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 would like to acknowledge financial support from NIA grants (R01AG060109, R01AG076830) as well as useful comments from participants at the Frontiers in Mortality, Risk and Insurance seminar series at UW.

† 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, individuals residing in urban areas, 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 increases in disability, reductions in household income, and increases in the probability of being incarcerated.

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 of 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., 2005). 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.[5] 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

[4] In Appendix A, we use the Numident data from the Censoc project to replicate the main results. This data is more restricted in its death window and covers the years 1988-2005 while it reports deaths to both females and males. 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.

[5] The Tycho database also reports polio cases for a subset of cities. There are two reasons that we prefer state-level analysis. First, the information regarding birth state is available both in the 1940 census and the DMF data while the information on birth city is not. Second, the state-level analysis produces much larger sample size that enables heterogeneity tests. However, in Appendix C, we employ hybrid analysis to exploit both city and state level variations in polio cases from this database. These estimates reveal similar patterns and point to negative impacts across childhood ages although the magnitudes are slightly larger than the main results.

of birth. We also include several state-level characteristics in our analysis, which are taken from the 1910-1920 full count censuses and linearly interpolated for inter-decennial years. All full-count censuses are extracted from Ruggles et al. (2020).

The top and bottom panels of Figure 1 illustrate the geographic distribution of polio rate and age at death based on the state of birth of individuals 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. Visually and in a cross-sectional manner, 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 916.5 months (76.4 years). The average polio rate in the 1916 pandemic is 1.3 per 100K population. About 6.9, 13.9, 18.1, and 12.8 percent of individuals are aged [-1,0], [1,4], [5,10], and [11,16] in 1916. Roughly 95.1 and 4.7 percent of individuals are whites and Blacks, respectively. Since family characteristics are observed in 1940, when many of the cohorts in the final sample had already left their original household, a significant portion of the sample have missing values for their parental characteristics. Among observations with non-missing values for parental characteristics, 33.1 and 2 percent have mothers with less than a high school education and any college education, respectively. Using data from Tycho (2021), we also calculate state-level rates of influenza and pneumonia for the years 1918-1919, covering the Spanish flu era. The data suggests 88.4 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:

$$y_{ist} = \alpha_0 + \sum_{g=1}^{\overline{T}} \beta_g I(1916 - t = g) \times PR_s + \alpha_1 X_i + \alpha_2 Z_{st} + \xi_s + \zeta_t + \varepsilon_{ist} \qquad (1)$$

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 standardize this variable using its mean and standard deviation in the final sample. 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. Finally, ε is a disturbance term. We cluster standard errors on birth-state and birth-year to control for spatial and 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 four arbitrary categories 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 slightly 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 a 0.4-month reduction in longevity of individuals with age at exposure of [-1,0], [5,10], and [11,16] for a one-standard-deviation rise in the polio rate.

The difference between the group of states in the top decile of polio rate and those in the bottom decile is roughly 9.5 cases per 100K population. This is approximately 3 times the standard deviation in the final sample. Therefore, the top-versus-bottom decile comparison suggests a 1.2-month reduction in longevity for the above-mentioned ages at exposure.

To understand the magnitude of a 1.2-month reduction 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 40-70% 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. 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 10% 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 coefficient of age at exposure [-1,0] in the subsample of nonwhites is 6.5 times that of whites (column 2 versus column 1), suggesting a 2.2-month reduction in longevity among nonwhites. However, the coefficients of age at exposure of [5,10] and [11,16] are comparable across the two subsamples. One possible story is that the effects on ages 5-16 are solely driven by the consequences of isolation and school closures and that is why we observe similar impacts among whites and nonwhites while the effect on ages [-1,0] is possibly driven by the direct exposure, hence larger effect among nonwhites.

In columns 3-4, we replicate the results among residents of rural and urban areas. We observe a considerably larger effect in urban areas for age at exposure of [-1,0] and [11,16]. This is not unexpected given the higher incidence of polio in areas with higher population density (Noori et al., 2017; Paul, 1947).

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 5-6, we replicate the results for the subsample of low- and high-socioeconomic-status fathers. We observe considerably larger coefficients among children of low socioeconomic status fathers. For instance, for age at exposure of [5,10] and [11,16], children of low socioeconomic status fathers reveal significant reductions in longevity of 0.5 and 0.3 months while children of high socioeconomic status fathers reveal insignificant decreases of about 0.08 and 0.01 months.

Although we include controls for the Spanish flu but flexibly interacting birth year fixed effects with state level influenza rate during 1918-1919, there still 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 7-8, 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 5-16. However, the effects suggest larger impacts among the high influenza rate states 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 region by birth year fixed effects to account for unobserved factors that influence outcomes and are common to specific census regions within the same cohorts. 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 replace the outcome with the log of age at death. We should note that the coefficients of this column are multiplied by 100 for the 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 a one-standard-deviation rise in polio rate for age at exposure of [-1,0] in column 1 is 0.044% and its respective coefficient in column 4 suggests a 0.045% reduction in age at death. In column 5, we replace the outcome with a binary variable that indicates longevity beyond age 80. The pattern in the estimated coefficients is similar to that of column 1. A one-standard-deviation increase in polio rate for age at exposure of [-1,0], [5,10], and [11,16] is associated with 17.4, 20.2, and 19.4 basis points reduction in the probability of longevity past age 80, equivalent to 0.49, 0.59, and 0.56 percent change with respect to the mean of the outcome, respectively.

In column 6, we use Censoc-provided weights that adjust representation of different cohorts in a way that makes the sample's cohort-specific longevity be representative of the longevity reports of the Human Mortality Database. We observe comparable point estimates to those of column 1.

While in the main results, we employ two-way clustering by birth state and birth year, in column 7, we use standard errors that are clustered on birth state. Although standard errors slightly rise in magnitude compared to those of column 1, the coefficients remain statistically significant at conventional levels.

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.

A final concern arises from the fact that the DMF data covers a truncated death window, between 1975-2005. In Appendix B, we show 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.

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 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 some associations with father's schooling. However, the implied changes with respect to the mean of the outcome suggest quite small impacts. For instance, a one-standard-deviation rise in polio rate for age at exposure of [5,10] and [11,16] is associated with 0.3 and 0.4 percent change in the outcome (column 4). Moreover, we find small and insignificant changes for father socioeconomic index (column 4). We observe negative and significant associations for father socioeconomic index being missing, which likely picks up on the fact that many of cohorts in the control group (i.e., born after 1916) still live in the original household and for whom father's characteristics are non-missing. 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. In column 2, we observe statistical associations for age at exposure between 5-16. However, the implied percentage changes are quite small with respect to the outcome mean, suggesting changes of about 0.6-1 percent. Further, we do not observe discernible and significant changes for age at exposure of [-1,0] or [1,4]. 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). One noticeable difference in the pattern of these coefficients and those of the main results is that the coefficient of age at exposure [1,4] becomes meaningful, significant, and comparable to other ages.

The magnitude of the estimated coefficients for school-age children in column 1 of table 7 implies reductions of about 0.04-0.08 years of schooling for a one standard deviation increase in 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.

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, Fletcher & Noghanibehambari (2023) examines the effects of college openings on college education and mortality and estimate that college-educated individuals leave about a one-year additional life. Based on column 4 of Table 7, a one-standard-deviation change in polio rate during in utero and 1st year of life is associated with about 1 percentage point reduction in the probability of college education. Based on Fletcher & Noghanibehambari (2023)'s estimates, this change results in a reduction in longevity by about 0.11 months.

Halpern-Manners et al. (2020) uses twin fixed effects strategy to examine the education-longevity relationship and find that an additional year of schooling is associated with roughly 4 months higher longevity. The estimated effects of column 1 for age at exposure of [5,10] and [11,16] suggest 0.08 and 0.04 years lower schooling. Therefore, we can estimate reduction in longevity of about 0.3 and 0.16 months if the effects operate solely through reductions in schooling. These estimates are between 40-80 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 [5,10] and [11,16], we observe 0.3 and 0.2 percentiles drop, equivalent to a reduction of 0.48 and 0.32 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.

A narrow strand of research points to a link between incarceration and post-release health and longevity (Daza et al., 2020). Incarceration could disrupt the trajectory of skill development and impact lifecycle employment and occupational standing with potential implications for health outcomes. In column 8, we find some evidence of increases in the probability of being incarcerated for age at exposure of [1,4]. The point estimate suggests an increase of about 16% with respect to the mean of the outcome.

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 reductions in probability of being in the enlistment data, estimates are small and statistically insignificant. In column 2, we find small and insignificant reductions in height of exposed individuals. In column 3, we find negative, significant, and economically meaningful reductions in the cognitive score. For instance, a one-standard-deviation change in polio rate for age at exposure of [-1,0] is associated with 1.8 units drop in cognitive score, equivalent to 2.4% change with respect to the mean of the outcome. 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 across different ages at exposure. For instance, a one-standard-deviation change in polio rate for age at exposure of [-1,0] results in roughly 0.4 percentage points decreases in the probability of normal BMI, implying about 0.5% change with respect to the mean of the outcome.

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 columns 1-2, we observe increases in the incidences of independent living difficulty and self-care difficulty. These effects are more concentrated among ages at exposure of 5-16. For instance, the coefficient of age at exposure [11,16] suggests increases in the probability of independent living difficulty and self-care difficulty by 4 and 3.6 percentage points, equivalent to 21 and 28 percent change, respectively. These results, combined with the literature that links disability with mortality, point to possible pathways between early-life pandemic exposure and later-life longevity (Pongiglione et al., 2016). We also find significant and economically large reductions in household income (column 3). In column 4 of Table 9, 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 [5,10] and [11,16].

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 the 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 is 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 top decile versus bottom decile states based on the 1916 polio rate, our estimates suggest 1.2 months reductions in longevity. We found considerably larger impacts among nonwhites, residents of urban areas, and people from lower socioeconomic status families. Further, in search of mechanism channels, we documented reductions in years of schooling, socioeconomic measures, wage income, household income, cognitive score, and measures of anthropometric outcomes. We also found increases in disability related to self-care difficulty and independent living difficulty, observed in late-life years.

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 in the US increased from 45.5 in 1890 to 58 in 1930, a change of 12.5 years across cohorts in the final sample. The observed effect of pandemic is roughly 1% of the overall rise in life expectancy 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 that the exposure to the pandemic (top-versus-bottom decile of exposure) cost roughly $13,093 per person. In our final sample, roughly 2.2 million individuals were exposed during early-life and childhood to the 1916 pandemic. Therefore, we estimate $28.8 billion lost due to longevity reductions as a result of exposure to the pandemic.

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 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

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).

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., Peirolo, 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

Daza, S., Palloni, A., & Jones, J. (2020). The Consequences of Incarceration for Mortality in the United States. *Demography*, *57*(2), 577–598. https://doi.org/10.1007/S13524-020-00869-5

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/200111APPENDIX.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 socio-economic attainments: Evidence from a health shock. *Population Studies* , *73*(1), 139–148. https://doi.org/10.1080/00324728.2018.1498533

Fletcher, J. M. (2018b). 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. (2018c). 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. (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

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

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

Pongiglione, B., De Stavola, B. L., Kuper, H., & Ploubidis, G. B. (2016). Disability and all-cause mortality in the older population: evidence from the English Longitudinal Study of Ageing. *European Journal of Epidemiology*, *31*(8), 735–746. https://doi.org/10.1007/S10654-016-0160-8/TABLES/4

Ristaino, J. B., Anderson, P. K., Bebber, 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

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., Paolella, 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. (2005). 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

Tycho. (2021). *Project Tycho*. https://www.tycho.pitt.edu/

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

Yeung, G. Y. C., den Berg, G. J., Lindeboom, M., & Portrait, F. R. M. (2014). The impact of early-life economic conditionson 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)	916.484	117.98	529	1356
Birth Year	1913.04	10	1890	1930
Death Year	1989.422	8.725	1975	2005
STD Polio Rate × Age-at-Exposure [-1,0]	.001	.26	-.389	6.11
STD Polio Rate × Age-at-Exposure [1,4]	.002	.375	-.389	6.367
STD Polio Rate × Age-at-Exposure [5-10]	.006	.453	-.388	6.753
STD Polio Rate × Age-at-Exposure [11,16]	.008	.408	-.388	7.163
Age-at-Exposure [-1,0]	.069	.253	0	1
Age-at-Exposure [1,4]	.139	.346	0	1
Age-at-Exposure [5-10]	.181	.385	0	1
Age-at-Exposure [11,16]	.128	.334	0	1
1916 Polio Rate (per 100K)	1.292	3.265	.01	26.537
White	.951	.217	0	1
Black	.047	.211	0	1
Father's Socioeconomic Index Missing	.594	.491	0	1
Father's Socioeconomic Index 1st Quartile	.112	.316	0	1
Father's Socioeconomic Index 2nd Quartile	.104	.305	0	1
Father's Socioeconomic Index 3rd Quartile	.094	.292	0	1
Father's Socioeconomic Index 4th Quartile	.096	.295	0	1
Mother Education < High School	.331	.471	0	1
Mother Education = High School	.098	.297	0	1
Mother Education > High School	.02	.14	0	1
Mother Education Missing	.551	.497	0	1
State Covariates:				
Share of Literate	.937	.057	.615	.991
Share of Married People	.584	.028	.481	.65
Average Socioeconomic Index	25.176	4.206	13.906	36.053
1918-19 Flu Rate (per 100K)	88.388	80.348	1	235
Observations		5,802,813		

Notes. STD stands for standardized variable with respect to mean and standard deviation over the final sample.

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

	Outcome: Age-at-Death (Months)			
	(1)	(2)	(3)	(4)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.3383*	-.37136*	-.42255**	-.40614**
	(.19181)	(.1916)	(.2058)	(.18202)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.02403	-.05625	-.0546	-.04235
	(.16471)	(.165)	(.16489)	(.16234)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.44615***	-.46223***	-.36974***	-.36762***
	(.10351)	(.10563)	(.10697)	(.10376)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.53653***	-.54849***	-.37329***	-.38729***
	(.10456)	(.10569)	(.11113)	(.09802)
Observations	5802813	5802813	5802813	5802813
R-squared	.38057	.38093	.38098	.38101
Mean DV	916.484	916.484	916.484	916.484
Birth-State FE	✓	✓	✓	✓
Birth-Year FE	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓
Family Controls		✓	✓	✓
1918-19 Flu by Birth-Year FE			✓	✓
State Controls				✓

Notes. Standard errors, clustered on birth-state-year, are in parentheses. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal education 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

	Outcome: Age-at-Death (Months), Subsamples:							
	Whites	Nonwhites	Rural	Urban	Low Socioeconomic Status Father	High Socioeconomic Status Father	1918-19 Flu Rate High	1918-19 Flu Rate Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.3426**	-2.22778***	-.22493	-.52353***	-.83168	-.48249***	-.37295*	-.10282
	(.17151)	(.76407)	(.35173)	(.16759)	(3.05132)	(.17472)	(.2102)	(.30913)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.03575	-.26563	.12035	-.13608	-.07763	-.03275	.06949	.18541
	(.16658)	(.69663)	(.2056)	(.1948)	(.25028)	(.18584)	(.14775)	(.31327)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.36113***	-.45013	-.37485**	-.37645***	-.52047***	-.08643	-.1037	-1.01815***
	(.11078)	(.57505)	(.17312)	(.12298)	(.19262)	(.11741)	(.10567)	(.35608)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.37431***	-.362	-.25695*	-.47552***	-.32179*	-.0098	-.15292	-1.0944***
	(.09659)	(.58474)	(.13663)	(.13802)	(.16889)	(.14185)	(.09721)	(.28127)
Observations	5516253	286560	2579034	3223779	2908721	2894092	2833952	2968861
R-squared	.37772	.4243	.39671	.36817	.38387	.29646	.37258	.38892
Mean DV	917.476	897.391	915.182	917.526	946.198	886.620	917.140	915.858

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year 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 education dummies. Regressions include state-level 1918-1919 influenza rate 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-Region-by-Birth-Year FE	Adding Birth-State Trend	Outcome: Log Age-at-Death
	(1)	(2)	(3)	(4)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.40614**	-.3681***	-.41169**	-.04491**
	(.18202)	(.12028)	(.19354)	(.02028)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.04235	.01036	-.06443	-.00266
	(.16234)	(.14476)	(.16025)	(.01783)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.36762***	-.30336***	-.40832***	-.03682***
	(.10376)	(.10694)	(.09441)	(.01158)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.38729***	-.3436***	-.42622***	-.03677***
	(.09802)	(.09054)	(.09455)	(.01067)
Observations	5802813	5802813	5802813	5802813
R-squared	.38101	.38106	.38106	.37373
Mean DV	916.484	916.484	916.484	681.181

	Outcome: Age-at-Death ≥ 80	Using Censoc Weights	Standard Errors Clustered on Birth-State	Accelerated Failure Time Model
	(5)	(6)	(7)	(8)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.00174*	-.33592**	-.40614*	-.045**
	(.001)	(.16928)	(.21284)	(.022)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.00078	-.04217	-.04235	-.004
	(.00088)	(.17326)	(.14738)	(.016)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.00202***	-.40992***	-.36762*	-.037*
	(.00055)	(.10337)	(.18175)	(.02)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.00194***	-.42133***	-.38729**	-.038**
	(.00062)	(.09501)	(.14752)	(.016)
Observations	5802814	4789594	5802813	5802813
R-squared	.27288	.22195	.38101	---
Mean DV	0.342	936.635	916.484	916.484

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year 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 education dummies. Regressions include state-level 1918-1919 influenza rate 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

	Outcomes:				
	White	Black	Father's Years of Schooling	Father's Socioeconomic Index	Father's Socioeconomic Index Missing
	(1)	(2)	(3)	(4)	(5)
Polio Rate (STD) × Age-at-Exposure [-1,0]	.00001	-.00007	-.07804***	-.01816	-.01678***
	(.00045)	(.00043)	(.02443)	(.08602)	(.00414)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.00004	-.00005	-.05236***	-.07103	-.01372***
	(.00043)	(.00043)	(.0176)	(.06167)	(.00222)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.00038	.00017	-.0222*	.00645	-.00473***
	(.00031)	(.00031)	(.01256)	(.08857)	(.00122)
Polio Rate (STD) × Age-at-Exposure [11,16]	.00023	-.00036	.0309*	.1001	-.00097
	(.00036)	(.00036)	(.01663)	(.1229)	(.00123)
Observations	5802814	5802814	2300740	2118768	5802814
R-squared	.13383	.14146	.04608	.02879	.44525
Mean DV	0.951	0.047	7.187	27.143	0.594

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

	Outcome: Successful Merging between Death Records and 1940 Census	
	(1)	(2)
Polio Rate (STD) × Age-at-Exposure [-1,0]	.00118	.00099
	(.00095)	(.00076)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.00062	-.00084
	(.00084)	(.00066)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.00082	-.00103**
	(.00056)	(.00046)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.0017**	-.00182**
	(.00067)	(.00076)
Observations	35752953	35752953
R-squared	.02011	.02143
Mean DV	0.160	0.160
Birth-State FE	✓	✓
Birth-Year FE	✓	✓
Controls		✓

Notes. Standard errors, clustered on birth-state-year, are in parentheses. Controls include dummies for race and ethnicity and 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

	Outcomes:							
	Years of Schooling	Education < High School	Years of Schooling < 12	Education: College-More	Socioeconomic Index	Occupational Income Score	Wage Income Percentile	Being Incarcerated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.13951***	.02312***	.02038***	-.00917***	-.72266***	-.20618***	-.39613***	-.00006
	(.02348)	(.00362)	(.00244)	(.00135)	(.14493)	(.05815)	(.12501)	(.00007)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.12243***	.0214***	.0165***	-.00757***	-.57803***	-.18516***	-.35939***	.00016**
	(.01721)	(.00292)	(.00172)	(.00115)	(.09335)	(.03864)	(.06747)	(.00007)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.08464***	.01544***	.01066***	-.00445***	-.41313***	-.16424***	-.28914***	.0001
	(.01216)	(.00221)	(.00136)	(.00077)	(.06635)	(.02879)	(.05158)	(.00006)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.04119***	.00815***	.00528***	-.00207***	-.24123***	-.09029***	-.20551***	.00003
	(.0128)	(.00243)	(.00134)	(.00071)	(.05675)	(.02539)	(.05212)	(.00006)
Observations	3586436	3655169	3655169	3655169	3436224	3495321	2713454	3655169
R-squared	.14236	.10138	.07673	.04321	.08758	.10337	.1929	.00226
Mean DV	9.740	0.438	0.645	0.141	30.145	24.840	60.730	0.001

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year 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 education dummies. Regressions include state-level 1918-1919 influenza rate 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 8 - Exploring Mechanism Channels Using World War II Enlistment Data

	Outcomes:			
	Being in the Enlistment Data	Height	Cognitive Score	Normal Body Mass Index
	(1)	(2)	(3)	(4)
Polio Rate (STD) × Age-at-Exposure [-1,0]	.00261	-.00441	-1.79426**	-.004***
	(.00228)	(.01165)	(.8037)	(.00083)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.00035	-.00482	-1.34021*	-.00342***
	(.00158)	(.01041)	(.70672)	(.00124)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.00073	.00385	-1.09677*	-.00368***
	(.00151)	(.01232)	(.64054)	(.00134)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.00178	-.02172	.57848	-.00304
	(.00174)	(.01606)	(1.7552)	(.0027)
Observations	5137027	573861	101546	572844
R-squared	.13865	.04809	.22822	.0256
Mean DV	0.162	68.323	76.476	0.756

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year 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 education dummies. Regressions include state-level 1918-1919 influenza rate 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

	Outcomes:			
	Disability: Independent living difficulty	Disability: Self-care difficulty	Household Income (in 2020 thousand dollars)	Residence State Average Income Per Capita
	(1)	(2)	(3)	(4)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.00496	-.00681	-2.98341*	-.16886**
	(.00604)	(.00475)	(1.58998)	(.06617)
Polio Rate (STD) × Age-at-Exposure [1,4]	.00497	.0009	-3.16354*	-.28415***
	(.00866)	(.00606)	(1.70058)	(.06628)
Polio Rate (STD) × Age-at-Exposure [5-10]	.00773	.01354*	-3.83084***	-.59603***
	(.00939)	(.00713)	(1.28347)	(.22806)
Polio Rate (STD) × Age-at-Exposure [11,16]	.04045***	.03615**	-3.98401***	-1.05554***
	(.01378)	(.01785)	(1.50451)	(.15159)
Observations	6040724	6040728	9967954	9680409
R-squared	.13178	.07638	.07823	.65403
Mean DV	0.190	0.129	62.815	12.688

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year fixed effects, individual controls, 1918-1919 flu rate controls, and state controls. Individual controls include dummies for race and ethnicity. Regressions include state-level 1918-1919 influenza rate 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

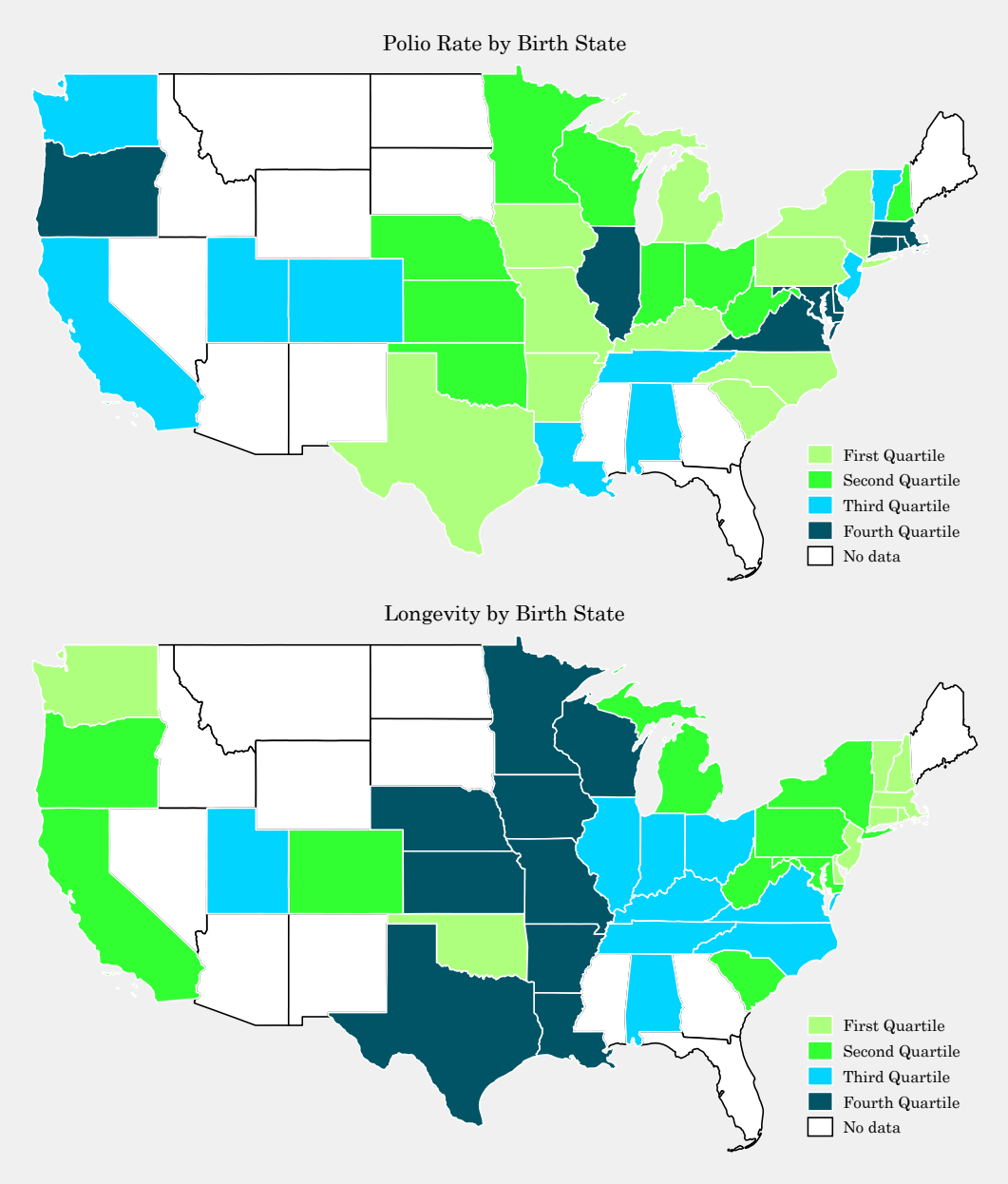

Figure 1 - Geographic Distribution of Polio Rateand Longevity across States

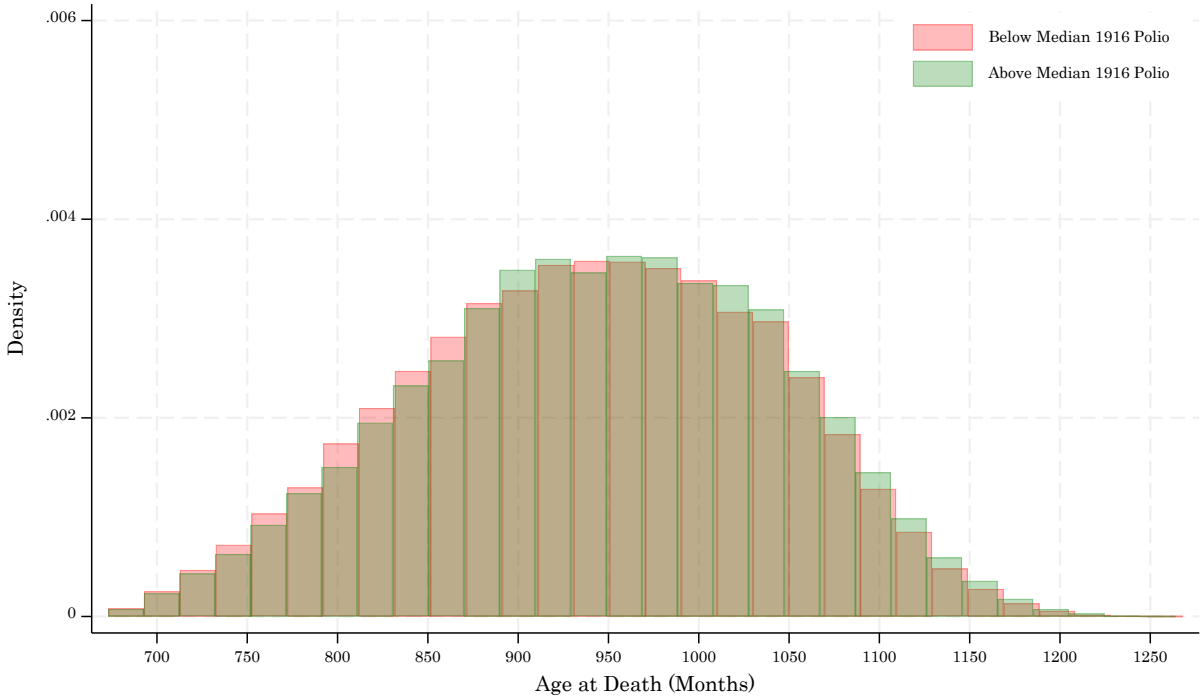

Figure 2 - Distribution of Age-at-Death across States with High and Low Polio Rate in 1916

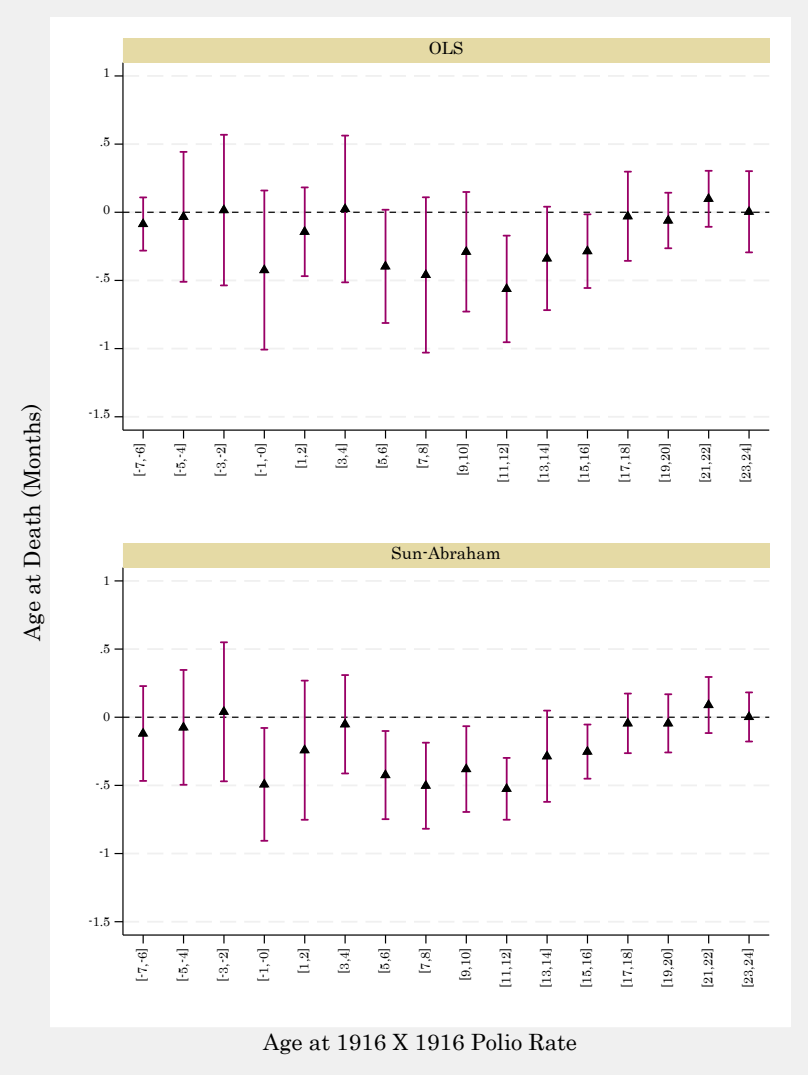

Figure 3 - Exposure to 1916 Polio Rate across Different Ages and Old-Age Longevity

Notes. Point estimates and standard errors are depicted. Standard errors are clustered on state-year. All regressions include birth-state fixed effects, birth-year fixed effects, individual controls, family controls, 1918-1919 flue rate controls, and state controls. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal education dummies. Regressions include state-level 1918-1919 influenza rate 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. 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. In this appendix, we compare the results of the DMF with those of Numident. These results are reported in Appendix Table A-1. In column 1, we replicate the main results of the paper using the DMF data. In column 2, we restrict the DMF sample to death years 1988-2005, a death window comparable to that of Numident. Although we observe negative coefficients for our exposure variables, they are considerably smaller in magnitude and statistically insignificant. This fact suggests that the truncation of the data with narrower death window may significantly bias the coefficients downward. In column 3, we focus on the subsample of male individuals in the Numident data. We observe negative and mostly significant coefficients. One notable difference is the negative coefficient of age at exposure of [1,4] which is considerably larger in magnitude compared with that of column 1 and statistically significant. In column 4, we restrict the movement and sample to female deaths. We observe quite small and insignificant coefficients which suggests that the observed long-run impacts are concentrated among male individuals only. In column 5, we pool male and female subsamples and report the results. We observe negative and significant coefficients only for ages at exposure of [-1,0] and [1,4].

Appendix Table A-1 – Comparing the Results with Numident Data

	Outcomes:				
	DMF, Males, Years 1975-2005	DMF, Males, Years 1988-2005	Numident, Males, Years 1988-2005	Numident, Females, Years 1988-2005	Numident, Males & Females, Years 1988-2005
	(1)	(2)	(3)	(4)	(5)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-.40614**	-.19577	-.23887***	-.07982	-.1581**
	(.18202)	(.14196)	(.07013)	(.09877)	(.0632)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.04235	-.15855	-.29975***	-.02899	-.15799**
	(.16234)	(.11423)	(.09732)	(.06423)	(.06471)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.36762***	-.07517	-.46546**	.11311	-.05891
	(.10376)	(.08191)	(.19599)	(.11013)	(.11621)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.38729***	-.13567	-.26525	.14247	.05504
	(.09802)	(.08782)	(.1802)	(.08833)	(.08009)
Observations	5802813	3216407	3295392	3699888	6995280
R-squared	.38101	.65764	.51615	.66831	.62942
Mean DV	916.484	955.069	920.923	965.091	944.284

Notes. Standard errors, clustered on birth-state-year, are in parentheses. All regressions include birth-state fixed effects, birth-year 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 education dummies. Regressions include state-level 1918-1919 influenza rate 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 B

There are 2 limitations with the DMF data. The death window of this data is truncated to the years 1975 – 2005. Further, it only includes male individuals. In order to understand the implications of these limitations, we use two alternative sources of data. First, we use the Berkeley Unified Numident Mortality Database (BUNMD) that covers deaths to both males and females. Although BUNMD covers much earlier death years, its coverage is considerably limited. Therefore, we focus on post-1970 death years as an arbitrary cutoff point. Further, the data is limited to pre-2007 deaths. Second, we use death certificates of the National Center for Health Statistics (NCHS). 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 1979-2019. We implement similar sample selection strategies and empirical method as in the main results of the paper. The results for the BUNMD and NCHS data are reported in Appendix Table B-1 and Appendix Table B-2, respectively. In both tables, we report the results on male individuals, female individuals, and the full sample in columns 1, 2, and 3, respectively. The BUNMD estimates suggest considerably larger coefficients for male individuals. The pattern in coefficients followed the same pattern as in the main results, i.e., larger effects for early life exposure and school-age children. We also find relatively larger coefficients among male individuals of the NCHS sample. The combination of these 2 tables implies that the truncation of the DMF data may underestimate the true effects, as suggested by other studies (Lleras-Muney et al., 2022). However, both samples fail to provide statistically significant coefficients for female individuals as well as the full sample.

Appendix Table B-1 - Replicating the Main Results Using BUNMD Data 1970-2007

	Outcome: Age-at-Death (Months)		
	Males	Females	Full-Sample
	(1)	(2)	(3)
Polio Rate (STD) × Age-at-Exposure [-1,0]	-2.11967***	-.50559	-1.1353**
	(.60974)	(.50995)	(.45401)
Polio Rate (STD) × Age-at-Exposure [1,4]	-1.70069***	-.0776	-.59953
	(.44713)	(.57176)	(.42086)
Polio Rate (STD) × Age-at-Exposure [5-10]	-4.15741***	1.06362	.26386
	(1.29405)	(.76192)	(.72219)
Polio Rate (STD) × Age-at-Exposure [11,16]	-6.17413***	.75037	.29109
	(1.64875)	(.71452)	(.7087)
Observations	8951994	9695849	18647843
R-squared	.33293	.48647	.44597
Mean DV	895.422	949.455	923.516
Birth-State FE	✓	✓	✓
Birth-Year FE	✓	✓	✓
Individual Controls	✓	✓	✓
Family Controls	✓	✓	✓
1918-19 Flu by Birth-Year FE	✓	✓	✓
State Controls	✓	✓	✓

Notes. Standard errors, clustered on birth-state-year, are in parentheses. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal education dummies. 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 B-2 - Replicating the Main Results Using NCHS Data 1979-2019

	Outcome: Age-at-Death (Months)		
	Males	Females	Full-Sample
	(1)	(2)	(3)
Polio Rate (STD) × Age-at-Exposure [-1,0]	.07771	-1.67514**	-.76451
	(.62578)	(.68711)	(.46653)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.92523	-1.38522**	-1.05506*
	(.57082)	(.59568)	(.54893)
Polio Rate (STD) × Age-at-Exposure [5-10]	-1.04123*	-.63165	-.70501
	(.62596)	(.46058)	(.48084)
Polio Rate (STD) × Age-at-Exposure [11,16]	-1.6318***	.08621	-.67414*
	(.56629)	(.38759)	(.40706)
Observations	21437720	25606359	47044079
R-squared	.13521	.13559	.17363
Mean DV	944.058	995.169	971.878
Birth-State FE	✓	✓	✓
Birth-Year FE	✓	✓	✓
Individual Controls	✓	✓	✓
Family Controls	✓	✓	✓
1918-19 Flu by Birth-Year FE	✓	✓	✓
State Controls	✓	✓	✓

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

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

Appendix C

The Tycho project provides disease cases information for a subset of cities (Tycho, 2021). This provides us with sub-state variation that we can exploit. In so doing, we link individuals from 1940 census to historical censuses 1900 – 1920 in order to infer their city of birth – childhood. We then link these individuals with the 1916 polio rate information at the city level calculated based on case reports of Tycho project. We are able to link about 1.4 million observations using this method. For the remainder of individuals in the final sample, we use exposure measure at the birth state level. We then employ regressions similar to equation 1 with both city and state fixed effects. We report these results in Appendix Table C-1. Although the point estimates across childhood ages are negative and mostly significant, the resulting pattern varies slightly from that of the main results. Specifically, we find that a one standard deviation change in polio rate results in 0.5, 0.7, and 0.3 months lower longevity for age groups 1-4, 5-10, and 11-16.

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

	Outcome: Age-at-Death (Months)			
	(1)	(2)	(3)	(4)
Polio Rate (STD) × Age-at-Exposure [-1,0]	.00175	-.03222	-.06269	-.06345
	(.15864)	(.15998)	(.16128)	(.16017)
Polio Rate (STD) × Age-at-Exposure [1,4]	-.38657***	-.4369***	-.45108***	-.45272***
	(.13483)	(.13641)	(.14167)	(.13862)
Polio Rate (STD) × Age-at-Exposure [5-10]	-.64269***	-.68888***	-.65215***	-.6519***
	(.14351)	(.15279)	(.15768)	(.15539)
Polio Rate (STD) × Age-at-Exposure [11,16]	-.33871**	-.38703**	-.28915*	-.28091*
	(.14888)	(.16143)	(.1509)	(.14931)
Observations	5690977	5690977	5641582	5641582
R-squared	.38453	.38488	.38483	.38484
Mean DV	915.713	915.713	915.711	915.711
Birth-City/State FE	✓	✓	✓	✓
Birth-Year FE	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓
Family Controls		✓	✓	✓
1918-19 Flu by Birth-Year FE			✓	✓
State Controls				✓

Notes. Standard errors, clustered on birth-city/state, are in parentheses. Individual controls include dummies for race and ethnicity. Family controls include dummies for paternal socioeconomic status and maternal education dummies. Birth-state controls include literacy rate, share of married people, and average socioeconomic index.
**** p<0.01, ** p<0.05, * p<0.1*