommegna, P. (2024). *Aging in the United States | PRB*. Population Reference Bureau. <https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/>
- Mazumder, B., Almond, D., Park, K., Crimmins, E. M., & Finch, C. E. (2010). Lingering

- prenatal effects of the 1918 influenza pandemic on cardiovascular disease. *Journal of Developmental Origins of Health and Disease*, 1(1), 26–34.
<https://doi.org/10.1017/S2040174409990031>
- Meyers, K., & Thomasson, M. A. (2021). Can pandemics affect educational attainment? Evidence from the polio epidemic of 1916. *Cliometrica*, 15(2), 231–265.
<https://doi.org/10.1007/S11698-020-00212-3/TABLES/14>
- Myrskylä, M., Mehta, N. K., & Chang, V. W. (2013). Early life exposure to the 1918 influenza pandemic and old-age mortality by cause of death. *American Journal of Public Health*, 103(7). <https://doi.org/10.2105/AJPH.2012.301060>
- Nasajpour, M., Pouriyeh, S., Parizi, R. M., Dorodchi, M., Valero, M., & Arabnia, H. R. (2020). Internet of Things for Current COVID-19 and Future Pandemics: an Exploratory Study. *Journal of Healthcare Informatics Research*, 4(4), 325–364.
<https://doi.org/10.1007/S41666-020-00080-6/FIGURES/23>
- Nathanson, N., & Kew, O. M. (2010). From Emergence to Eradication: The Epidemiology of Poliomyelitis Deconstructed. *American Journal of Epidemiology*, 172(11), 1213–1229.
<https://doi.org/10.1093/AJE/KWQ320>
- NCHS. (2020). Multiple Cause-of-Death Data Files. In *National Center for Health Statistics*.
- Noghanibehambari, H. (2022). In utero exposure to natural disasters and later-life mortality: Evidence from earthquakes in the early twentieth century. *Social Science & Medicine*, 307, 115189. <https://doi.org/10.1016/J.SOCSCIMED.2022.115189>
- Noghanibehambari, H., & Fletcher, J. (2023a). Childhood exposure to birth registration laws and old-age mortality. *Health Economics*, 32(3), 735–743. <https://doi.org/10.1002/HEC.4643>
- Noghanibehambari, H., & Fletcher, J. (2023b). In utero and childhood exposure to alcohol and old age mortality: Evidence from the temperance movement in the US. *Economics & Human Biology*, 50, 101276. <https://doi.org/10.1016/J.EHB.2023.101276>
- Noghanibehambari, H., & Fletcher, J. (2023c). Long-Term Health Benefits of Occupational Licensing: Evidence from Midwifery Laws. *Journal of Health Economics*, 92, 102807. <https://doi.org/10.1016/J.JHEALECO.2023.102807>
- Noghanibehambari, H., & Fletcher, J. (2024). Dust to Feed, Dust to Gray: The Effect of in Utero

- Exposure to the Dust Bowl on Old-Age Longevity. *Demography*.
<https://doi.org/10.1215/00703370-11140760>
- Noghanibehambari, H., & Fletcher, J. M. (2023d). The Long Shadow of the Past: Early-Life Disease Environment and Later-Life Mortality. *SSRN Electronic Journal*.
<https://doi.org/10.2139/SSRN.4597720>
- Noghanibehambari, H., & Fletcher, J. M. (2023e). Walking with Feet Tied to the Past: Childhood Exposure to the 1918 Influenza Pandemic and Later-Life Male Mortality. *Working Paper*.
- Noghanibehambari, H., Fletcher, J., Schmitz, L., Duque, V., & Gawai, V. (2024). Early-life economic conditions and old-age male mortality: evidence from historical county-level bank deposit data. *Journal of Population Economics*, 37(1), 1–33.
<https://doi.org/10.1007/S00148-024-01007-W/TABLES/7>
- Noori, N., Drake, J. M., & Rohani, P. (2017). Comparative epidemiology of poliovirus transmission. *Scientific Reports 2017 7:1*, 7(1), 1–12. <https://doi.org/10.1038/s41598-017-17749-5>
- NYNS. (1910). *Epidemic poliomyelitis: Report on the New York epidemic of 1907* (N. Y. neurological society. Collective investigation committee (ed.); Issue 6). Johnson Reprint Corporation.
- Onyeaka, H., Anumudu, C. K., Al-Sharify, Z. T., Egele-Godswill, E., & Mbaegbu, P. (2021). COVID-19 pandemic: A review of the global lockdown and its far-reaching effects. *Science Progress*, 104(2).
https://doi.org/10.1177/00368504211019854/ASSET/IMAGES/LARGE/10.1177_00368504211019854-FIG3.JPEG
- Ornoy, A., & Tenenbaum, A. (2006). Pregnancy outcome following infections by coxsackie, echo, measles, mumps, hepatitis, polio and encephalitis viruses. *Reproductive Toxicology*, 21(4), 446–457. <https://doi.org/10.1016/J.REPROTOX.2005.12.007>
- OWID. (2024). *Polio case and death rates in the United States*. Our World in Data.
<https://ourworldindata.org/grapher/prevalence-of-polio-rates-in-the-united-states>
- Park, J. H. (2021). Who is hardest hit by a pandemic? Racial disparities in COVID-19 hardship

- in the U.S. *International Journal of Urban Sciences*, 25(2), 149–177.
<https://doi.org/10.1080/12265934.2021.1877566>
- PAUL, J. R. (1947). The Peripheral Spread of Poliomyelitis Through Rural and Urban Areas: Observations from a Recent Epidemic. *The Yale Journal of Biology and Medicine*, 19(4), 521. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2602387/>
- Petursson, H., Sigurdsson, J. A., Bengtsson, C., Nilsen, T. I. L., & Getz, L. (2011). Body Configuration as a Predictor of Mortality: Comparison of Five Anthropometric Measures in a 12 Year Follow-Up of the Norwegian HUNT 2 Study. *PLOS ONE*, 6(10), e26621. <https://doi.org/10.1371/JOURNAL.PONE.0026621>
- Price, J., Buckles, K., Van Leeuwen, J., & Riley, I. (2021). Combining family history and machine learning to link historical records: The Census Tree data set. *Explorations in Economic History*, 80, 101391. <https://doi.org/10.1016/J.EEH.2021.101391>
- Ristaino, J. B., Anderson, P. K., Bebbler, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., Martin, M. D., MacDonald, G. K., Neenan, P., Records, A., Schmale, D. G., Tateosian, L., & Wei, Q. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences of the United States of America*, 118(23), e2022239118. https://doi.org/10.1073/PNAS.2022239118/SUPPL_FILE/PNAS.2022239118.SAPP.PDF
- Ruggles, S., Flood, S., Goeken, R., Grover, J., & Meyer, E. (2020). IPUMS USA: Version 10.0 [dataset]. *Minneapolis, MN: IPUMS*. <https://doi.org/10.18128/D010.V10.0>
- Santibañez, L., & Guarino, C. M. (2021). The Effects of Absenteeism on Academic and Social-Emotional Outcomes: Lessons for COVID-19. *Educational Researcher*, 50(6), 392–400. https://doi.org/10.3102/0013189X21994488/SUPPL_FILE/SJ-PDF-1-EDR-10.3102_0013189X21994488.PDF
- Social Security. (2024). *Life Tables for the United States Social Security Area 1900-2100*. https://www.ssa.gov/oact/NOTES/as120/LifeTables_Tbl_11.html
- Steckel, R. H. (2009). Heights and human welfare: Recent developments and new directions. *Explorations in Economic History*, 46(1), 1–23. <https://doi.org/10.1016/J.EEH.2008.12.001>

- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
<https://doi.org/10.1016/J.JECONOM.2020.09.006>
- Suvisaari, J., Haukka, J., Tanskanen, A., Hovi, T., & Lönnqvist, J. (1999). Association between prenatal exposure to poliovirus infection and adult schizophrenia. *American Journal of Psychiatry*, 156(7), 1100–1102.
<https://doi.org/10.1176/AJP.156.7.1100/ASSET/IMAGES/LARGE/AO23T1.JPEG>
- Tessum, C. W., Apte, J. S., Goodkind, A. L., Muller, N. Z., Mullins, K. A., Paoella, D. A., Polasky, S., Springer, N. P., Thakrar, S. K., Marshall, J. D., & Hill, J. D. (2019). Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure. *Proceedings of the National Academy of Sciences of the United States of America*, 116(13), 6001–6006.
https://doi.org/10.1073/PNAS.1818859116/SUPPL_FILE/PNAS.1818859116.SM01.MOV
- Torche, F. (2011). The effect of maternal stress on birth outcomes: exploiting a natural experiment. *Demography*, 48(4), 1473–1491. <https://doi.org/10.1007/s13524-011-0054-z>
- Trevelyan, B., Smallman-Raynor, M., & Cliff, A. D. (2005a). The Spatial Dynamics of Poliomyelitis in the United States: From Epidemic Emergence to Vaccine-Induced Retreat, 1910–1971. *Annals of the Association of American Geographers*, 95(2), 269–293.
<https://doi.org/10.1111/J.1467-8306.2005.00460.X>
- Trevelyan, B., Smallman-Raynor, M., & Cliff, A. D. (2005b). The Spatial Structure of Epidemic Emergence: Geographical Aspects of Poliomyelitis in North-Eastern USA, July–October 1916. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 168(4), 701–722. <https://doi.org/10.1111/J.1467-985X.2005.00372.X>
- Tycho. (2021). *Project Tycho*. <https://www.tycho.pitt.edu/>
- Vaiserman, A. (2021). Season-of-birth phenomenon in health and longevity: epidemiologic evidence and mechanistic considerations. *Journal of Developmental Origins of Health and Disease*, 12(6), 849–858. <https://doi.org/10.1017/S2040174420001221>
- Van Den Berg, G. J., Lindeboom, M., Portrait, F., Berg, G. J. Van Den, Lindeboom, M., Portrait, F., den Berg, G. J., Lindeboom, M., & Portrait, F. (2006). Economic Conditions Early in

Life and Individual Mortality. *American Economic Review*, 96(1), 290–302.

<https://doi.org/10.1257/000282806776157740>

Xuan, L. T. T., Egondi, T., Ngoan, L. T., Toan, D. T. T., & Huong, L. T. (2014). Seasonality in mortality and its relationship to temperature among the older population in Hanoi, Vietnam. *Global Health Action*, 7(1). <https://doi.org/10.3402/GHA.V7.23115>

Yeung, G. Y. C., den Berg, G. J., Lindeboom, M., & Portrait, F. R. M. (2014). The impact of early-life economic conditions on cause-specific mortality during adulthood. *Journal of Population Economics*, 27(3), 895–919.

Tables

Table 1 - Descriptive Statistics

	Mean	SD	Min	Max
Death age (month)	908.349	117.619	529	1356
Birth Year	1914.131	9.613	1890	1930
Death Year	1989.835	8.756	1975	2005
Normalized Polio Rate $\times$ Age-at-Exposure [-1,0]	.009	.08	0	1
Normalized Polio Rate $\times$ Age-at-Exposure [1,16]	.053	.186	0	1
Age-at-Exposure [-1,0]	.074	.261	0	1
Age-at-Exposure [1,16]	.434	.496	0	1
1916 Polio Rate (per 100K)	1.62	3.526	.012	13.342
White	.963	.189	0	1
Black	.035	.184	0	1
Father's Socioeconomic Index	32.258	22.561	3	96
Father's Socioeconomic Index Missing	.161	.368	0	1
Father Literate	.762	.426	0	1
Mother Literate	.77	.421	0	1
Mother Literate Missing	.196	.397	0	1
State Covariates:				
Share of Literate	.949	.04	.574	.991
Share of Married People	.58	.029	.423	.65
Average Socioeconomic Index	26.864	3.664	13.906	36.053
1918-19 Flu Rate (per 100K)	206.56	91.9	16.23	361.19
1918-19 Flu Rate above-median	.464	.499	0	1
Observations	2,800,778			

Table 2 - The Impacts of in-Utero and Childhood Exposure to the 1916 Polio on Later-Life Longevity

	<i>Outcome: Age-at-Death (Months)</i>				
	(1)	(2)	(3)	(4)	(5)
Polio Rate (NRM) $\times$ Age-at-Exposure [-1,0]	-1.9*	-2***	-2***	-2.1***	-2.2***
	(1)	(.6)	(.6)	(.6)	(.5)
Polio Rate (NRM) $\times$ Age-at-Exposure [1,16]	-1.7**	-1.6*	-1.6*	-1.6*	-1.8**
	(.7)	(.9)	(.9)	(.9)	(.8)
Observations	2,800,778	2,800,778	2,800,778	2,800,778	2,800,778
R-squared	.4	.4	.4	.4	.4
Mean DV	908.4	908.4	908.4	908.4	908.4
Birth-State FE	✓	✓	✓	✓	✓
Birth-Year FE	✓	✓	✓	✓	✓
Birth-Region-Year FE		✓	✓	✓	✓
Individual/Family Controls			✓	✓	✓
1918-19 Flu by Birth-Year FE				✓	✓
State Controls					✓

Notes. Standard errors, clustered on birth-state, are in parentheses. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Table 3 - Heterogeneity in the Impacts of In-Utero and Childhood Exposure to the 1916 Polio on Later-Life Longevity

	<i>Outcome: Age-at-Death (Months), Subsamples:</i>					
	Whites	Nonwhites	High Socioeconomic Status Father	Low Socioeconomic Status Father	1918-19 Flu Rate High	1918-19 Flu Rate Low
	(1)	(2)	(3)	(4)	(5)	(6)
Polio Rate (NRM) $\times$ Age-at- Exposure [-1,0]	-1.9*** (.5)	-15.7*** (4.2)	-2.3** (1.1)	-2.1** (.8)	-1.8*** (.2)	-2 (1.2)
Polio Rate (NRM) $\times$ Age-at- Exposure [1,16]	-1.5** (.7)	-6.5*** (1.4)	.3 (1)	-2.9*** (.7)	-.8 (.8)	-3.7** (1.4)
Observations	2696888	103890	1126804	1673974	1299745	1501033
R-squared	.4	.4	.3	.4	.3	.4
Mean DV	909.1	889.1	905.7	910.1	909.6	907.3

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4 - Robustness Checks

	Column 4 of Table 2	Adding Birth-division- by-Birth-Year FE	Adding Birth-State Trend	Interacting Birth-State FE by Individual/Family Dummies
	(1)	(2)	(3)	(4)
Polio Rate (NRM) $\times$ Age-at-Exposure [-1,0]	-2.17*** (.5)	-2.42** (1.03)	-2.14*** (.5)	-2.16*** (.48)
Polio Rate (NRM) $\times$ Age-at-Exposure [1,16]	-1.75** (.78)	-1.87** (.79)	-1.64** (.75)	-1.74** (.78)
Observations	2800778	2800778	2800778	2800778
R-squared	.36	.36	.36	.36
Mean DV	908.35	908.35	908.35	908.35
	Adding Birth-Month and Death-Month FE	Adding Birth-County FE	Outcome: Log Age-at- Death	Accelerated Failure Time Model
	(5)	(6)	(7)	(8)
Polio Rate (NRM) $\times$ Age-at-Exposure [-1,0]	-2.11*** (.5)	-1.93*** (.37)	-.25*** (.06)	-.25*** (.06)
Polio Rate (NRM) $\times$ Age-at-Exposure [1,16]	-1.7** (.77)	-1.67 (1.11)	-.2** (.08)	-.2** (.08)
Observations	2800718	2800654	2800778	2800778
R-squared	.36	.36	.35	---
Mean DV	908.35	908.35	680.28	908.35

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 - Balancing Tests: In-Utero and Childhood Exposure to the 1916 Polio and Sociodemographic Characteristics

	<i>Outcomes:</i>						
	White	Black	Father's Socioeconomic Index (Historical Censuses)	Father's Socioeconomic Index (1940 Census)	Father Literate	Father Information Missing	Mother Literate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Polio Rate (NRM) × Age-at- Exposure [-1,0]	-.00198 (.00295)	.0016 (.00288)	-.30191 (.3912)	-.25746 (.46067)	.00593 (.00683)	-.00805 (.01053)	-.00084 (.0062)
Polio Rate (NRM) × Age-at- Exposure [1,16]	.00132 (.00663)	-.00145 (.00673)	-.70867 (.60604)	-.65485 (.50529)	.01956*** (.00656)	-.02022** (.00809)	.02607*** (.00618)
Observations	2800778	2800778	2349520	1122099	2800778	2800778	2800778
R-squared	.12423	.13171	.00982	.00941	.05664	.07075	.05784
Mean DV	0.963	0.035	32.258	32.971	0.762	0.196	0.770

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects and birth-year-by-birth-region fixed effects.

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Table 6 - Exploring Endogenous Merging between the Original Population in 1940 and DMF Death Records

	<i>Outcome: Successful Merging between Death Records and 1940 Census</i>		
	(1)	(2)	(3)
Polio Rate (NRM) $\times$ Age-at-Exposure [-1,0]	.00284 (.01585)	.01709 (.01714)	.01346 (.01474)
Polio Rate (NRM) $\times$ Age-at-Exposure [1,16]	.02003 (.01432)	.02687 (.01593)	.01614 (.01373)
Observations	22540790	22540790	22540790
R-squared	.0456	.04568	.06842
Mean DV	0.260	0.260	0.260
Birth-State FE	✓	✓	✓
Birth-Region-Year FE	✓	✓	✓
1918/19 Influenza Controls		✓	✓
Individual/Family/State Controls			✓

Notes. Standard errors, clustered on birth-state, are in parentheses. Individual/family controls include dummies for race and ethnicity, paternal socioeconomic status, and maternal education dummies. Birth-state controls including literacy rate, share of married people, and average socioeconomic index.

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Table 7 - Exploring Mechanism Channels Using 1940 Census

	<i>Outcomes:</i>						
	Years of Schooling	Education < High School	Years of Schooling < 12	Education: College-More	Socioeconomic Index	Occupational Income Score	Wage Income Percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Polio Rate (NRM) × Age-at- Exposure [-1,0]	-.234*** (.056)	.037** (.016)	.024*** (.003)	-.012 (.008)	-.722* (.357)	-.153* (.08)	-.787** (.379)
Polio Rate (NRM) × Age-at- Exposure [1,16]	-.12*** (.036)	.019*** (.006)	.005 (.004)	-.008* (.005)	-.71* (.417)	-.422** (.16)	-1.23*** (.303)
Observations	1831334	1863131	1863131	1863131	1700951	1736191	1500932
R-squared	.124	.113	.09	.055	.118	.091	.256
Mean DV	10.546	0.319	0.557	0.183	35.376	27.391	66.373

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index. The sample is restricted to individuals above age 21 in 1940.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8 - Exploring Mechanism Channels Using World War II Enlistment Data

	<i>Outcomes:</i>			
	Being in the Enlistment Data	Height (STD)	AGCT Score	Normal Body Mass Index
	(1)	(2)	(3)	(4)
Polio Rate (NRM) $\times$ Age-at- Exposure [-1,0]	.009 (.013)	-.025** (.009)	-6.01*** (1.27)	-.01*** (.002)
Polio Rate (NRM) $\times$ Age-at- Exposure [1,16]	.008 (.01)	-.019 (.012)	-4.16 (6.69)	-.006 (.006)
Observations	2447414	298807	34300	298242
R-squared	.12	.94	.22	.02
Mean DV	0.19	-0.17	75.04	0.73

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9 - Exploring Mechanism Channels Using Censuses 1980-2000 and American Community Survey 2001-2005

	<i>Outcomes:</i>		
	Household Income (in 2020 thousand dollars)	Home Owner	Residence State Average Income Per Capita
	(1)	(2)	(3)
Polio Rate (NRM) $\times$ Age-at-Exposure [- 1,0]	-.95 (1.24)	-.01 (.01)	-.17*** (.02)
Polio Rate (NRM) $\times$ Age-at-Exposure [1,4]	-2.56* (1.34)	-.03** (.01)	-.48*** (.06)
Observations	9669405	9669558	9837070
R-squared	.09	.04	.94
Mean DV	63.076	0.816	12.676

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Figures

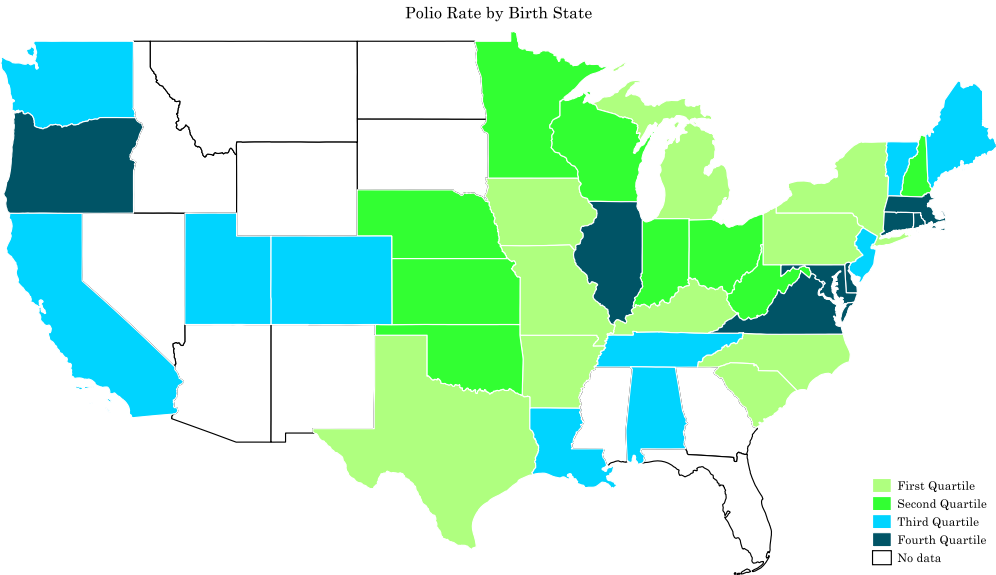


Figure 1 - Geographic Distribution of Polio Rateand Longevity across States

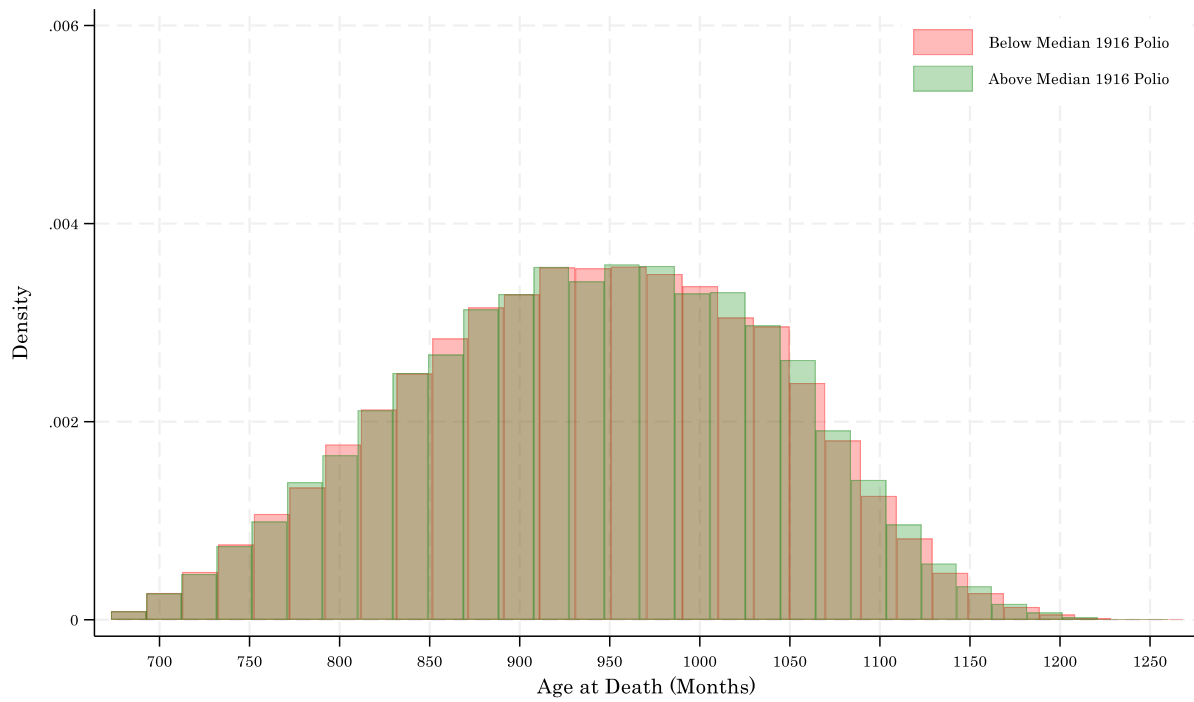


Figure 2 - Distribution of Age-at-Death across States with High and Low Polio Rate in 1916

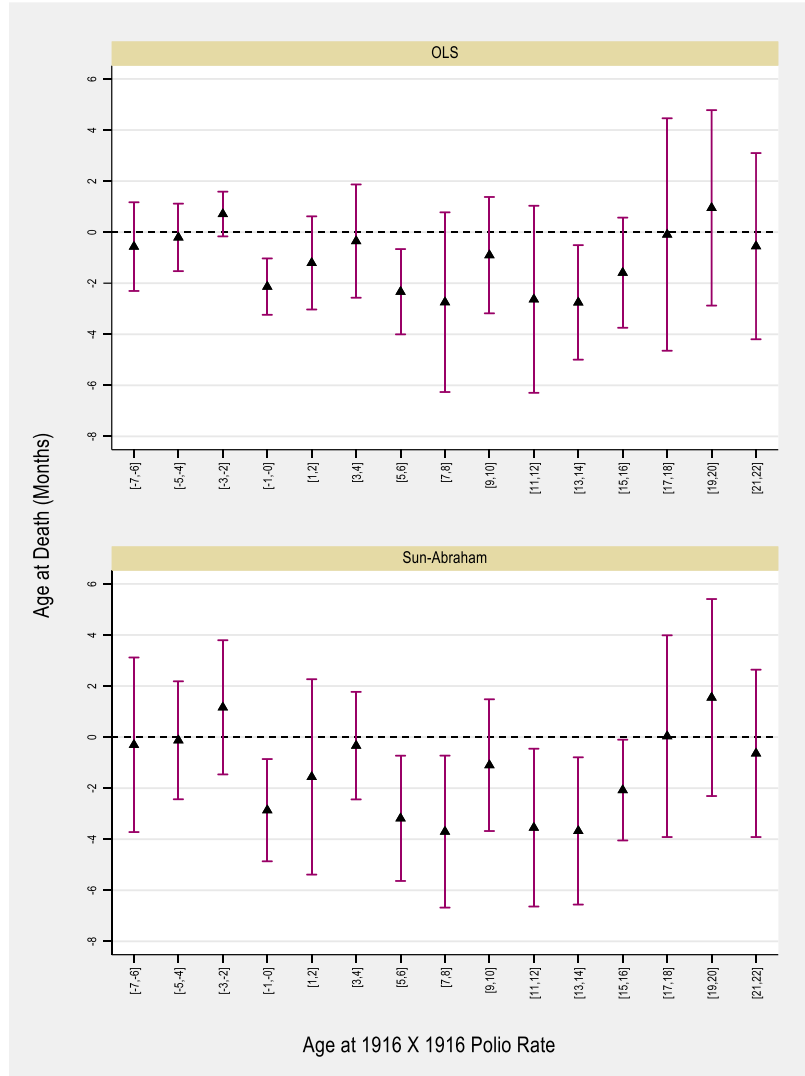


Figure 3 - Exposure to 1916 Polio Rate across Different Ages and Old-Age Longevity

Notes. Point estimates and standard errors are depicted. Standard errors are clustered on state. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

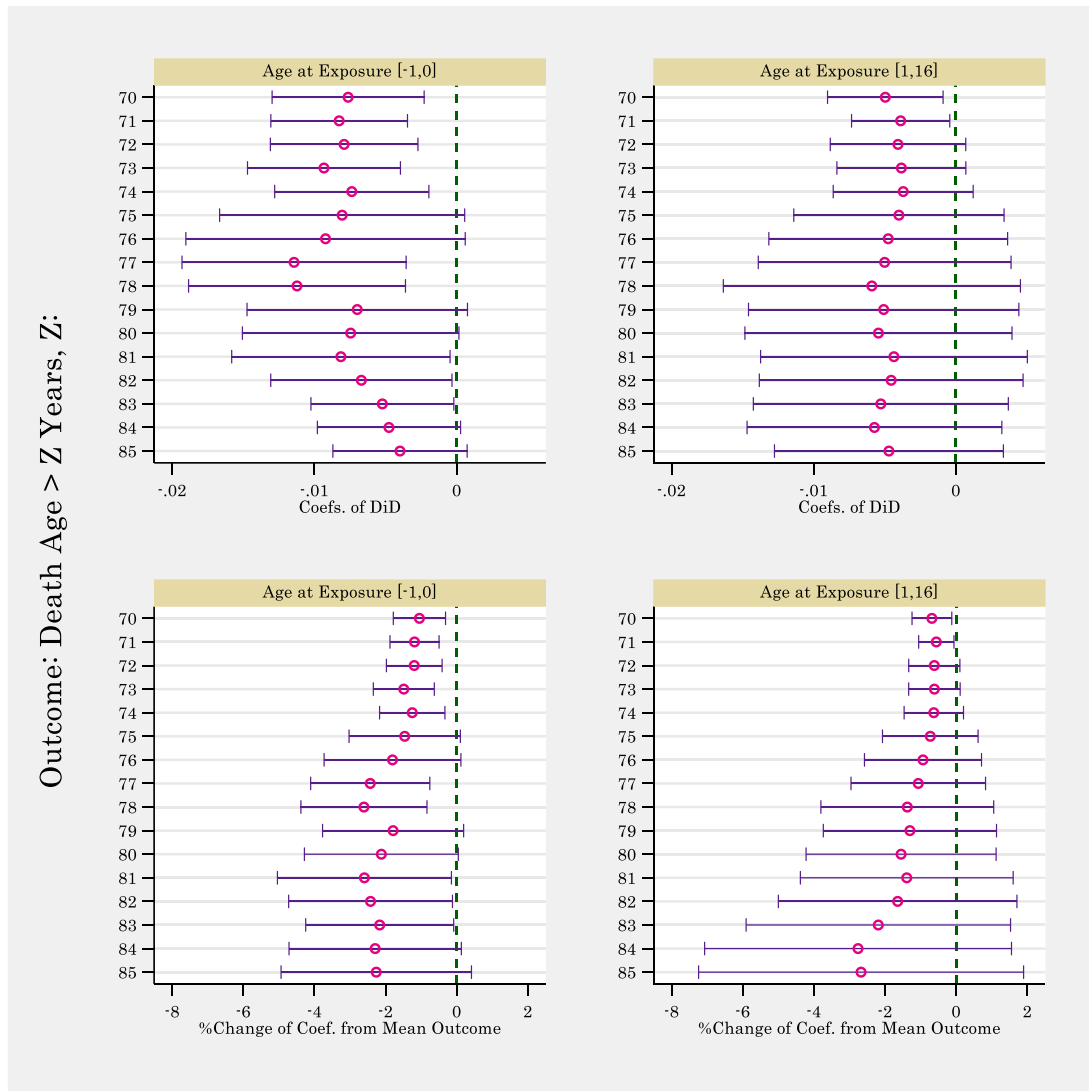


Figure 4 - Exposure to 1916 Polio Rate across Different Ages and Old-Age Longevity

Notes. Point estimates and standard errors are depicted. Standard errors are clustered on state. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

Appendix A

For the main analyses of the paper, we used the DMF data that covers death to male individuals who died between 1975 and 2005. In this appendix, we employ free alternative data sources to examine the role of double-truncation death data as well as the effects on females. These data sources are explained below.

The Censoc project also provides the Numerical Identification (so-called Numident) data that covers death to both genders. The main disadvantage of this data is that it covers deaths for the years 1988 to 2005, considerably narrower window compared with that of DMF. Another difference relates to the average longevity of each cohort in these two data sets. On average, death records of the DMF reveal higher longevity of about two – three years even for similar cohorts. After sample selections, the DMF data for death years 1980 – 2005 show 2.5 years higher longevity than the Numident data.

Second, we use the Berkeley Unified Numident Mortality Database (BUNMD) that covers deaths to both males and females (Breen & Goldstein, 2022). Although BUNMD covers much earlier death years, its coverage is considerably limited. Therefore, we focus on post-1950 death years as an arbitrary cutoff point. The data coverage is negligible for earlier years, and even up until 1970s, the data remains sparse. Further, the data is limited to pre-2007 deaths. Therefore, we are able to extract this data prior to the start of the DMF (i.e., 1975) and after the death window of the DMF closes (i.e., 2005). When we restrict the BUNMD data to DMF death years and gender, we observe an average longevity that is 1.5 years higher than the DMF data.

Third, we use the National Center for Health Statistics (NCHS) multiple cause-of-death records (NCHS, 2020). The NCHS data covers the universe of deaths in the US. While it reports post-2007 deaths, it starts to report birth state variable from 1979. Therefore, the NCHS sample

for the analysis of this appendix covers the years of 1979-2020. One notable difference between NCHS (and BUNMD) with the DMF data is that they are not linked to the 1940 census and therefore we are not able to extract parental information. Additionally, the NCHS data covers slightly older ages at death. Restricting this data to male deaths between 1979 – 2005, we observe an average age at death that is roughly 8 months higher than the DMF data for the years 1979 – 2005.

For all three data sources, we implement similar sample selection strategies and empirical method as in the main results of the paper. We draw cross data comparisons across different columns of Appendix Table A-1. First, we replicate the main results in column 1. Since the DMF data is linked to the 1940 census, we can use cross census linking to locate individuals in historical censuses and extract the urban status of place of birth. In the main results of the paper, as argued in section 3.1, we focused on those who were born in urban places. However, we cannot do this sample selection for the NCHS and BUNMD data. In column 2, we include all individuals in the DMF data and replicate the main results. The coefficients drop by about 27-39%. Therefore, it is not unreasonable to assume that the effects in the NCHS and BUNMD data could be somewhat attenuated due to this aggregation. In column 3, we restrict the starting year of the DMF data to 1979, similar to the starting year of the NCHS data. We observe a roughly 20-33% attenuation in the coefficient sizes. In column 4, we restrict the starting year of the DMF data to 1988, similar to the starting year of the Numident data. The coefficients drop by about 54 – 61% and become statistically insignificant. When we focus on the Numident data and restrict the sample to males born in urban areas, the coefficients rise in magnitude (compared with column 4) and become statistically significant (column 5). The BUNMD data for male deaths between the years 1975 – 2005 without restricting the urban places suggests a similar coefficient for age at exposure of [-

1,0] but a substantially smaller coefficient for age at exposure of [1,16] (column 6). However, when we expand the BUNMD data to the years 1950-2007, we observe a rise in the coefficients' magnitudes (column 7). Comparing columns 6 and 7 provides suggestive evidence that truncation of DMF death window might have attenuated the estimates.

In column 8-9, we show the results of male individuals in the NCHS data for death years 1979 – 2005 and 1979 – 2020, respectively. We can observe that post DMF expansion of death window significant increases the magnitude of the coefficient for age at exposure of [1,16] but reduces the coefficient size of age at exposure of [-1,0]. In the full count 1940 census, the population of cohorts born between 1890 – 1930 adds up to roughly 80M persons. In the American Community Survey data for the year 2005, there are 14M surviving individuals from these cohorts. Therefore, a large portion of these people have died by the end of the DMF death window, leaving the primary concern of truncation related to deaths prior to 1975, which we already argued that might attenuate our estimates.

In columns 10-12, we report female death in the Numident, BUNMD, and the NCHS data, respectively. In all three data sources, the coefficients are smaller in magnitude and statistically insignificant, limiting further interpretations.

Overall, although different data sets provide somewhat different estimates mostly due to differences in truncation of the death window, the overall pattern suggests that expanding the death window available in the DMF data increases the coefficient sizes and that the truncation results in underestimation of the true effects in the main results of the paper. The fact is also suggested by other studies (Breen et al., 2023; Goldstein et al., 2021; Lleras-Muney et al., 2022).

Appendix Table A-1 - Replicating the Main Results Using Combination Of City-State Polio Rate in Constructing the Exposure Measure

	<i>Outcome: Age-at-Death (Months), Data and Subsample:</i>					
	DMF	DMF	DMF	DMF	Numident	BUNMD
	Death Years: 1975-2005	Death Years: 1975-2005	Death Years: 1979-2005	Death Years: 1988-2005	Death Years: 1988-2005	Death Years: 1975-2005
	Males	Males	Males	Males	Males	Males
	Birthplace: Urban	Birthplace: All	Birthplace: Urban	Birthplace: Urban	Birthplace: Urban	Birthplace: All
	(1)	(2)	(3)	(4)	(5)	(6)
Polio Rate (NRM) × Age-at-Exposure [-1,0]	-2.2*** (.5)	-1.6*** (.4)	-1.8*** (.5)	-1 (.8)	-1.1*** (.4)	-2.4* (1.2)
Polio Rate (NRM) × Age-at-Exposure [1,16]	-1.8** (.8)	-1.1* (.6)	-1.2** (.6)	-.7 (.5)	-1.3*** (.4)	-.02 (2.8)
Observations	2800778	5856937	2364971	1605589	1708917	8475696
R-squared	.4	.4	.5	.6	.5	.3
Mean DV	908.4	916.6	923.6	949.3	919.3	890.1
	BUNMD	NCHS	NCHS	Numident	BUNMD	NCHS
	Death Years: 1950-2007	Death Years: 1979-2005	Death Years: 1979-2020	Death Years: 1988-2005	Death Years: 1950-2007	Death Years: 1979-2020
	Males	Males	Males	Females	Females	Females
	Birthplace: All	Birthplace: All	Birthplace: All	Birthplace: Urban	Birthplace: All	Birthplace: All
	(7)	(8)	(9)	(10)	(11)	(12)
Polio Rate (NRM) × Age-at-Exposure [-1,0]	-3.3** (1.3)	-.8 (1.5)	-.5 (1.1)	.07 (.5)	-.7 (1.6)	-1.1 (1.1)
Polio Rate (NRM) × Age-at-Exposure [1,16]	-1.1 (2.9)	-.9 (1.3)	-3.2 (2.3)	.4 (.5)	.9 (1.9)	-.6 (1.3)
Observations	9112918	17107758	21602075	1965531	9798224	25809393
R-squared	.3	.5	.1	.7	.5	.1
Mean DV	898.3	915.6	944.1	961.1	955.1	995.2

Notes. Standard errors, clustered on birth-state, are in parentheses. All regressions include birth-state fixed effects, birth-year-by-birth-region fixed effects, individual controls, family controls (DMF and Numident columns only), 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal literacy dummies. Regressions include a dummy variable indicating whether the state-level 1918-1919 influenza rate is above the median, interacted with birth-year fixed effects. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix Table A-2 - Summary Statistics of NCHS, BUNMD, and DMF Data for Males

	Mean	Std. Dev.	Min	Max
<i>CDC Males Death Age 1979-2020</i>				
Death Age Month	944.107	117.488	588	1416
Birth year	1916.824	9.242	1890	1930
Death year	1995.5	10.858	1979	2020
Age-at-Exposure [-1,0]×Polio Rate	.004	.044	0	1
Age-at-Exposure [1,16] ×Polio Rate	.02	.092	0	1
Age-at-Exposure [-1,0]	.071	.257	0	1
Age-at-Exposure [1,16]	.347	.476	0	1
Polio Rate	1.209	3.089	.012	20.483
White	.919	.272	0	1
Black	.077	.266	0	1
<i>BUNMD Males Death Age 1950-2007</i>				
Death Age Month	898.3	114.411	235	1367
Birth year	1920.502	6.231	1890	1930
Death year	1995.366	8.493	1950	2007
Age-at-Exposure [-1,0]×Polio Rate	.007	.055	0	1
Age-at-Exposure [1,16] ×Polio Rate	.012	.073	0	1
Age-at-Exposure [-1,0]	.106	.307	0	1
Age-at-Exposure [1,16]	.184	.388	0	1
Polio Rate	1.263	3.181	.012	20.483
White	.889	.314	0	1
Black	.081	.272	0	1
<i>DMF Males Death Age 1975-2005</i>				
Death Age Month	908.349	117.619	529	1356
Birth year	1914.131	9.613	1890	1930
Death year	1989.835	8.756	1975	2005
Age-at-Exposure [-1,0]×Polio Rate	.009	.08	0	1
Age-at-Exposure [1,16] ×Polio Rate	.053	.186	0	1
Age-at-Exposure [-1,0]	.074	.261	0	1
Age-at-Exposure [1,16]	.434	.496	0	1
Polio Rate	1.62	3.526	.012	13.342
White	.963	.189	0	1
Black	.035	.184	0	1

Number of observations for the NCHS data, BUNMD data, and DMF data are 21602075, 9112919, and 2800778, respectively.

Appendix Table A-3 - Summary Statistics of NCHS and BUNMD for Females

	Mean	Std. Dev.	Min	Max
<i>CDC Females Death Age 1979-2020</i>				
Death Age Month	995.241	114.643	588	1392
Birth year	1914.838	9.978	1890	1930
Death year	1997.775	11.074	1979	2020
Age-at-Exposure [-1,0]×Polio Rate	.004	.043	0	1
Age-at-Exposure [1,16] ×Polio Rate	.023	.099	0	1
Age-at-Exposure [-1,0]	.068	.252	0	1
Age-at-Exposure [1,16]	.389	.488	0	1
Polio Rate	1.219	3.109	.012	20.483
White	.918	.274	0	1
Black	.078	.269	0	1
<i>BUNMD Females Death Age 1950-2007</i>				
Death Age Month	955.052	111.655	246	1395
Birth year	1917.294	8.221	1890	1930
Death year	1996.887	7.159	1950	2007
Age-at-Exposure [-1,0]×Polio Rate	.007	.055	0	1
Age-at-Exposure [1,16] ×Polio Rate	.018	.09	0	1
Age-at-Exposure [-1,0]	.108	.31	0	1
Age-at-Exposure [1,16]	.31	.462	0	1
Polio Rate	1.234	3.141	.012	20.483
White	.862	.345	0	1
Black	.079	.269	0	1

Number of observations for the NCHS data and BUNMD data are 25809393 and 9798224, respectively.