NBER WORKING PAPER SERIES

WALKING WITH FEET TIED TO THE PAST: CHILDHOOD EXPOSURE TO THE 1918 INFLUENZA PANDEMIC AND LATER-LIFE MALE MORTALITY

Hamid Noghanibehambari
Jason Fletcher

Working Paper 35471
http://www.nber.org/papers/w35471

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2026

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 the University of Wisconsin Madison and the Data-Intensive Research Conference. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Hamid Noghanibehambari and Jason Fletcher. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Walking with Feet Tied to the Past: Childhood Exposure to the 1918 Influenza Pandemic and Later-Life Male Mortality

Hamid Noghanibehambari and Jason Fletcher

NBER Working Paper No. 35471
July 2026
JEL No. I1, J10

ABSTRACT

Many previous studies have used cross-cohort comparisons and a focus on in utero exposure to document the lingering effects of the 1918 influenza pandemic on survivors' outcomes later in life. A subset has examined impacts on old-age health and mortality, finding mixed results. This paper contributes to the literature by extending the analysis to evaluate early-life and childhood exposure to the pandemic on old-age longevity along with a set of candidate mechanisms. We employ Social Security Administration death records linked to the full-count 1940 census and employ event-study and difference-in-difference regressions to compare longevity of cohorts with different age-at-exposure who were born in cities with high versus low exposure based on 1918 influenza-related mortality. Our treatment-on-treated calculations suggest that cohorts with age-at-exposure of 3-8 and 9-11 experience 6.4 and 11.2 months reductions in old age longevity; importantly, we note that these cohorts are often used as "controls" in the broader literature. While our main analysis focuses on males due to data constraints, supplementary results suggest smaller and less precise effects for females. An analysis of census data suggests reductions in health, education and socioeconomic scores as potential mechanisms.

Hamid Noghanibehambari
Austin Peay State University
noghanih@apsu.edu

Jason Fletcher
University of Wisconsin - Madison
and NBER
jfletcher@lafollette.wisc.edu

1. Introduction

The years 1918-1919 witnessed the largest pandemic in recorded history, the infamous Spanish flu. The estimates suggest that about one third of the world population contracted the disease and death tolls are estimated to be about 50 million, with approximately 650K deaths in the US (Gagnon et al., 2013; Patterson & Pyle, 1991). While the effects on the US economy was relatively modest[4] and local areas that were hit harder recovered faster and even experienced larger wage growth during the 1920s (Brainerd & Siegler, 2003), the effects on individuals lasted for decades.

A strand of research examines the long-term effects on individuals' later-life health and socioeconomic measures with a specific focus on early-life exposure to the pandemic. For instance, Almond (2006) documents reductions in education and income and increases in disability for in-utero exposed cohorts. Cook et al. (2019) show that these negative effects spill over to the next generations. Fletcher (2018b) shows that women exposed during the first trimester marry lower educated men. Children and grandchildren of cohorts exposed to the pandemic during early-life reveal reductions in education and income. A narrow strand of this literature evaluates the effects on old-age health and mortality outcomes. Almond & Mazumder (2005) find deteriorations in subjective and physical health of exposed cohorts compared with the adjacent cohorts. Although there are evidence of the effects on education and socioeconomic measures, the impacts on mortality and longevity is mixed and inconclusive.[5] Fletcher (2018b) employs data from the National Longitudinal Mortality Study and fails to find significant impacts on mortality for in-utero exposure. However, Fletcher (2018a) uses similar data and find small increases in mortality for those born in 1918 versus those born in 1917 and 1919. Mazumder et al. (2010) employ data from National Health Interview Surveys and document increases in cardiovascular diseases for 1918-born cohorts compared with adjacent cohorts. Myrskylä et al. (2013) document mixed evidence on mortality by cause of death. They show that in-utero exposed cohorts reveal reductions in cancer mortality but increases in non-cancer deaths.

[4] Barro et al. (2020) estimate a roughly 1.5 percent drop in GDP and about 2 percent reduction in consumption for the US economy during the 1918-1919 pandemic.

[5] Studies of other countries also find similar pattern of effects for a wide array of later-life education, socioeconomic status, and health outcomes (Beach, Clay, et al., 2022; Helgertz & Bengtsson, 2019; Karlsson et al., 2014; Lin & Liu, 2014; Nelson, 2010).

This paper employs newly released data and explores the effects of early-life and childhood exposure to the 1918 pandemic on later-life longevity. Our data allows us to infer individuals' city-of-childhood and implement event-study difference-in-difference regressions to examine the effects across different ages-at-exposure in high/low affected cities. We find significant reductions in longevity of those with age-at-exposure of 3-11 and smaller and insignificant effects for those exposed between 0-2. We find considerably larger effects among those with lower maternal education and from lower socioeconomic status families. Our analysis from census data reveals a very similar pattern of negative effects on education and measures of socioeconomic status.

Our paper adds to the existing literature in three ways. First, previous studies rely on cohort differences in mortality using adjacent cohorts as the control group (Cook et al., 2019; Fletcher, 2018a, 2018b, 2018c; Mazumder et al., 2010; Myrskylä et al., 2013). Therefore, they fail to account for the dosage of exposure across different geographic regions. We take advantage of a city-level measure of influenza mortality and employ a difference-in-difference method which enables us to absorb unobserved cohort heterogeneity and place heterogeneity in longevity.

Second, the focus of this specific literature is almost exclusively on in-utero exposures relying on the relevance of insults during prenatal development for life-cycle outcomes (Almond & Currie, 2011; Barker, 1994, 1995, 1997). However, our focus spans several years of childhood and argues that childhood experiences of influenza pandemic, and disease environment in general, is also an important contributor for long-run outcomes, including old-age mortality (Almond et al., 2018; M. D. Hayward & Gorman, 2004). In our data, cohorts exposed at ages 3-8 and 9-11 during 1918 suffered the largest longevity losses. This pattern matches the broader evidence that shocks in early childhood have far larger long-run impacts than shocks later in life (Aizer et al., 2016; Almond et al., 2018; Fletcher & Noghanibehambari, 2024, 2025; Goodman-Bacon, 2021; Goodman et al., 2011; Linde & Egede, 2024; Noghanibehambari & Fletcher, 2024; Smith, 2009a, 2009b; Smith et al., 2009).

Third, our paper provides empirical evidence to clarify the mechanisms through which the pandemic shaped long-run outcomes. The 1918 influenza was a multifaceted early-life shock. It operated through a direct health shock, which connects our findings to the literature that studies how childhood health and disease shape long-term outcomes (Smith, 2009a; Smith & Smith, 2010). It also operated by exposing children to indirect effects of the disease environment,

including limitations in nutrition, access to healthcare, and local economic activity. Each of these channels can contribute to long-term mortality. Therefore, the findings of the paper also speak to the literature that examines how these factors shape long-term health and survival (Case et al., 2002; Case & Paxson, 2009, 2010; Hayward et al., 2013; Hoehn-Velasco, 2018, 2021; Lindeboom et al., 2010; Van Abeelen et al., 2012; Van Den Berg et al., 2006). Additionally, as highlighted by both the predictions of our theoretical model and the empirical evidence, the pandemic may have affected family structure through parental death. Therefore, our findings also contribute to the literature on the long-term effects of parental loss on later-life mortality (Case & Ardington, 2006; Dupraz & Ferrara, 2025; Goldstein, 2025; Luecken & Roubinov, 2012; Smith et al., 2014; Stikkelbroek et al., 2012). Our theoretical framework integrates these channels by modeling both scarring and selection and allowing their magnitudes to vary with socioeconomic status. The empirical patterns we document, including age gradients, SES gradients, and race-specific evidence of selection, are consistent with this broader view of the pandemic as a combined health, environment, and family resource shock. Overall, this perspective positions our study within the broader literature on early-life adversity and early-life interventions, demonstrating that the pandemic's effects operated well beyond in-utero exposure and shaped long-run outcomes through multiple pathways.

Finally, as noted by Beach, Brown, et al. (2022), it is important to control for family characteristics as the 1918 pandemic affected cohorts come from lower socioeconomic status families.[6] Our data is linked to historical censuses and allows us to observe a rich set of family characteristics, which we use as additional controls in our regressions. This aspect of our analysis is also unique as other studies of long-run effects of 1918 pandemic on mortality fail to control for family background characteristics.

The results of this paper have important policy implications. They point to the long-term costs of pandemics for individuals' outcomes. The magnitude of the effects could be used in cost-benefit analyses to justify policy interventions during and following pandemics. The results also reveal subpopulations at risk which help policymakers in identifying those who could most benefit from pandemic-related policy initiatives. Finally, our analysis shows that the 1918 pandemic had

[6] They show that the results of the seminal paper by Almond (2006) become small and insignificant when they control for family background characteristics.

long lasting health and economic effects on cohorts who are often used as "controls" in the literature.

The rest of the paper is organized as follows. Section 2 provides historical background. Section 3 discusses the empirical method. Section 4 introduces data sources. Section 5 provides a simple theoretical framework. Section 6 reviews the results. Section 7 discusses mechanism channels. Section 8 discusses the economic meaning of the estimated effects. Finally, section 9 concludes the paper.

2. Background

2.1. The 1918 Influenza Pandemic

The 1918–19 "Spanish Flu" was the deadliest pandemic on record. It infected an estimated one-third of the world's population, killing on the order of 50 million people globally. In the United States roughly 650,000–700,000 people died (Bristow, 2012). The pandemic came in waves (spring and fall 1918, plus a smaller 1919 wave) and was unusually virulent: it caused severe pneumonia in many patients, overwhelming WWI medical facilities (Beach, Clay, et al., 2022; Jester et al., 2019). Wartime troop movements and crowded camps helped spread the virus worldwide. Major U.S. cities (e.g. Philadelphia, St. Louis) imposed quarantines and school closures during the worst wave (Hatchett et al., 2007; Markel et al., 2007). Economically, the shock was large but brief, and studies find that most hard-hit local economies recovered quickly, but at the individual level the shock could still have long-lived effects (Brainerd & Siegler, 2003). Surviving a severe childhood infection often left scars: for example, pediatric pneumonia is now known to permanently reduce adult lung function (Chan et al., 2015; Johnston et al., 1998). Likewise, children whose parents fell ill (or died) suffered lasting family disruption. Many households lost breadwinners or had disabled caregivers after 1918. Although systematic data on parental disability from the 1918 flu is scarce, historical accounts note that many survivors suffered chronic respiratory and neurological problems. In short, the 1918 pandemic imposed an immense shock on the early-life and childhood environment, a shock whose consequences can be traced in later-life health and mortality.

2.2. Pathways to Old-Age Mortality

A substantial literature documents biological mechanisms through which early-life exposure to severe infections can generate long-lasting health deficits. Viral respiratory illnesses

such as influenza and pneumonia can cause permanent lung damage in childhood, reducing pulmonary capacity and increasing the risk of chronic respiratory disease in adulthood (Collaro et al., 2023). Early infections may also induce lasting inflammatory or immune-system changes or alter metabolic development, consistent with the broader "fetal origins" and developmental programming hypotheses (Carraro et al., 2014; Johnston et al., 1998; Nelson, 2010; Ubags et al., 2020). Evidence from the 1918 influenza pandemic supports these biological pathways: individuals exposed in utero have been shown to experience higher rates of adult disability, chronic disease, and poor physiological markers, while studies from other countries find smaller body size and elevated risks of kidney disease and diabetes among cohorts gestated during the pandemic (Almond & Mazumder, 2005; Beach, Clay, et al., 2022; Guimbeau et al., 2020; Helgertz & Bengtsson, 2019; Lin & Liu, 2014; Mazumder et al., 2010; Myrskylä et al., 2013). Taken together, these findings indicate that severe early-life viral shocks can impair organ development and physiological resilience in ways that raise morbidity and mortality later in life.

In addition to biological scarring, early-life pandemics can affect long-run outcomes through disruptions to human capital formation and family environments. Childhood illness or community-level closures may reduce schooling, delay cognitive development, and lower educational attainment (Barro, 2022; Beach et al., 2016; Chapelle, 2022; Markel et al., 2007; Smith & Smith, 2010). The evidence from the 1918 flu suggests that exposed birth cohorts obtained less schooling and had lower socioeconomic attainment in adulthood, factors strongly associated with poorer later-life health (Almond, 2006; Almond & Mazumder, 2005; Beach, Brown, et al., 2022; Beach, Clay, et al., 2022; Fletcher, 2018a, 2018b). Family-level shocks also play a central role: pandemics often increase parental illness, disability, and mortality, leaving children with fewer resources, less parental investment, and greater exposure to stress and poverty. A broader set of indirect pathways, including constrained healthcare access, nutritional stress, and local economic disruption, can further compound these disadvantages as the literature suggest a direct link between these adversities and later life health (Hoehn-Velasco, 2021; Lindeboom et al., 2010; Van Den Berg et al., 2006, 2011). Studies of later influenza epidemics, such as the 1957 Asian flu, similarly find adverse impacts on birth weight and later cognitive performance (Kelly, 2011). Across pandemics, the literature emphasizes that both biological and social mechanisms contribute to the elevated chronic disease and mortality risks observed among individuals exposed during early life.

3. Econometric Method

Our primary analysis leverages an event-study-type difference-in-difference method that compares the longevity of individuals at various ages of exposure in cities with higher versus lower 1918-1919 influenza mortality rates. Specifically, we implement the following regressions:

$$y_{icrb} = \alpha + \sum_{i=-5}^{i=17} \xi_i I(1918 - b = i) \times Exp_{cr} + \gamma X_i + \eta Z_{crb} + \lambda_{rb} + \theta_c + \varepsilon_{icrb} \qquad (1)$$

The outcome y is age-at-death (in months) for individual i in city c in census-division r born in year b. The function $I(.)$ takes a value of one if the inside argument is true and zero otherwise. The parameter Exp indicates whether the city of birth/childhood is at the top three-quartiles of influenza mortality rate.[7] Therefore, the coefficients ζ represent the effects of cohorts who were exposed in high-mortality cities at various ages (0 to 17) and those born after the pandemic. In X, we include individual race, maternal education dummies, and paternal socioeconomic status dummies. In Z, we include city-year covariates extracted from full-count decennial censuses 1880-1940 and interpolated for inter-decennial years, including the share of married people, share of literate people, and average socioeconomic score. Birth-cohort by census-division-of-birth[8] fixed effects (λ) control for secular trends in longevity with heterogenous evolutionary paths across different regions in the country. City fixed effects (θ) absorb time-invariant unobserved heterogeneity of city influences in health and longevity. Finally, ε is a disturbance term. We cluster standard errors at the city and cohort level to account for both serial and spatial correlations in error terms.

[7] The use of a dummy variable to indicate high versus low exposure eases the interpretation of the results. However, one may be concerned with differential effects of exposure across different quartiles of influenza rate and also the extent of nonlinearities in the effects. In Appendix C, we extensively discuss this by replicating the main results across different age of exposure and different quartiles of influenza rate. We find that for several age-at-exposures, the effects of top-quartile cities become slightly smaller than those second and third quartiles, showing the joint roles of scarring and mortality selection. Therefore, it is not surprising to observe insignificant and smaller effects when we look at exposure using linear variables.

[8] A census division is one of nine Census-defined state groupings used for statistical reporting, including New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.

We complement the event studies with a series of difference-in-difference regressions in which we group several event-studies and focus on three age groups[9], as follows:

$$y_{icrb} = \alpha + \zeta_1 Age_b^{0-2} Exp_{cr} + \zeta_2 Age_b^{3-8} Exp_{cr} + \zeta_3 Age_b^{9-11} Exp_{cr} + \gamma X_i + \eta Z_{crb} + \lambda_{rb} + \theta_c + \varepsilon_{icrb} \quad (2)$$

Each Age variable is a dummy that equals one if the individual belongs to age-at-exposure range shown in superscripts. All other parameters are as in equation 1.

4. Data and Sample Selection

The primary source of data is Death Master Files (DMF) of Social Security Administration records linked to the full-count 1940 census extracted from the CenSoc Project (J. R. Goldstein et al., 2021). The data reports deaths to male[10] individuals that occurred between years 1975-2005.[11] We use cross-census linking rules provided by the Census Linking Project to locate individuals in censuses 1900-1930 (Abramitzky et al., 2020). The cross-census data allows us to observe city of birth-childhood of individuals.[12] This is an important identifier to measure the intensity of treatment and enables us to implement cross-city cross-cohort comparisons in a difference-in-difference setting.

We extract city-level influenza mortality data from Beach, Clay, et al. (2022). We use the mortality counts due to outbreaks of influenza for about 489 cities over the years 1918-1919. We calculate quartiles of influenza mortality rate for these cities as the dose-response measure in our

[9] The choice of age groups is guided by the flexible specification results from equation 1, shown in Figure 3. The age group 0–2 captures in utero and early-life exposure, allowing the estimated coefficients to align with and provide a benchmark for the broader literature on the 1918 influenza, which places substantial emphasis on these periods. The age group 3–8 reflects early childhood exposure and follows the pattern of point estimates through the trough of negative effects in Figure 3. Because the point estimates for ages 12–17 are generally similar in magnitude and mostly statistically insignificant, we designate ages 9–11 as the late-childhood exposure group.

[10] In Appendix A, we employ an alternative data that also covers females but includes a narrower death window. We argue that the effects are relatively smaller among females compared with males. Moreover, the effects are smaller for narrower death windows, suggesting that the inclusion of earlier deaths increase the magnitude of coefficients.

[11] The DMF provides the most complete and reliable mortality coverage during the years 1975 to 2005. As documented by the Social Security Administration, inclusion in the DMF in this period was driven by standardized administrative reporting, resulting in near universal coverage for individuals with Social Security numbers. After 2005, changes to state reporting rules substantially reduced the completeness of the file. Focusing on the 1975 to 2005 window therefore minimizes measurement error in mortality and offers the cleanest comparison across exposure groups. Nevertheless, in Appendix Table E-1 we replicate the main results using the extended DMF from the CenSoc project, which provides coverage through 2023. The estimated effects remain highly similar, and in some cases even larger, while preserving the overall pattern documented in the preferred specification.

[12] We use the city in the first census of observation as the city of birth/childhood. For instance, if an individual is born in 1901, the first census of observation is 1910. We assign the city-of-residence in 1910 as the city of birth/childhood.

regressions. Figure 1 shows the geographic distribution of influenza rate across counties (top panel) and location of cities within these counties in the final sample (bottom panel).

We limit the sample to pre-1926 cohorts to have several cohorts born post-pandemic. We also remove cohorts aged above 20 in 1918. Therefore, our sample covers birth cohorts of 1898-1926 who died between 1975-2005. Summary statistics of the final sample are reported in the top-panel of Table 1. The final sample covers 1,471,443 observations. Average age-at-death is 76.4 years (916.4 months). Figure 2 shows the density distribution of age-at-death for those above and below-median of the 1918 influenza rate. About 10, 18, and 7 percent of individuals live in high exposure cities and are aged 0-2, 3-8, and 9-11 years old, respectively. Although the sample underrepresents blacks and overrepresents whites (about 3 percent blacks and 96.7 percent whites), both subpopulations have similar characteristics as the original cohorts in the original 1940 census (Breen & Osborne, 2022). In section 6.2, we further delve into this issue and argue that these differential representations due to data linking process is not likely to confound our estimates.

We employ alternative data sources to examine mechanism channels. We use the information reported in the 1940 census to examine the effects on educational achievements and other socioeconomic measures. Further, we also use census data for the years 1980-2000 combined with the American Community Survey data for the year 2005 extracted from Ruggles et al. (2020). The reason for this selection of years is to have a sample consistent with death years of the DMF data (1975-2005). We also use information on World War II enlistment records that are linked to the 1940 census and our final sample. This data is extracted from Goldstein et al. (2023). The bottom panels of Table 1 report summary statistics of these data sources. We will further discuss these data and their relevance in section 7.

5. A Simple Theoretical Framework

We outline a simple conceptual framework to rationalize the heterogeneous long-run effects of 1918 influenza exposure documented in our empirical analysis. Our framework combines essential ideas from the dynamic human-capital models of Cunha & Heckman (2007) and the mortality-selection logic developed in the early-life health literature (e.g., Currie et al. (2014) and Cunha & Heckman (2009)).

Each child begins life with a vector of capabilities, i.e., health, cognitive skills, and non-cognitive traits. Denote this composite stock as H_t, where t indexes the child's age category. The child development is dynamic and follows that:

$$H_{t+1} = f(H_t, I_t, E_t)$$

Where I_t is parental investments (e.g., nutrition, time, healthcare, schooling) and E_t represent environmental conditions, including disease environment. This simple framework contains two crucial properties. First, it introduces self-productivity that barely health and skills persist and reinforce themselves over time. In other words, self-productivity posits that "capabilities beget capabilities". For instance, an influenza infection that reduces lung capacity or cardiac robustness at age 5 (H_5) results in a lower health stock at age 6 (H_6), age 7, and so on, even if the disease environment E_t returns to baseline. Second, the framework provides a dynamic complementarity aspect whereby early deficits lower the productivity of later investments. Therefore, the compensatory behavior of parental investment in later stages of development (I_t) produces lower marginal return to replace the health shock.

5.1. Pandemic Exposure and Mortality Selection

At age j a high-mortality event (pandemic) strikes, imposing additional health and survival risk. Let $S_{ij} = 1$ indicate survival of child i at age j, with $P(S_{ij} = 1)$ increasing in H_{ij} (children in better health survive with higher probability). Equivalently, one can define a survival threshold $\overline{H}_j$ so that $S_{ij} = 1$ if $H_{ij} > \overline{H}_j$. Because infants and very young children had much higher baseline all-cause mortality than older children, the survival threshold for newborns was high. This pattern is evident in historical U.S. vital statistics: in 1914, before the influenza pandemic, the all-cause death rate in the death-registration states was 107.2 per 1,000 for infants under age 1, compared with 10.2 per 1,000 at ages 1–4 and 2.5 per 1,000 at ages 5–14 (U.S. Bureau of the Census, 1949). This steep age gradient is consistent with the high level of early-childhood mortality in the late nineteenth and early twentieth centuries documented by Preston & Haines (2014). Thus, an infant shock culls a large fraction of low-health children, leaving only robust survivors (hence, strong selection). In contrast, for older children baseline death rates are low, so relatively few die from the shock, suggesting $\overline{H}_j$ is low, selection is weak, and scarring dominates. Formally, if Y_i is an eventual outcome (say adult health or education) increasing in H_i, then among survivors, the net effect of a shock experienced at age j can be decomposed into a selection term (survivors' average

health/ability is shifted upward by culling) and a scarring term (the shock reduces each survivor's skills):

$$\begin{aligned}\Delta_{ij} &= E\left[Y_i \middle| S_{ij} = 1\right]_{shock} - E[Y_i]_{no\ shock} \\ &= E\left[Y_i(H_i - \delta) \middle| H_i \geq \bar{H}_j\right] - E[Y_i(H_i)|H_i \geq 0] \\ &= \left\{E\left[Y_i(H_i) \middle| H_i \geq \bar{H}_j\right] - E[Y_i(H_i)|H_i \geq 0]\right\} + \{E\left[Y_i\left(H_i - \delta_j\right) \middle| H_i \geq \bar{H}_j\right] \\ &\quad - E\left[Y_i(H_i) \middle| H_i \geq \bar{H}_j\right]\}\end{aligned} \tag{3}$$

The first term captures how early-life exposure alters the survival threshold and thus changes the composition of the surviving cohort ("selection"). The second term captures the direct physiological and developmental reductions in capability among survivors caused by the shock ("scarring"). In this formulation, the parameter δ represents the shock at age j due to exposure to the pandemic.

5.2. Heterogeneity by Socioeconomic Status

A substantial literature in economics and public health underscores that socioeconomic status (SES) is a key determinant of children's vulnerability to early-life shocks (e.g., Currie 2009; Currie et al. 2014). Therefore, SES introduces an important layer of heterogeneity in the effects of early-life exposure. Because SES shapes baseline health, living conditions, and access to care, the survival threshold during the pandemic and the severity of post-exposure scarring differ systematically across households. For instance, high-SES families can offset part of the biological shock through compensatory investments while low-SES children face both weaker protection ex ante and fewer compensatory resources ex-post. In order to update the theoretical net effect term, we can allow SES to shift the effective survival threshold by augmenting health (H) or altering parental investment (I). Therefore, it follows that:

$$S_{ij} = 1\ if\ H_{ij} + \phi Q_i \geq \bar{H}_j$$

Where Q represents SES and $\phi > 0$ means higher SES relaxes the survival condition (e.g., better food, housing, medical care, isolation from contagion). Equivalently, the survival threshold now depends on SES, as follows:

$$\bar{H}_j(Q_i) = \bar{H}_j - \phi Q_i$$

Similarly, the scarring term becomes:

$$\delta(Q_i) = \delta - \alpha Q_i$$

Therefore, the net effect is:

$$\begin{aligned}\Delta_{ij} = &\left\{E\left[Y_i(H_i)\middle|H_i \geq \bar{H}_j(Q_i)\right] - E[Y_i(H_i)|H_i \geq 0]\right\} \\ &+ \left\{E\left[Y_i\left(H_i - \delta_j - \gamma PL_{ij}\right)\middle|H_i \geq \bar{H}_j(Q_i)\right] - E\left[Y_i(H_i)\middle|H_i \geq \bar{H}_j(Q_i)\right]\right\}\end{aligned} \tag{4}$$

The first term captures how SES and parental loss shift the survival threshold and therefore the composition of survivors ("selection"). The second term captures how SES and parental loss shape the magnitude of developmental damage experienced by survivors ("scarring"). The difference between these two components determines the net reduced-form effect observed in the data.

The model implies predictions about heterogeneity by SES. As SES (Q_i) increases, the survival threshold ($\bar{H}_j(Q_i)$) falls, so selection becomes weaker. At the same time, higher SES reduces the size of the scarring shock, since $\delta(Q_i) = \delta - \alpha Q_i$ is smaller, so the scarring term becomes less negative with SES.

5.3. Parental Loss

Parental loss affects survival probabilities through reductions in resource investments (I_{ij}). We can update the survival condition as:

$$S_{ij} = 1 \; if \; H_{ij} + \psi PL_{ij} \geq \bar{H}_j$$

Here, PL is an indicator of losing at least one parent during the pandemic. The parameter ψ captures that parental loss reduces the child's effective survival capacity. In other words, the effective survival threshold is:

$$\bar{H}_j^* = \bar{H}_j - \psi$$

Since $\psi < 0$, this implies that $\bar{H}_j^* > \bar{H}_j$. Accordingly, we can update the original scarring term as a result of the shock (i.e., δ) to $\delta_j + \gamma_j PL_{ij}$, in which $\gamma > 0$. Therefore, we can update the net effect formulation as:

$$\Delta_{ij} = \left\{E\left[Y_i(H_i)\middle|H_i \geq \bar{H}_j^*\right] - E[Y_i(H_i)|H_i \geq 0]\right\} + \left\{E\left[Y_i\left(H_i - \delta_j - \gamma PL_{ij}\right)\middle|H_i \geq \bar{H}_j^*\right] - E\left[Y_i(H_i)\middle|H_i \geq \bar{H}_j^*\right]\right\} \quad (5)$$

Where $\bar{H}_j^* = \bar{H}_j - \psi$ and $\psi < 0$. Given the increase in the survival threshold, we observe stronger positive selection and stronger scarring, which magnifies the downward shift in long-run mortality.

6. Results

6.1. Concerns over Endogeneity

The Spanish flu resulted in influenza mortality rates that varied substantially by individuals' age and sociodemographic backgrounds (Clay et al., 2019; Gagnon et al., 2013; Noymer & Garenne, 2000). The differential mortality may confound the estimates of equations 1 and 2 if the exposure results in excess mortality of people with specific characteristics that are related to their longevity. For instance, higher exposure may result in overrepresentation of people from higher socioeconomic status families due to their higher survival during the pandemic. Since these people have, on average, higher longevity for other usually unobserved reasons, our results may underestimate the true (negative) effects of exposure in longevity. We empirically test this using balancing tests in which we regress a series of pre-determined observable characteristics on exposure measures of equation 2, conditional on city and census-division-year fixed effects. The results are reported in Table 2.

The top panel of Table 2 reports estimates using information from the 1940 Census. We do not find statistically significant associations between exposure and being white or Black, nor do we find significant effects on father's schooling, mother's schooling, or father's socioeconomic score. The magnitudes are also economically small relative to the mean outcomes shown in the last row of the table. For example, exposure at ages 0–2 and 3–8 in high-mortality cities is associated with statistically insignificant decreases of only about 0.06 and 0.14 units, respectively, in father's socioeconomic score, compared with a mean of 32.7. The corresponding estimate for ages 9–11 is a small and statistically insignificant increase of 0.30 units. However, several coefficients on the indicators for missing paternal and maternal characteristics are statistically significant (columns 4, 6, and 8). It is important to note that the reference group in equation 2 consists of individuals born after the pandemic (cohorts 1919 to 1925) as well as those aged 18 to

19 at the time of the pandemic. Because parental characteristics are observed in the 1940 Census and the 1919 to 1925 cohorts are less likely to have left the parental household by 1940, it is not surprising that these later-born cohorts have fewer missing values, which mechanically produces positive coefficients on the missing indicators.

To further assess observable differences between the treatment and control groups, we link individuals across historical censuses using the matching procedure described in Abramitzky et al. (2020). This allows us to observe parental characteristics in the 1900 to 1930 censuses. The procedure mirrors the approach used to recover city-of-birth/childhood location described in Section 4. Results from these historical characteristics are reported in the bottom panel of Table 2. Individuals exposed in early childhood are less likely to have been born in the urban centers of cities (column 9). The evidence on parental literacy is mixed: for exposure at ages 0-2, paternal and maternal literacy rates are lower, whereas for exposure at ages 9-11, both parents appear more likely to be literate (columns 10 and 12). For father's socioeconomic index in the historical censuses, we do not find statistically significant associations across any age-at-exposure group (column 15). However, we do detect that fathers of individuals exposed at ages 3-8 and 9-11 in high-disease cities were more likely to hold white-collar occupations (column 14).

Several considerations reduce concern that the patterns in Table 2 threaten our identification strategy. First, most coefficients are economically small and estimated imprecisely, particularly for race, parental education, and paternal socioeconomic index. Second, the signs of the estimates do not point in a consistent direction; some are positive and some are negative, suggesting no coherent pattern of selection on observables. Third, in cases where statistically significant differences emerge, they typically imply that the exposed cohorts came from more advantaged families, as indicated by higher parental literacy and a greater likelihood of fathers working in white-collar occupations. Because children from more advantaged families tend to have higher longevity on average, these differences would, if anything, bias our estimated effects downward. Thus, the observable characteristics documented in Table 2, if they influence estimates at all, would lead us to understate the true long-run impacts.

6.2. Endogenous Merging

Individuals in the final sample come from death records that are linkable to the 1940 census as well as cross-census linking between 1940 and 1900-1930. This sample selection may cause

concerns about generalizability of the results. Further, the selection may also confound the estimates if those who are merged have some characteristics that are correlated with their exposure and longevity. In this section, we empirically explore whether there is any association between merging and our exposure measure. The idea behind this test is that if more/less exposed individuals are more/less likely to be in the final sample, then the estimates might pick up on characteristics of linked individuals. We start our analysis by using the full-count 1940 census and limit the sample to cohorts of 1897-1925 who reside in cities that are covered in our final sample. We then merge this with the final sample and generate a variable indicating successful merging. We then regress this successful merging indicator on the right-hand side variables of equation 2. The results are reported in Table 3. We find quite small and statistically insignificant associations. For instance, the those in high exposure cities with age-at-exposure of 0-2, 3-8, and 9-11 reveal 8.3, 2.1, and 11 basis-points change, off a mean of 0.088. Since these results do not support an association with our main independent variables, there is little concern about endogenous merging.[13]

6.3. Main Results

The main results of equation 1 are depicted in the top-panel of Figure 3. We remove the coefficient of 1919 cohorts (age-at-exposure of -1) so that all coefficients are in comparison with longevity of those born in 1919.[14] The effects on individuals born after the pandemic (age-at-exposure of -2 to -9) reveal no difference compared with those born in 1919. Moreover, the results do not reveal a significant impact for those with age-at-exposure of 0-1. This is in contrast with several studies that show excess mortality for in-utero exposure to the 1918 influenza (Fletcher, 2018a; Myrskylä et al., 2013) but in line with Fletcher (2018b) that reports null results. However, we find significant impacts for age-at-exposure of 2-11. For age-at-exposures of 9-11, we observe reductions in longevity compared with the reference point, but the differences are relatively small

[13] We should also note that both DMF-census and cross-census linking rules are based on name commonalities rather than specific individuals' characteristics that are correlated with their health. Therefore, it is unlikely that linking rules confound the exposure-longevity relationships.

[14] Based on the results of Figure 3, we observe no discernible and statistically significant difference between longevity of cohorts born in 1918, 1919, and several years afterward. One notable outlier is for cohorts of 1920 which depicts positive effect versus those of 1919 cohort. Moreover, we should note that the 1919 cohort is also treated as pandemic continued toward the early months of 1919 and many individuals were born in this year had been in utero during 1918 or the early months of 1919. Therefore, this cohort may also be considered in utero exposed. If there are indeed some influences for in utero exposure for this cohort, our results produce conservative estimates of the true effects as a portion of the control group has also received the treatment.

and insignificant. To account for potential heterogeneity in the treatment effects that are ignored by ordinary least square estimations, we implement the method developed by Sun & Abraham (2021) and replicate the main event-study results in the bottom-panel of Figure 3. We observe a very similar pattern as those produced by OLS.

One plausible explanation for the small effects at ages 0-1 is the very high mortality rate for this age group relative to other childhood ages. Thomas (1951) documents roughly 112 deaths per 1,000 infants under age 1 in 1918, compared with 15.7 deaths per 1,000 for ages 1-4 and 4.1 deaths per 1,000 for ages 5-14. This difference is even more evident when focusing on deaths from influenza and pneumonia. Rogers (1920) tabulates deaths by cause and age for 1918 in Philadelphia, Kansas, and Indiana and documents substantially higher influenza- and pneumonia-related mortality among ages 0-1 than among later childhood ages. For example, in Indiana, the influenza-related death rate per 100,000 for males aged 0 and 1 was 3,483 and 2,683, respectively, compared with only 409 for males aged 5-9. This extremely high mortality among infants may have generated strong positive selection among survivors. As noted by Bozzoli et al. (2009), when infant mortality rates are high, selection effects may dominate scarring effects, leaving the fittest infants to survive into adulthood. This logic is also consistent with the theoretical framework in Section 5: a severe shock during gestation or infancy raises the survival threshold, so survivors are disproportionately drawn from children with higher underlying health stocks. Therefore, it is not surprising that we do not observe longevity reductions among these individuals. By contrast, at lower mortality rates, especially among children aged 2-11, mortality selection is weaker and scarring effects are more likely to dominate, producing negative effects on old-age longevity. Finally, the 1918 influenza mortality rate displayed an unusual W-shaped age profile, with high mortality among infants, older adults, and young adults aged 15-44 (Gagnon et al., 2013; Noymer & Garenne, 2000), which is consistent with smaller estimated effects for age groups where mortality selection was stronger.

We group the event-study point estimates in the manner discussed in equation 2 and report the results in Table 4. Column 1 report the results of regressions with only city and birth-year fixed effects. We add more covariates across consecutive columns. We find relatively robust coefficients. Based on the full specification of column 4, individuals with age-at-exposure of 3-8 and 9-11 in high exposure cities reveal 2.8 and 1.6 months reductions in longevity, respectively. We find an insignificant effect of -0.79 for age-at-exposure of 0-2.

Several studies show that during the 1918 influenza pandemic, there were varying rates of infection and mortality based on socioeconomic indicators (Beach, Clay, et al., 2022; Mamelund et al., 2021). Therefore, one expects to observe similar heterogeneity for long-run outcomes. In Table 5, we explore the heterogeneity by race, maternal education, and paternal socioeconomic score. The heterogeneity results by race provide mixed but informative evidence. For age-at-exposure of 0–2, we observe a positive coefficient among Blacks. Because Black infants faced extremely high baseline mortality during the 1918 pandemic, this pattern is consistent with strong mortality selection: exposure at ages 0–2 likely removed the frailest infants and left a more robust pool of survivors. In contrast, for age-at-exposure of 3–8, the coefficient for Blacks is larger in magnitude than that for whites, although the estimate is imprecise. Turning to socioeconomic markers, we divide the sample by maternal education using information from the 1940 Census. Children of low-educated mothers (fewer than 6 years of schooling) exhibit substantially larger effects than children of higher-educated mothers (columns 3 and 4). For example, for age-at-exposure of 3–8, the coefficient is –5.1 for low-educated mothers compared with –1.9 for higher-educated mothers. A similar pattern emerges when we use father's socioeconomic index from historical censuses. Children from lower-SEI families display larger and more consistently significant estimated effects than children from higher-SEI families. However, the negative effects are not limited to the lower-SEI subgroup. For exposure at ages 3–8, the coefficient is also negative and statistically significant among children with high-SEI fathers, with an estimated reduction of 2.1 months. The corresponding estimate for children with low-SEI fathers is larger, at 3.1 months, and is also statistically significant. For the lower-SEI subgroup, all age-at-exposure categories yield negative and statistically significant coefficients.[15]

These patterns are consistent with the predictions of our theoretical framework (section 5.2), in which socioeconomic status shapes both the survival threshold and the severity of scarring, leading lower-SES children to experience larger long-run losses from pandemic exposure.

[15] Another potential source of heterogeneity comes from the historical fact that World War I veterans were more likely to contract the disease, carry the virus upon their return, and expose their family members. In Appendix D, we examine the source of heterogeneity and find mixed evidence with insignificant results. Despite the coefficients being larger for age at exposure of 0-2, the point estimate is statistically insignificant.

6.4. Robustness Checks

In Table 6, we explore the robustness of results to alternative specifications. In column 1, we add birth-state by birth-year fixed effects to compare cross-cohort outcomes across cities within a state. This model will absorb all time-varying policy and economic shocks at the birth-state level that could influence the penetration of pandemic and appear in longevity equations. The effects only slightly drop in magnitude compared with those of column 4 of Table 4.

In column 2, we allow for city fixed effects to have heterogenous impacts by race and parental characteristic dummies. We observe increases in the magnitudes. In column 3, we control for unobserved city characteristics that evolve linearly over cohorts by adding a city-specific trend in birth year. The coefficients become slightly smaller than the main model of Table 4.

To account for seasonality in birth and seasonal factors in mortality that may confound longevity equations, we include birth-month and death-month fixed effects. The results, reported in column 4, are very similar to the main results. In column 5, we add a battery of additional city controls, including population, labor force participation rate, average occupational income score, share of different race groups, age groups, and occupation groups. The coefficients remain fairly comparable to the main results.

In columns 6-8, we examine the sensitivity of the functional form. In column 6, we replace the outcome with log of age-at-death. The coefficients of age-at-exposure 3-8 and 9-11 suggest 0.3 and 0.16 percent change, almost identical to the implied percentage change with respect to the mean of the outcome in Table 4. In columns 7-8, we replace the outcome with binary variables indicating longevity beyond 75 and 80 years, respectively. We observe larger effects for longevity at older ages, suggesting that the negative effects are delayed until much older ages. This is true for both point estimates and their implied changes with respect to the outcomes' mean.

Since our death window is limited, we do not observe the full history of longevity of each cohort. Goldstein et al. (2021) use the Human Mortality Database to estimate gender-specific cohort-specific longevity. Based on these estimates, they provide a weighting variable for a subset of observations that take into account differential longevity of cohorts based on the observation window. In column 9, we weight regressions by their provided "weight" variable. We observe quite similar effects to those reported in Table 4. Finally, in column 10, we show that standard errors are quite robust when we cluster at the city level.

A potential concern in interpreting the results in light of the balancing tests in Table 2 is the significantly positive coefficients on the missing indicators for parental characteristics. The fact that these indicators are larger for younger ages (0–2) than for older ages is unexpected, since younger children were more likely to reside with parents in 1940 (although all age groups 0–11 are more than 22 years old in 1940, and therefore older than the usual age of living in the parental household). This pattern may reflect unobserved selection related to parental mortality, differential household formation, or migration following the pandemic. To evaluate this, Appendix Table F-1 replicates the main results separately for subsamples defined by whether parental characteristics are missing or non-missing in both the 1940 census and the historical censuses. The estimates are highly similar across these subsamples, indicating that the main findings are not driven by patterns of missingness. Appendix Table F-2 further incorporates flexible controls for parental background (using both historical and 1940 census information) and interacts these measures with city fixed effects to allow for heterogeneous baseline differences across cities. The results remain stable, providing no evidence that unobserved selection arising from parental loss or migration is biasing the estimated long-run effects.

Another argument of concern is that older individuals or individuals born after the pandemic may be systematically different from those who were in early childhood at the time of the pandemic. For example, children born after the pandemic might come from families with different sociodemographic characteristics, making them a less suitable control group (Beach, Clay, et al., 2022). In particular, if the pandemic generated lasting maternal health impairments or disabilities that reduced post-1918 childbearing, the cohorts born after the pandemic could be positively selected on maternal health, especially in cities that experienced more severe outbreaks (Almond, 2006; Almond & Mazumder, 2005). Such selection would make the post-1918 cohorts healthier on average, potentially exaggerating the relative declines in longevity observed for the exposed cohorts. Similarly, older cohorts aged 12 to 20 face particularly strong survivorship requirements to appear in the DMF data. For example, individuals aged 18 in 1918 were born in 1900 and would need to survive to at least age 75 to be observed in the DMF, which exceeds the average life expectancy of these cohorts. This generates a survivorship selection concern.

To address the concern regarding the unsuitability of control cohorts in the main results, we examine robustness across many alternative control cohorts. First, we gradually remove older cohorts to focus on comparisons of childhood ages at the time of the pandemic with those born

after the pandemic. We then remove those born after the pandemic and make the comparison with older cohorts only. These results are reported in Appendix Table H-1. Across all specifications, the estimated coefficients for ages 3–8 and 9–11 remain negative, stable in magnitude, and statistically significant, regardless of the control cohorts used. The consistency of these effects across a wide range of alternative cohort restrictions indicates that the main findings are not driven by the particular choice of comparison cohorts.

To further address this issue, we use information from the 1940 census to include mother fixed effects in the regressions. These results are reported in Appendix Table G-1. The coefficient for age at exposure 0–2 is somewhat larger (still negative), but the estimates for ages 3–8 and 9–11 are quite comparable to the main results. However, their lack of statistical significance limits further interpretation. With caution, we can argue that these results suggest the main findings are not driven by unobserved family-level factors.

A related concern is that the wide range of birth cohorts in the main sample may introduce asymmetric follow-up windows given the limited death window of the DMF (1975–2005). For instance, individuals born in 1898 must survive into their late 70s to appear in the DMF data, whereas individuals born in 1918 must survive only into their late 50s. This may generate selection for the earliest cohorts in ways that differ from later-born cohorts. The results of Appendix Table H-1 partly alleviate this concern, as we focus on varying birth cohort windows. To complement that analysis and examine this issue more directly, we replicate the analysis using more balanced samples in which all individuals are observed over a comparable age range at death. Appendix Table I-1 restricts the sample to cohorts born between 1905 and 1920 and limits the outcome to ages at death between 70 and 85, while Appendix Table I-2 further restricts the sample to cohorts born between 1910 and 1920 with ages at death between 65 and 85. Across these alternative windows, the estimated coefficients for the key childhood exposure groups remain negative and similar in magnitude to the main results, particularly for ages 3–8. These results therefore indicate that the main findings are not driven by asymmetric follow-up periods or selective survival among the earliest birth cohorts.

7. Mechanisms

There are several channels through which the exposure to influenza during childhood may affect later life health outcomes. One possible channel is the interruption of schooling due to

business closures, city lockdowns, and general Non-Pharmaceutical Interventions (NPI). Although studies show the potential benefits of NPIs during the pandemics on mortality rates, they are also associated with adverse effects on mental health outcomes and interruptions in schooling and other social interactions all of which may have long-lasting consequences (Garrett, 2007; Lai et al., 2020). For instance, city lockdowns are documented to interrupt children's socioemotional and cognitive developments. Therefore, these NPIs may lead to lower human capital accumulation and reductions in schooling which in turn affects later-life health (Adams-Prassl et al., 2022; Goodman et al., 2011; Spinelli et al., 2020). Another pathway is that the pandemic may result in parental death. Studies document that the experience of parental death during early life and childhood have long-lasting legacies, specifically for mental health outcomes and suicide (Rostila & Saarela, 2011; Smith et al., 2014). Further, childhood exposure to pandemic might be accompanied by significant psychosocial stress. Prolonged exposure to this stress during the developmental period may have long-lasting impacts on health outcomes (Brew et al., 2022). Furthermore, infection during the pandemic among children may trigger inflammatory responses that accumulates and contributes to chronic inflammation, which is documented to be associated with later life cardiovascular disease, diabetes, and other autoimmune diseases (Burgner et al., 2015). However, the changes in immune system programming can also be beneficial specifically if infected during childhood. Exposure to certain pathogens may contribute to changes in immune response system and make individuals less susceptible to future infectious diseases, bringing benefits for exposed children in the long-run (Jain, 2020).

Although the scarcity of data limits our ability to test many of these pathways, we can employ the information available in censuses and World War II enlistment data to provide suggestive channels.

7.1. Parental Death

A narrow strand of research suggests the long-run impacts of exposure to parental death during infancy and childhood for later life health outcomes, specifically mental health and suicide (Berg et al., 2014; Smith et al., 2014). We do not have any information on parental death in the census. However, we can observe the change in parental absence in census records of children in the same age range from 1910 to 1920. In so doing, we use the full count censuses 1910 and 1920 and limit the sample to male individuals aged 0-17. We then merge this data with the influenza

rate data to cover similar cities as in the final sample of the paper. We use "person number" of parents in the household to detect presence of parents. Thus, a missing value of this variable may represent, among alternative explanations, parental death. We then regress parental death dummy on covariates and fixed effects similar to the main regressions of the paper. We also add year fixed effects as the comparison is derived from cross city and over-time variation. The results are reported in columns 1-2 of Table 7. We observe positive coefficients across different ages at exposure. The coefficients suggest larger effects for mother absence (compared with the mean of outcomes). Since fathers (relative to mothers) are more likely to be absent for non-death reasons, the fact that we observe larger changes among mothers may point to possible influence of excess death observe in 1920 versus 1910. Moreover, these coefficients suggest relatively large impacts. For instance, we observe 27-71 basis points increases in the likelihood of maternal absence, off a mean of 0.063. This is equivalent to 4.2-11.1% increases with respect to the mean of maternal absence.

7.2. Education and SES Profile

Changes in education and socioeconomic measures may operate as mediatory channels of impact as studies document the associations between these outcomes and longevity (Chetty et al., 2016; Fletcher, 2015; Lleras-Muney, 2005; Miller & Bairoliya, 2021). Since the final sample is linked to the 1940 census, we can examine the effects of exposure on these outcomes, as observed in early adulthood. To avoid those with uncompleted education, we limit this sample to those born before 1922. We report these results in Table 8.

We find a positive but statistically insignificant coefficient for exposure at ages 0–2. This is in contrast with findings of Almond (2006) and Fletcher (2018b) but in line with the more recent study of Beach, Clay, et al. (2022) that find small and insignificant effects on educational outcomes after adjusting for proxies of parental characteristics.

In contrast, we observe negative and larger-magnitude coefficients for exposure at ages 3–8 and 9–11. Although the age 9–11 estimate is statistically insignificant, the age 3–8 estimate is significant at the 5 percent level. This pattern mirrors the reduced-form effects on longevity. In columns 2–5, we document declines across multiple measures of occupational status and socioeconomic index. The coefficients are particularly large and statistically significant for those exposed at ages 3–8 and 9–11, again consistent with the reduced-form mortality results. For

instance, relative to the mean outcome, exposure at ages 0–2, 3–8, and 9–11 in high-exposure cities corresponds to declines of 0.3 percent, 3.1 percent, and 2.8 percent, respectively, in the Duncan socioeconomic index.

7.3. World War II Enlistment Data

To complement the analysis of this section and further explore the mechanisms, we turn to World War II enlistment data. There are two advantages of using this data source. First, the data is linked to the 1940 full census and therefore we can link our final sample to the enlistment data at the individual level. Second, the data reports several important mediatory outcomes related to anthropometric measures and cognitive ability. Specifically, the data reports the Army General Classification Test (AGCT) scores. The AGCT was a comprehensive test developed by Army technicians in order to evaluate intelligence and other cognitive abilities of individuals in order to assign them to specific tasks and occupations within the Army. This measure is useful because intelligence and cognitive abilities are linked to early life exposures and also, they are documented to influence later life health and longevity (Deeg et al., 1990; Kremen et al., 2019; Lee et al., 2022). In addition, the data also reports height and weight of individuals. Earlier research points to the association between several anthropometric outcomes on schooling, socioemotional, and health outcomes (Deaton, 2007; Kim & Fletcher, 2021). Therefore, these measures provide us with an opportunity to explore several alternative pathways.

We use enlistment-based outcomes in regressions similar to the main results and report them in Table 9. In column 1, we explore the effects on the probability that an individual in the final sample is observed in the enlistment data. We find that higher exposure to the 1918 influenza results in a lower probability of being in the enlistment data with the largest effects for ages 0-2 and 3-8. Although the coefficients are relatively large with respect to the mean of the outcome, they are noisy and limit further interpretation.

In column 2, we explore the results on the standardized height-for-age variable. We observe significant reductions for age groups 3-8 and 9-11. The coefficients suggests reductions of about 6.5% and 5.1% of the standard deviation of height-for-age. In column 3, we examined the effects on AGCT score. We observe reductions in the score for all age groups, but the estimated effect is statistically significant only for those exposed at ages 3–8 and 9-11, indicating a decline of roughly 5.7 and 4.8 units relative to a mean of 81.1.

Next, we use information on height and weight to construct the Body Mass Index (BMI). We then generate a dummy variable that indicates the BMI being in the normal range (i.e., above 18.5 and below 24.5). The results are reported in column 4. All coefficients are negative but imprecisely estimated, limiting conclusions. Notably, however, the point estimate for the 3–8 age group is larger in magnitude than for the other groups. Overall, the pattern across all outcomes in this table broadly mirrors the results reported in Table 4.

8. Discussion

The results suggest reductions in longevity of children of ages 3-11 who were exposed to the 1918 influenza in high mortality cities. We should note that the documented estimates are intent-to-treat effects across the whole population of children. Although obtaining the number of contracted cases is difficult, CDC (2018) estimates that about 25 percent of the US population were inflicted with the disease. Assuming the same rate for all age groups, we can use this number as the first stage to inflate the effects of reduced-form of Table 4. A back-of-the-envelope calculation to convert intent-to-treat effects into treatment-on-treated effects suggests that those with age-at-exposure of 3-8 and 9-11 in high exposure cities experienced 11.2 and 6.4 months reductions in longevity, respectively. The original population of these cohorts in high exposure cities add up to about 2.498 and 1.06 million individuals. Based on the reduced-form treatment-on-treated estimates, we calculate roughly 2.89 million life-years lost to these cohorts due to childhood exposure to the pandemic.

One way to put the magnitudes into perspective is to employ Value of Statistical Life (VSL) estimates. The average age-at-death of exposed individuals (age-at-exposure of 0-11) in the final sample is roughly 77 years. The average life expectancy of males survived to 77 is estimated at roughly 10 years (Arias, 2014). Further, studies suggest an average VSL of about $10.9 million for US-born individuals (Colmer, 2020; Kniesner & Viscusi, 2019; Viscusi, 2018). A simple calculation yields $141,558 for each year of life. Using the life-years lost calculations, we estimate a total loss of $409.1 billion due to longevity reductions of cohorts exposed to the pandemic at ages 3-11.

We can also understand the magnitude of the effects by comparing these effects with other determinants of mortality. For instance, Halpern-Manners et al. (2020) documents that each additional year of schooling is associated with about 4 months additional age at death. Therefore,

the calculated treatment-on-treated effects are equivalent to a drop of about 1.6-2.8 years of schooling. Aizer et al. (2016) show that children of mothers who received a 3-year cash transfer under the Mother's Pension program during early decades of the 20th century live, on average, 11.6 months additional life. These cash transfers accounted for about 30-40 percent of pre-transfer maternal income. Therefore, the effects of childhood exposure to the pandemic are about 55-97 percent of a relatively large and long-lasting cash transfer program.

In addition to these discussions on longevity, we can also compare the magnitudes of our mechanism-channel estimates with the literature on alternative childhood disease exposures. Beach et al. (2016) study the long-term consequences of early-life exposure to typhoid fever and exploit plausibly exogenous variation in downstream water contamination. Their instrumental-variable strategy uses the presence and proximity of upstream cities discharging sewage into the same river system as an instrument for local typhoid mortality. Their IV estimates indicate that the eradication of typhoid fever would increase completed years of schooling by approximately 0.12 years, relative to a sample mean of 9.5 years, a change of roughly 1.3 percent. Using the same logic discussed above to convert our intent-to-treat estimates into changes in human-capital outcomes, our results imply that influenza exposure at ages 3–8 in high-mortality cities reduces completed schooling by roughly 0.18 years, relative to a mean of 10.3 years (Table 8). This represents an implied change of approximately 1.7 percent, which is slightly larger than the IV-based effect reported by Beach et al. (2016).

It is important to interpret these treatment-on-the-treated effects with caution. The scaling exercise assumes that the reduced-form estimates reflect the consequences of individual infection, but the estimated effects may also capture broader family-level shocks associated with the pandemic, such as long-term impacts of parental disability or parental death, as well as community-level pathways such as healthcare crowd-out effects. Because the empirical design cannot fully disentangle individual exposure from these correlated household- or community-level mechanisms, the treatment-on-the-treated values should be viewed as upper-bound approximations of the true effects rather than precise treatment-on-the-treated estimates.

9. Conclusion

The 1918 influenza pandemic, as the largest pandemic in recorded history, claimed hundreds of thousands of lives in the US. However, its massive death toll and welfare impacts

were not limited to only short-run outcomes. Its impacts were also detectable in a wide array of outcomes for the survivors decades later. Studies suggest that early-life exposure to the pandemic is associated with lower education, socioeconomic measures, and deteriorated health outcomes (Beach, Clay, et al., 2022). In this paper, we provided empirical evidence that childhood exposure to the 1918 influenza resulted in significant reductions in longevity.

We used social security administration death records linked with the full-count 1940 census. We implemented event-study and difference-in-difference method to compare longevity across different age-at-exposure of individuals living in cities with high influenza-related mortality versus low exposure cities. Our results suggest reductions in longevity of about 2.8 and 1.6 months for age-at-exposure of 3-8 and 9-11, respectively. We calculated treatment-on-treated effects of about 11.2 and 6.4 months, respectively. Further, we used information in the 1940 census and document reductions in education and socioeconomic scores as potential mechanism channels.

Our results can be used in cost-benefit calculations to justify government interventions during pandemics. Further, since the results reveal childhood ages as the primary age of impact, the findings may also help identifying the population at higher risks who would benefit most from interventions during major pandemics.

Reference

Abramitzky, R., Boustan, L., & Rashid, M. (2020). *Census Linking Project: Version 1.0 [dataset]*. https://doi.org/https://censuslinkingproject.org

Adams-Prassl, A., Boneva, T., Golin, M., & Rauh, C. (2022). The impact of the coronavirus lockdown on mental health: evidence from the United States. *Economic Policy*, *37*(109), 139–155. https://doi.org/10.1093/EPOLIC/EIAC002

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. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–172. https://doi.org/10.1257/JEP.25.3.153

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360–1446.

Almond, D., & Mazumder, B. (2005). The 1918 influenza pandemic and subsequent health outcomes: An analysis of SIPP data. *American Economic Review*, *95*(2), 258–262.

Arias, E. (2014). United States Life Tables, 2009. *Public Health Resources*. https://digitalcommons.unl.edu/publichealthresources/322

Bailey, M., Clay, K., Fishback, P., Haines, M., Kantor, S., Severnini, E., & Wentz, A. (2016). U.S. County-Level Natality and Mortality Data, 1915-2007. *Inter-University Consortium for Political and Social Research*. https://doi.org/https://doi.org/10.3886/E100229V4

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

Barro, R. J. (2022). Non-pharmaceutical interventions and mortality in U.S. cities during the great influenza pandemic, 1918–1919. *Research in Economics*, *76*(2), 93–106. https://doi.org/10.1016/J.RIE.2022.06.001

Barro, R. J., Ursúa, J. F., Weng, J., Ursúa Dodge, J. F., & Weng EverLife, J. (2020). *The Coronavirus and the Great Influenza Pandemic: Lessons from the "Spanish Flu" for the Coronavirus's Potential Effects on Mortality and Economic Activity*. https://doi.org/10.3386/W26866

Beach, B., Brown, R., Ferrie, J., Saavedra, M., & Thomas, D. (2022). Reevaluating the Long-Term Impact of In Utero Exposure to the 1918 Influenza Pandemic. *Journal of Political Economy*, *130*(7), 1963–1990. https://doi.org/10.1086/719757/SUPPL_FILE/2016445DATA.ZIP

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

Beach, B., Ferrie, J., Saavedra, M., & Troesken, W. (2016). Typhoid Fever, Water Quality, and Human Capital Formation. *The Journal of Economic History*, *76*(1), 41–75. https://doi.org/10.1017/S0022050716000413

Berg, L., Rostila, M., Saarela, J., & Hjern, A. (2014). Parental Death During Childhood and Subsequent School Performance. *Pediatrics*, *133*(4), 682–689. https://doi.org/10.1542/PEDS.2013-2771

Bozzoli, C., Deaton, A., & Quintana-Domeque, C. (2009). Adult height and childhood disease. *Demography*, *46*(4), 647–669. https://doi.org/10.1353/DEM.0.0079

Brainerd, E., & Siegler, M. V. (2003). The economic effects of the 1918 influenza epidemic. *Available at SSRN 394606*.

Breen, C. F., & Osborne, M. (2022). *An Assessment of CenSoc Match Quality*. https://doi.org/10.31235/OSF.IO/BJ5MD

Brew, B. K., Lundholm, C., Caffrey Osvald, E., Chambers, G., Öberg, S., Fang, F., & Almqvist, C. (2022). Early-Life Adversity Due to Bereavement and Inflammatory Diseases in the

Next Generation: A Population Study in Transgenerational Stress Exposure. *American Journal of Epidemiology*, *191*(1), 38–48. https://doi.org/10.1093/AJE/KWAB236

Bristow, N. (2012). *American pandemic: The lost worlds of the 1918 influenza epidemic*. Oxford University Press.

Burgner, D. P., Cooper, M. N., Moore, H. C., Stanley, F. J., Thompson, P. L., De Klerk, N. H., & Carter, K. W. (2015). Childhood Hospitalisation with Infection and Cardiovascular Disease in Early-Mid Adulthood: A Longitudinal Population-Based Study. *PLOS ONE*, *10*(5), e0125342. https://doi.org/10.1371/JOURNAL.PONE.0125342

Carraro, S., Scheltema, N., Bont, L., & Baraldi, E. (2014). Early-life origins of chronic respiratory diseases: understanding and promoting healthy ageing. *European Respiratory Journal*, *44*(6), 1682–1696. https://doi.org/10.1183/09031936.00084114

Case, A., & Ardington, C. (2006). The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography*, *43*(3), 401–420. https://doi.org/10.1353/DEM.2006.0022

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review*, *92*(5), 1308–1334. https://doi.org/10.1257/000282802762024520

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

Case, A., & Paxson, C. (2010). Causes and consequences of early-life health. *Demography 2010 47:1*, *47*(1), S65–S85. https://doi.org/10.1353/DEM.2010.0007

CDC. (2018). *History of 1918 Flu Pandemic*. https://www.cdc.gov/flu/pandemic-resources/1918-commemoration/1918-pandemic-history.htm

Census, U. S. B. of the. (1949). *Historical Statistics of the United States, 1789-1945: A Supplement to the Statistical Abstract of the United States, Prepared with the Cooperation of the Social Science Research Council* (Vol. 2). U.S. Government Printing Office.

Chan, J. Y. C., Stern, D. A., Guerra, S., Wright, A. L., Morgan, W. J., & Martinez, F. D. (2015). Pneumonia in Childhood and Impaired Lung Function in Adults: A Longitudinal Study. *Pediatrics*, *135*(4), 607–616. https://doi.org/10.1542/PEDS.2014-3060

Chapelle, G. (2022). The medium-term impact of non-pharmaceutical interventions. The case of the 1918 influenza in US cities. *Economic Policy*, *37*(109), 43–81. https://doi.org/10.1093/EPOLIC/EIAC003

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

Clay, K., Lewis, J., & Severnini, E. (2019). What explains cross-city variation in mortality during the 1918 influenza pandemic? Evidence from 438 U.S. cities. *Economics & Human Biology*, *35*, 42–50. https://doi.org/10.1016/J.EHB.2019.03.010

Collaro, A. J., McElrea, M. S., Marchant, J. M., Chatfield, M. D., Sondergeld, P., Perret, J. L., Vicendese, D., Anuntaseree, W., Dharmage, S. C., & Chang, A. B. (2023). The effect of early childhood respiratory infections and pneumonia on lifelong lung function: a systematic review. *The Lancet Child and Adolescent Health*, *7*(6), 429–440. https://doi.org/10.1016/S2352-4642(23)00030-5

Colmer, J. (2020). What is the meaning of (statistical) life? Benefit–cost analysis in the time of COVID-19. *Oxford Review of Economic Policy*, *36*(Supplement_1), S56–S63. https://doi.org/10.1093/OXREP/GRAA022

Cook, C. J., Fletcher, J. M., & Forgues, A. (2019). Multigenerational Effects of Early-Life Health Shocks. *Demography*, *56*(5), 1855–1874. https://doi.org/10.1007/S13524-019-00804-3

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

Cunha, F., & Heckman, J. J. (2009). Human capital formation in childhood and adolescence. *CESifo DICE Report*, *7*(4), 22–28.

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

Deeg, D. J. H., Hofman, A., & Van Zonneveld, R. J. (1990). THE ASSOCIATION BETWEEN CHANGE IN COGNITIVE FUNCTION AND LONGEVITY IN DUTCH ELDERLY. *American Journal of Epidemiology*, *132*(5), 973–982. https://doi.org/10.1093/OXFORDJOURNALS.AJE.A115740

Dupraz, Y., & Ferrara, A. (2025). Fatherless. *Journal of Human Resources*, *60*(4), 1126–1174. https://doi.org/10.3368/JHR.0122-12118R2

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). 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. (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). 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., & 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

Fletcher, J., & Noghanibehambari, H. (2025). Early-Life Lead Exposure and Male Longevity: Evidence from Historical Municipal Water Systems. *Explorations in Economic History*, 101701.

Gagnon, A., Miller, M. S., Hallman, S. A., Bourbeau, R., Herring, D. A., Earn, D. J. D., & Madrenas, J. (2013). Age-Specific Mortality During the 1918 Influenza Pandemic: Unravelling the Mystery of High Young Adult Mortality. *PLOS ONE*, *8*(8), e69586. https://doi.org/10.1371/JOURNAL.PONE.0069586

Garrett, T. A. (2007). Economic effects of the 1918 influenza pandemic. *Federal Reserve Bank*

of St. Louis, *26*, 74–94.

Goldstein, E. G. (2025). A Mighty Toll. *Journal of Human Resources*, 0922-12547R2. https://doi.org/10.3368/JHR.0922-12547R2

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

Goodman, A., Joyce, R., & Smith, J. P. (2011). The long shadow cast by childhood physical and mental problems on adult life. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6032–6037. https://doi.org/10.1073/PNAS.1016970108

Guimbeau, A., Menon, N., Musacchio, A., Alsan, M., Feigenbaum, J., Frydman, C., Hilt, E., Margo, R., & Severnini, E. (2020). *The Brazilian Bombshell? The Long-Term Impact of the 1918 Influenza Pandemic the South American Way*. https://doi.org/10.3386/W26929

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

Hatchett, R. J., Mecher, C. E., & Lipsitch, M. (2007). Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7582–7587. https://doi.org/10.1073/PNAS.0610941104/SUPPL_FILE/10941TABLE11.XLS

Hayward, A. D., Rickard, I. J., & Lummaa, V. (2013). Influence of early-life nutrition on mortality and reproductive success during a subsequent famine in a preindustrial population. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34), 13886–13891. https://doi.org/10.1073/PNAS.1301817110

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

Helgertz, J., & Bengtsson, T. (2019). The Long-Lasting Influenza: The Impact of Fetal Stress During the 1918 Influenza Pandemic on Socioeconomic Attainment and Health in Sweden, 1968–2012. *Demography*, *56*(4), 1389–1425. https://doi.org/10.1007/S13524-019-00799-X

Hoehn-Velasco, L. (2018). Explaining declines in US rural mortality, 1910–1933: The role of county health departments. *Explorations in Economic History*, *70*, 42–72. https://doi.org/10.1016/J.EEH.2018.08.003

Hoehn-Velasco, L. (2021). The Long-term Impact of Preventative Public Health Programs. *The Economic Journal*, *131*(634), 797–826. https://doi.org/10.1093/EJ/UEAA071

Jain, N. (2020). The early life education of the immune system: Moms, microbes and (missed) opportunities. *Gut Microbes*, *12*(1). https://doi.org/10.1080/19490976.2020.1824564

Jester, B., Uyeki, T. M., Jernigan, D. B., & Tumpey, T. M. (2019). Historical and clinical aspects of the 1918 H1N1 pandemic in the United States. *Virology*, *527*, 32–37.

Johnston, I. D. A., Strachan, D. P., & Anderson, H. R. (1998). Effect of Pneumonia and Whooping Cough in Childhood on Adult Lung Function. *New England Journal of Medicine*, *338*(9), 581–587. https://doi.org/10.1056/NEJM199802263380904;PAGEGROUP:STRING:PUBLICATION

Karlsson, M., Nilsson, T., & Pichler, S. (2014). The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of Health Economics*, *36*(1), 1–19. https://doi.org/10.1016/J.JHEALECO.2014.03.005

Kelly, E. (2011). The Scourge of Asian Flu. *Journal of Human Resources*, *46*(4), 669–694. https://doi.org/10.3368/JHR.46.4.669

Kim, J., & Fletcher, J. M. (2021). The effects of relative body weight on socioemotional and schooling outcomes among female adolescents in the United States. *Social Science & Medicine*, *289*, 114418. https://doi.org/10.1016/J.SOCSCIMED.2021.114418

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

Kremen, W. S., Beck, A., Elman, J. A., Gustavson, D. E., Reynolds, C. A., Tu, X. M., Sanderson-Cimino, M. E., Panizzon, M. S., Vuoksimaa, E., Toomey, R., Fennema-Notestine, C., Hagler, D. J., Fang, B., Dale, A. M., Lyons, M. J., & Franz, C. E. (2019). Influence of young adult cognitive ability and additional education on later-life cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(6), 2021–2026. https://doi.org/10.1073/PNAS.1811537116/SUPPL_FILE/PNAS.1811537116.SAPP.PDF

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., Wesolowski, A., Santillana, M., Zhang, C., Du, X., Yu, H., & Tatem, A. J. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature 2020 585:7825*, *585*(7825), 410–413. https://doi.org/10.1038/s41586-020-2293-x

Lee, H., Lee, M. W., Warren, J. R., & Ferrie, J. (2022). Childhood lead exposure is associated with lower cognitive functioning at older ages. *Science Advances*, *8*(45), 5164. https://doi.org/10.1126/SCIADV.ABN5164

Lin, M.-J., & Liu, E. M. (2014). Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *Journal of Health Economics*, *37*, 152–163.

Linde, S., & Egede, L. E. (2024). Individual-Level Exposure to Residential Redlining in 1940 and Mortality Risk. *JAMA Internal Medicine*. https://doi.org/10.1001/JAMAINTERNMED.2024.4998

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

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

Luecken, L. J., & Roubinov, D. S. (2012). Pathways to Lifespan Health Following Childhood Parental Death. *Social and Personality Psychology Compass*, *6*(3), 243–257.

https://doi.org/10.1111/J.1751-9004.2011.00422.X

Mamelund, S. E., Shelley-Egan, C., & Rogeberg, O. (2021). The association between socioeconomic status and pandemic influenza: Systematic review and meta-analysis. *PLOS ONE*, *16*(9), e0244346. https://doi.org/10.1371/JOURNAL.PONE.0244346

Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M., & Cetron, M. S. (2007). Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. *JAMA*, *298*(6), 644–654. https://doi.org/10.1001/JAMA.298.6.644

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

Miller, R., & Bairoliya, N. (2021). Health, Longevity, and Welfare Inequality of Older Americans. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/REST_A_01103

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

Nelson, R. E. (2010). Testing the Fetal Origins Hypothesis in a developing country: evidence from the 1918 Influenza Pandemic. *Health Economics*, *19*(10), 1181–1192. https://doi.org/10.1002/HEC.1544

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

Noymer, A., & Garenne, M. (2000). The 1918 Influenza Epidemic's Effects on Sex Differentials in Mortality in the United States. *Population and Development Review*, *26*(3), 565–581. https://doi.org/10.1111/J.1728-4457.2000.00565.X

Patterson, K. D., & Pyle, G. F. (1991). The geography and mortality of the 1918 influenza pandemic. *Bulletin of the History of Medicine*, *65*(1), 4–21.

Preston, S. H., & Haines, M. R. (2014). *Fatal years: Child mortality in late nineteenth-century America*. Princeton University Press.

Rogers, S. L. (1920). Special tables of mortality from influenza and pneumonia. *Indiana, Kansas, and Philadelphia, PA*.

Rostila, M., & Saarela, J. M. (2011). Time Does Not Heal All Wounds: Mortality Following the Death of a Parent. *Journal of Marriage and Family*, *73*(1), 236–249. https://doi.org/10.1111/J.1741-3737.2010.00801.X

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

Smith, J. P. (2009a). Reconstructing childhood health histories. *Demography 2009 46:2*, *46*(2), 387–403. https://doi.org/10.1353/DEM.0.0058

Smith, J. P. (2009b). The Impact of Childhood Health on Adult Labor Market Outcomes. *The Review of Economics and Statistics*, *91*(3), 478–489. https://doi.org/10.1162/REST.91.3.478

Smith, J. P., & Smith, G. C. (2010). Long-term economic costs of psychological problems during childhood. *Social Science & Medicine*, *71*(1), 110–115. https://doi.org/10.1016/J.SOCSCIMED.2010.02.046

Smith, K. R., Hanson, H. A., Norton, M. C., Hollingshaus, M. S., & Mineau, G. P. (2014). Survival of offspring who experience early parental death: Early life conditions and later-life mortality. *Social Science & Medicine*, *119*, 180–190. https://doi.org/10.1016/J.SOCSCIMED.2013.11.054

Smith, K. R., Mineau, G. P., Garibotti, G., & Kerber, R. (2009). Effects of childhood and middle-adulthood family conditions on later-life mortality: Evidence from the Utah Population Database, 1850–2002. *Social Science & Medicine*, *68*(9), 1649–1658. https://doi.org/10.1016/J.SOCSCIMED.2009.02.010

Spinelli, M., Lionetti, F., Pastore, M., & Fasolo, M. (2020). Parents' Stress and Children's Psychological Problems in Families Facing the COVID-19 Outbreak in Italy. *Frontiers in Psychology*, *11*, 556582. https://doi.org/10.3389/FPSYG.2020.01713/BIBTEX

Stikkelbroek, Y., Prinzie, P., de Graaf, R., ten Have, M., & Cuijpers, P. (2012). Parental death during childhood and psychopathology in adulthood. *Psychiatry Research*, *198*(3), 516–

520. https://doi.org/10.1016/J.PSYCHRES.2011.10.024

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

Thomas, B. (1951). Historical Statistics of the United States, 1789–1945. *Journal of the Royal Statistical Society: Series A (General)*, *114*(1), 110–110. https://doi.org/10.2307/2980918

Ubags, N. D. J., Alejandre Alcazar, M. A., Kallapur, S. G., Knapp, S., Lanone, S., Lloyd, C. M., Morty, R. E., Pattaroni, C., Reynaert, N. L., Rottier, R. J., Smits, H. H., de Steenhuijsen Piters, W. A. A., Strickland, D. H., & Collins, J. J. P. (2020). Early origins of lung disease: towards an interdisciplinary approach. *European Respiratory Review*, *29*(157), 1–14. https://doi.org/10.1183/16000617.0191-2020

Van Abeelen, A. F. M., Veenendaal, M. V. E., Painter, R. C., De Rooij, S. R., Dijkgraaf, M. G. W., Bossuyt, P. M. M., Elias, S. G., Grobbee, D. E., Uiterwaal, C. S. P. M., & Roseboom, T. J. (2012). Survival effects of prenatal famine exposure. *The American Journal of Clinical Nutrition*, *95*(1), 179–183. https://doi.org/10.3945/AJCN.111.022038

Van Den Berg, G. J., Doblhammer-Reiter, G., Christensen, K., den Berg, G. J., Doblhammer-Reiter, G., Christensen, K., van den Berg, G. J., Doblhammer-Reiter, G., Christensen, K., den Berg, G. J., Doblhammer-Reiter, G., & Christensen, K. (2011). Being born under adverse economic conditions leads to a higher cardiovascular mortality rate later in life: Evidence based on individuals born at different stages of the business cycle. *Demography*, *48*(2), 507–530. https://doi.org/10.1007/s13524-011-0021-8

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

Viscusi, W. K. (2018). Best Estimate Selection Bias in the Value of a Statistical Life. *Journal of Benefit-Cost Analysis*, *9*(2), 205–246. https://doi.org/10.1017/BCA.2017.21

Tables

Table 1 - Summary Statistics

	Mean	SD	Min	Max
DMF-1940-Census Data:				
Death Age (Months)	916.38	108.67	589	1286
Birth Year	1914.3	7.47	1898	1926
Death Year	1989.67	8.57	1975	2005
Age 0-2	.14	.35	0	1
Age 3-8	.25	.43	0	1
Age 9-11	.1	.3	0	1
1918 Influenza Quartile 2-4	.72	.45	0	1
Age 0-2 × Q 2-4	.1	.3	0	1
Age 3-8 × Q 2-4	.18	.38	0	1
Age 9-11 × Q 2-4	.07	.26	0	1
White	.97	.18	0	1
Black	.03	.17	0	1
Father SEI 1st Quartile	.11	.31	0	1
Father SEI 2nd Quartile	.12	.32	0	1
Father SEI 3rd Quartile	.1	.3	0	1
Father SEI 4th Quartile	.09	.29	0	1
Father SEI Missing	.58	.49	0	1
Mother Education < High School	.38	.48	0	1
Mother Education = High School	.09	.29	0	1
Mother Education > High School	.02	.13	0	1
Mother Education Missing	.52	.5	0	1
City Literacy Rate	.95	.03	.7	1.07
City Share of Married	.56	.04	.18	.67
City Socioeconomic Index	31.48	3.09	9.54	47.27
Observations		1,471,443		
Mechanism Data (1940 Census):				
Schooling	10.33	2.83	0	20
Occupational Income Score	26.83	9	3	80
S&T Occupational Score	3.71	.84	.56	4.5
Duncan Socioeconomic Index	33.56	21.43	3	96
Occupational Prestige Score	34.48	12.52	0	81.5
Occupational Earnings Score	56.59	22.34	.6	100
Occupational Education Score	16.1	20.05	1	93.9
Nam-Powers-Boyd Occupational Status	56.54	20.89	.8	100
White-Collar Occupation	.15	.35	0	1
Observations		1,013,221		
Mechanism Data (1980-2000 Census and 2005 ACS):				
Schooling	12.89	2.59	0	20
Occupational Income Score	26.59	9.54	3	80
S&T Occupational Score	3.81	.94	.84	4.59
Duncan Socioeconomic Index	42.65	22.83	4	96
Occupational Prestige Score	38.95	14.09	0	81.5
Occupational Earnings Score	48.04	66.14	3.5	999.9
Occupational Education Score	47.6	24.94	6.6	100
Nam-Powers-Boyd Occupational Status	49.17	25.95	.4	100
White-Collar Occupation	.19	.39	0	1
Observations		1,096,712		
Enlistment Data:				
Height-for-Age (STD)	0	1	-2.68	3.39
AGCT Score	81.11	43.37	1	150
Normal BMI	.72	.45	0	1
Observations		219,380		

Table 2 - Balancing Tests

	Outcomes in 1940 Census:							
	White	Black	Father SEI	Father SEI Missing	Father Schooling	Father Schooling Missing	Mother Schooling	Mother Schooling Missing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age 0-2 × Q 2-4	-.02173*	.02008	-.05725	.03774***	.02567	.03607***	.01948	.03514***
	(.01316)	(.01292)	(.21659)	(.01129)	(.03369)	(.01125)	(.0332)	(.01208)
Age 3-8 × Q 2-4	-.02051	.01906	-.13861	.01952***	.08326	.02104***	.09774	.01992***
	(.01306)	(.0127)	(.28085)	(.00703)	(.0562)	(.00662)	(.06481)	(.00714)
Age 9-11 × Q 2-4	-.01908	.01776	.30209	-.00198	.1497	.00225	.10635	-.00168
	(.01203)	(.01165)	(.42055)	(.00586)	(.09214)	(.00486)	(.08629)	(.00372)
Observations	1471443	1471443	540648	1471443	604840	1471443	713093	1471443
R-squared	.12348	.13124	.02826	.35881	.06888	.34531	.08769	.37341
Mean DV	0.967	0.031	32.689	0.579	6.998	0.589	6.919	0.515
	Outcomes in Historical Censuses:							
	Parents Reside in Urban Area	Father literate	Father literacy missing	Mother literate	Mother literacy missing	Father in White-collar occupation	Father socioeconomic index	Father socioeconomic index Missing
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Age 0-2 × Q 2-4	.0057	-.01606***	.0194***	-.01232**	.01821***	.00052	.19141	.01609***
	(.00398)	(.00568)	(.0042)	(.00495)	(.00401)	(.00282)	(.20141)	(.00359)
Age 3-8 × Q 2-4	-.00863*	.00268	.00551*	.00867**	.00436	.00884***	.18003	.00126
	(.00512)	(.00369)	(.00334)	(.00347)	(.0029)	(.00293)	(.19959)	(.00301)
Age 9-11 × Q 2-4	-.03818***	.00898**	-.00151	.01837***	-.00199	.00888*	-.15618	-.00894**
	(.00973)	(.00429)	(.00457)	(.0051)	(.00443)	(.00453)	(.27662)	(.0042)
Observations	1471443	1471443	1471443	1471443	1471443	1471443	1269521	1471443
R-squared	.17959	.02348	.01676	.03289	.01489	.0184	.0261	.02442
Mean DV	0.851	0.789	0.152	0.791	0.128	0.209	30.083	0.137

Notes. Standard errors, clustered on city and year, are in parentheses. Regressions include city and census-division-by-year fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

Table 3 - The Association between Exposure to 1918 Influenza and Successful DMF-Census Merging

	Outcome: Successful Merging between Final Sample and Original Cohort of 1940 Census	
	(1)	(2)
Age 0-2 × Q 2-4	.00062	.00083
	(.00091)	(.00065)
Age 3-8 × Q 2-4	-.00001	-.00021
	(.00069)	(.0005)
Age 9-11 × Q 2-4	-.00123	-.0011
	(.00109)	(.00085)
Observations	16224287	16224287
R-squared	.00735	.11185
Mean DV	0.088	0.088
City FE	✓	✓
Census-division-year FE	✓	✓
Controls		✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Table 4 - Childhood Exposure to 1918 Influenza and Later-Life Male Mortality

	Outcome: Age at Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q 2-4	-1.00431	-.9744	-.8551	-.79942
	(.93051)	(.90367)	(.89066)	(.88398)
Age 3-8 × Q 2-4	-2.47005**	-2.95615***	-2.91689***	-2.83756***
	(1.01321)	(.75103)	(.78928)	(.78275)
Age 9-11 × Q 2-4	-.77486	-1.58069**	-1.63375**	-1.5966**
	(.58248)	(.59994)	(.64311)	(.64132)
Observations	1471443	1471443	1471443	1471443
R-squared	.2312	.23138	.23188	.23189
Mean DV	916.384	916.384	916.384	916.384
City FE	✓	✓	✓	✓
Year FE	✓			
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Table 5 - Heterogeneity Analysis of Childhood Exposure to 1918 Influenza and Later-Life Male Mortality

	Outcome: Age at Death (Months), Subsamples:					
	Whites	Blacks	High Educated Mother	Low Educated Mother	Father SEI High	Father SEI Low
	(1)	(2)	(3)	(4)	(5)	(6)
Age 0-2 × Q 2-4	-.96867	5.93843*	-.54749	-1.09899	1.56018	-1.636*
	(.89536)	(3.27109)	(.90943)	(1.6033)	(1.71318)	(.82926)
Age 3-8 × Q 2-4	-2.87093***	-3.87978	-1.93508	-5.13551***	-2.09128**	-3.14731***
	(.79889)	(3.76833)	(1.54446)	(.70243)	(1.00617)	(.78491)
Age 9-11 × Q 2-4	-1.72092**	.33913	-1.12387	-2.32257	-.84422	-1.86027**
	(.64862)	(6.92235)	(1.5218)	(2.03747)	(.93926)	(.71358)
Observations	1422937	45318	512267	200826	274271	995250
R-squared	.2292	.30205	.16418	.16497	.22705	.23287
Mean DV	916.929	899.633	884.315	890.030	916.121	914.232
City FE	✓	✓	✓	✓	✓	✓
Census-division-year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
Columns 3 and 4 use maternal education reported in 1940 to construct low- and high-education indicators. Mothers with fewer than six years of schooling (elementary school or less) are classified as low education, and those with seven or more years are classified as high education. Columns 5 and 6 use the father's socioeconomic index from historical censuses to define low- and high-SES groups, where low SES is an indicator for fathers in the bottom three quartiles of the SES distribution and high SES identifies fathers in the top quartile.
*** p<0.01, ** p<0.05, * p<0.1

Table 6 - Robustness Checks

	Outcome: Age at Death (Months)				
	Birth-State by Birth-Year FE	Interacting City FE by Covariate Dummies	Adding City Trend	Adding Death-Month and Birth-Month FE	Adding More City Controls
	(1)	(2)	(3)	(4)	(5)
Age 0-2 × Q 2-4	-1.31778	-.86808	-.81914	-.75925	-.86073
	(1.04921)	(.88697)	(.88524)	(.88192)	(.88975)
Age 3-8 × Q 2-4	-2.37167***	-3.00268***	-2.63117***	-2.80982***	-2.49467***
	(.79898)	(.73827)	(.74158)	(.78442)	(.62966)
Age 9-11 × Q 2-4	-1.06563	-1.76083***	-1.19204**	-1.56966**	-1.01846*
	(.74478)	(.60055)	(.52178)	(.63581)	(.56805)
Observations	1471443	1471409	1471443	1471443	1471443
R-squared	.23233	.23327	.23215	.23253	.23193
Mean DV	916.384	916.384	916.384	916.384	916.384
	Outcome: Log Age at Death	Outcome: Age at Death > 75	Outcome: Age at Death > 80	Using CenSoc Weights	Clustering on City
	(6)	(7)	(8)	(9)	(10)
Age 0-2 × Q 2-4	-.00084	-.00172	-.00569*	-.51562	-.79942
	(.00099)	(.00344)	(.00327)	(.3956)	(.66594)
Age 3-8 × Q 2-4	-.00303***	-.0122***	-.01191**	-2.63347***	-2.83756***
	(.0008)	(.00399)	(.00497)	(.88628)	(.80865)
Age 9-11 × Q 2-4	-.00161**	-.00315	-.00843	-1.4837*	-1.5966**
	(.00064)	(.0035)	(.00532)	(.82116)	(.76069)
Observations	1471443	1471443	1471443	1295978	1471443
R-squared	.23151	.1308	.12514	.13296	.23189
Mean DV	6.813	0.542	0.322	937.868	916.384

Notes. Standard errors, clustered on city and year, are in parentheses. Regressions include birth-city and birth-year fixed effects. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate. City controls in column 5 include population, share of immigrants, average occupational income score, labor force participation rates, share of people in difference race groups, age groups, and occupation groups. *** p<0.01, ** p<0.05, * p<0.1

Table 7 - Exploring the Effects on Parental Death

	Data: Census 1910-1920 ***Outcomes:***	
	Father Absent	Mother Absent
	(1)	(2)
Age 0-2 × Q 2-4	.00435***	.00269**
	(.00146)	(.00118)
Age 3-8 × Q 2-4	.00602**	.00574***
	(.00284)	(.00158)
Age 9-11 × Q 2-4	.00426	.00705**
	(.00411)	(.00316)
Observations	20428847	20428847
R-squared	.0511	.02861
Mean DV	0.111	0.063

Notes regarding the left panel: Standard errors, clustered on city and year, are in parentheses. Regressions include city fixed effects, census-division-by-year fixed effects, and controls. Controls include race dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Table 8 - Mechanism Analysis: Childhood Exposure to 1918 Influenza and Later-Life Education and Socioeconomic Measures

	Outcomes of the Final Sample in 1940:					
	Schooling	Duncan Socioeconomic Index	Nam-Powers-Boyd Occupational Status	Occupational Earnings Score	Occupational Prestige Score	White-Collar Occupation
	(1)	(2)	(3)	(4)	(5)	(6)
Age 0-2 × Q 2-4	.0183	-.1155	-.14475	.0825	-.11268	.0019
	(.02398)	(.2031)	(.16918)	(.16561)	(.13212)	(.00376)
Age 3-8 × Q 2-4	-.04531**	-1.04395***	-.38071***	-.41854***	-.62644***	-.01336***
	(.02191)	(.16271)	(.13464)	(.1316)	(.10099)	(.00286)
Age 9-11 × Q 2-4	-.04604	-.95716***	-.06246	-.37066*	-.62651***	-.01474***
	(.03739)	(.23396)	(.18795)	(.19814)	(.15215)	(.0044)
Observations	1070132	1085390	1082244	1082244	1085390	1085390
R-squared	.08302	.08476	.09962	.09054	.09083	.03791
Mean DV	10.329	33.558	56.541	56.591	34.483	0.147

Notes. Standard errors, clustered on city and year, are in parentheses. Regressions include city fixed effects, census-division-by-year fixed effects, and controls. Controls include race dummies and county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Table 9 - Exploring the Effects on Mediatory Outcomes Using WWII Enlistment Data

	Outcomes:			
	Being in the Enlistment Data	Height-for-Age (Standardized)	AGCT Score	Normal BMI
	(1)	(2)	(3)	(4)
Age 0-2 × Q 2-4	-.11849	-.05585	-4.07514	-.00363
	(.23637)	(.03737)	(2.98106)	(.00639)
Age 3-8 × Q 2-4	-.11946	-.06537**	-5.73336**	-.01015
	(.11039)	(.03024)	(2.49627)	(.00758)
Age 9-11 × Q 2-4	-.09042	-.05142*	-4.75201*	-.00126
	(.12165)	(.02882)	(2.7148)	(.00916)
Observations	1471443	219379	37129	256516
R-squared	.14957	.94946	.29283	.06591
Mean DV	0.146	-0.000	81.126	0.618
City FE	✓	✓	✓	✓
Census-division-year FE	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓
City Controls	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Figures

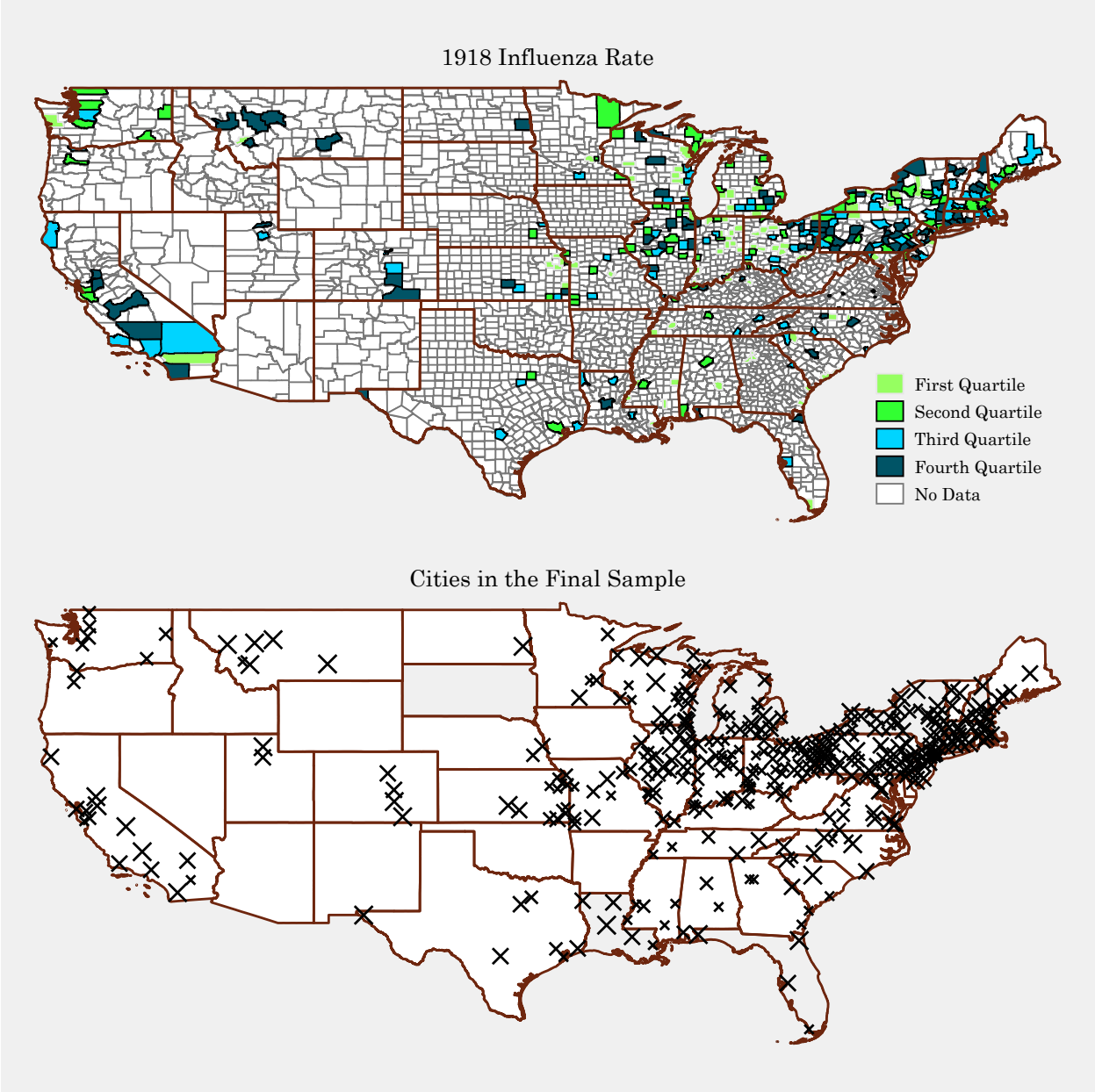


Figure 1 - Geographic Distribution of Cities in the Final Sample

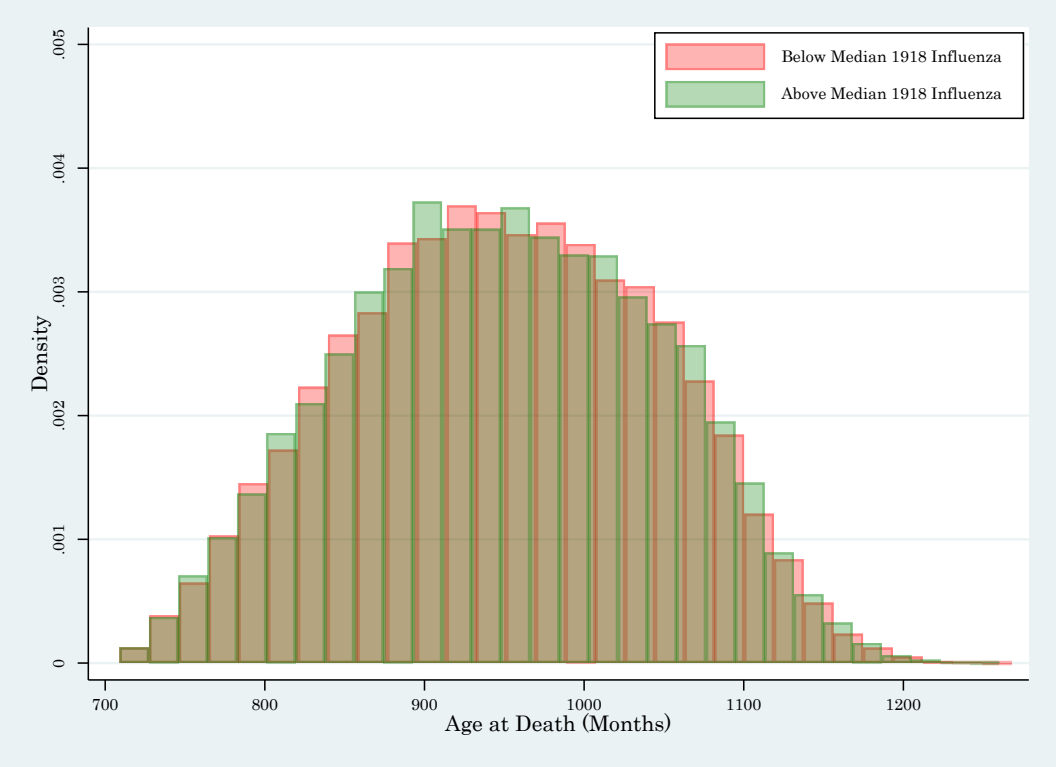


Figure 2 - Density Distribution of Age at Death based on Cumulative Childhood Exposure to the 1918 Influenza Pandemic

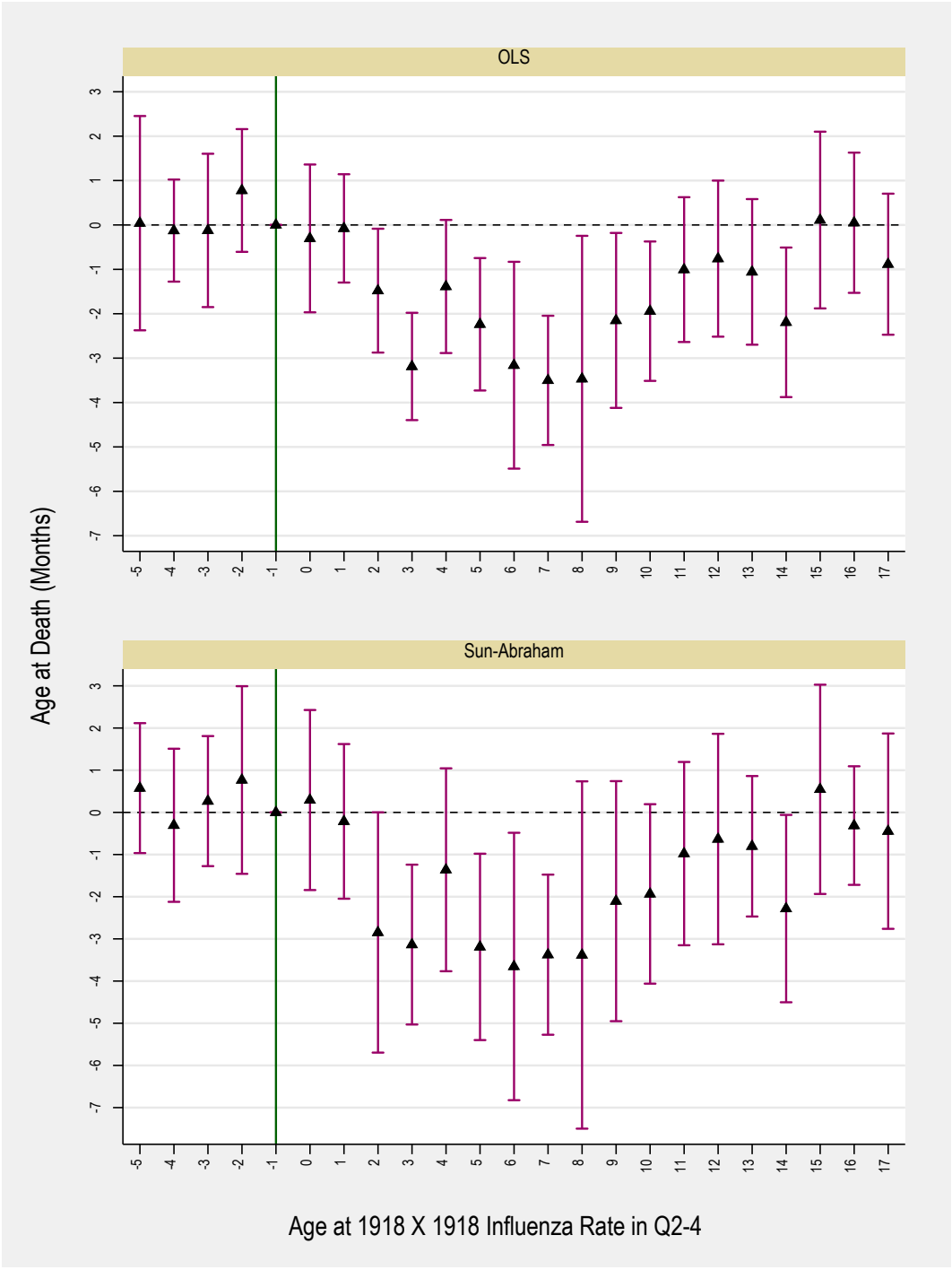


Figure 3 – Event Study of Childhood Exposure to 1918 Influenza and Later-Life Longevity

Notes. Point estimates and 90 percent confidence intervals are depicted. Standard errors are clustered on city and year. Regressions include birth-city and birth-division by birth-year fixed effects. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
Q2-4 represents top three quartiles.

Appendix A

The DMF data used in the paper covers deaths only to male individuals. To examine the effects on females, we use an alternative data source. We employ Numerical Identification (Numident) data extracted from Goldstein et al. (2021). The Numident data covers deaths that occur to both genders over the years 1988-2005. Similar to DMF, Numident is linked to the full-count 1940 census. We employ similar sample selections as in the DMF and implement equation 2 to compare the effects with those of DMF. The results are reported in Appendix Table A-1. In column 1, we replicate column 4 of Table 4. In column 2, we restrict DMF data to the Numident death years, i.e., years 1988-2005. We observe reductions in magnitude of the effects suggesting that the effects are larger as we expand death window and cover earlier death years. In column 3, we replicate the regressions using Numident data (death years 1988-2005) and restrict the sample to males to be able to compare the results with those of column 2. We observe almost identical effects for age 0-2 but substantially larger coefficient for age 3-8. In column 4, we restrict the Numident data to females. The effects drop by 25%-40% compared with those of column 3 suggesting heterogeneity in the effects by gender. Given that the point estimates for ages 3–8 and 9–11 in the full DMF data (column 1) are four to five times larger than those in the truncated DMF data (column 2), one might infer that if the death window for the Numident data for females were extended, the estimates of –0.9 and –0.7 for these age groups would also be several times larger, although still smaller in magnitude than the corresponding effects for males.

Appendix Table A-1 - Comparing the Results with the Numident Data Analysis for Death Years 1988-2005

	Outcome: Age at Death (Months)				
	DMF (1975-2005)	DMF (1988-2005)	Numident (1988-2005), Males Only	Numident (1988-2005), Females Only	Numident (1988-2005), Both Gender
	(1)	(2)	(3)	(4)	(5)
Age 0-2 × Q 2-4	-.79644	-.71474	-.60019	-.44448	-.54576*
	(.88496)	(.43828)	(.46458)	(.29549)	(.3107)
Age 3-8 × Q 2-4	-2.83927***	-.68882	-1.17606*	-.87016	-.95181*
	(.78352)	(.49289)	(.68976)	(.52692)	(.52756)
Age 9-11 × Q 2-4	-1.59962**	-.33373	-1.11026	-.66805	-.62275
	(.6441)	(.45454)	(1.76587)	(.75843)	(.80265)
Observations	1471443	839974	831405	1002701	1834106
R-squared	.23188	.53325	.39119	.53747	.50046
Mean DV	916.384	968.186	941.246	975.260	959.841
City FE	✓	✓	✓	✓	✓
Census-division-year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Regressions include birth-city and birth-year fixed effects. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix B

In this appendix, we employ a state-level analysis to examine the effects on infant mortality rates and birth rates. We use data from Bailey et al. (2016) that reports infant deaths and births by race for a subset of states for the years 1915-onwards. We focus on several years before and after the pandemic for a balanced panel of states. We implement regressions of the following forms:

$$y_{st} = \alpha + \beta_1 I(t = 1918) \times Exp_s + \beta_2 I(t = 1919) \times Exp_s + \gamma X_{st} + \lambda_s + \eta_t + \varepsilon_{st} \qquad \text{(B-1)}$$

Where y is log of infant mortality rate and birth rate for state s and year t. $I(.)$ is an indicator function. Exp is a dummy indicating the state is at the above-median of 1918 influenza mortality and zero otherwise. In X, we include state-level share of married, average socioeconomic score, and literacy rate. State and year fixed effects are included in λ and η, respectively. The sample covers the years 1916-1925. Standard errors are clustered on state.

The results are shown in Appendix Table B-1. Compared to adjacent years, we observe about 8.4 percent rise in infant mortality rate, 7.9 percent increase in white infant mortality rate, and 21 percent rise in black infant mortality rate (columns 1-3). However, we do not observe significant changes in birth rates and birth rate by race groups.

Appendix Table B-1 - State-Level Analysis of Exposure to the 1918 Influenza on Infant Mortality Rate and Birth Rate

	Outcomes:					
	Log Infant Mortality Rate	Log Infant Mortality Rate, Whites	Log Infant Mortality Rate, Blacks	Log Birth Rate	Log Birth Rate, Whites	Log Birth Rate, Blacks
	(1)	(2)	(3)	(4)	(5)	(6)
Year 1918 × Q 2-4	.08477**	.07953**	.21003**	.01198	.02211	-.02487
	(.03087)	(.03274)	(.07847)	(.01041)	(.01472)	(.06919)
Year 1919 × Q 2-4	-.00582	-.0034	.04165	-.01077	-.00771	-.03992
	(.01904)	(.01829)	(.10629)	(.02219)	(.02313)	(.0488)
Observations	235	235	235	235	235	235
R-squared	.95681	.95779	.68481	.96246	.95688	.99341
Mean DV	4.356	4.308	4.843	3.375	3.299	-0.263
State FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include state-level share of married, average socioeconomic score, and literacy rate. The sample covers the years 1916-1925.
*** p<0.01, ** p<0.05, * p<0.1

Appendix C

In the main results, we focused on cities at the top three quartiles of influenza rate during 1918-1919. We chose this threshold for ease of interpretation and comparison of highly affected areas with lower affected areas. In Appendix Figure C-1, We disaggregate the exposure measure across different quartiles of influenza rate and across different ages from 0 to 17. This figure reports the results of one regression that includes all covariates and fixed effects as the main results. We should note that in this figure we restrict the sample to cohorts born between 1900 and 1918. As a result, the comparison is no longer based on older cohorts or cohorts born after the pandemic; instead, identification comes solely from comparing individuals in the top quartiles of influenza mortality with those in the bottom quartile.

For most ages, the effects across different quartiles of influenza rate are quite close and comparable. However, we also observe slight discrepancies. For instance, for age at exposure of 9, 11, 12, and 13, we observe virtually zero effect for the top quartile cities while for quartiles two and three we observe negative effects. One possible explanation is the influence of mortality and selection in high exposed cities while the scarring effects for medium exposure cities. Nonetheless, the overall picture is closely aligned with the results of Figure 3.

To further explore the influence of linearity in the effects, in Appendix Table C-1, we replicate the main results for specifications that replace the top three quartiles dummy with log of influenza rate (column 1) and standardized influenza rate (column 2). Although we observe a very similar pattern as the main results, we do not find statistically significant coefficients. The difference between bottom quartile and top-three quartile cities and influenza rate is roughly 12.5 (per 1,000). This is also approximately equivalent to a change of 100 percent from the mean of influenza rate in the final sample. Based on the results of Table 4, this increase is equivalent to 1.8 and 2.7 months for age at exposure of 0-2 and 6-12, respectively. Using column 1 of Appendix Table C-1, a 100 percent increase in influenza rate is associated with 0.6 and 1 months reductions in age at death. These numbers are considerably smaller than those implied by the main results. One reason to explain this discrepancy is the observed discrepancies in the effects on different quartiles of exposure, as depicted in Appendix Figure C-1. One possible scenario is that top quartile of exposure cities have substantially higher mortality rates and higher selection, leading to much smaller effects on later life longevity, bringing down the overall effects when we examine the linear exposure measure as in Appendix Table C-1.

To complement this analysis, we replicate the results of Table 4 by interacting the age-at-exposure indicators with separate dummy variables for the second, third, and fourth quartiles of the influenza mortality distribution. The estimates are reported in Appendix Table C-2. Consistent with the patterns in Appendix Figure C-1, we find somewhat larger coefficients for the second and third quartiles relative to the top quartile. For example, for individuals exposed at ages 9–11 in second-quartile cities, we observe a sizable and statistically significant coefficient of –2.2 months.

In Appendix Table C-3, we replicate these results while restricting the sample to birth cohorts from 1907 through 1918, ensuring that the comparison reflects only Q2–Q4 cities relative to Q1 cities without incorporating older cohorts or those born after the pandemic; the resulting estimates are nearly identical to those reported in Appendix Table C-2.



Appendix Figure C-1 – Early-Life and Childhood Exposure to 1918 Influenza and Old Age Longevity by Age and by Quartile of City-Level Influenza Rate

Notes. Point estimates and 90 percent confidence intervals are depicted. Standard errors are clustered on city and year. Regressions include birth-city and birth-division by birth-year fixed effects. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.

Appendix Table C-1 - Replicating the Main Results Using Influenza Rate as a Linear Exposure Measure

	Outcome: Age-at-Death (Months)	
	(1)	(2)
Age 0-2 × Log Influenza Rate	-.12576	
	(.32437)	
Age 3-8 × Log Influenza Rate	-1.33585*	
	(.65968)	
Age 9-11 × Log Influenza Rate	-.20696	
	(.63915)	
Age 0-2 × Influenza Rate (STD)		-.47341*
		(.23453)
Age 3-8 × Influenza Rate (STD)		.10413
		(.33529)
Age 9-11 × Influenza Rate (STD)		-.69926**
		(.27664)
Observations	1471417	1471443
R-squared	.23187	.23187
Mean DV	916.384	916.384
City FE	✓	✓
Census-division-year FE	✓	✓
Controls	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table C-2 - Replicating the Main Results Using Influenza Rate as a Linear Exposure Measure

	Outcome: Age-at-Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q2	-1.18418	-1.06386	-.99347	-.95572
	(1.39009)	(1.37921)	(1.36761)	(1.36351)
Age 3-8 × Q2	-2.88994***	-2.95979***	-2.93172***	-2.90247***
	(1.00261)	(.75827)	(.79972)	(.79475)
Age 9-11 × Q2	-1.5101**	-2.18511***	-2.22237***	-2.23129***
	(.60058)	(.72009)	(.76524)	(.75107)
Age 0-2 × Q3	-.99288	-.93089	-.80621	-.73909
	(1.01287)	(.91945)	(.89979)	(.89197)
Age 3-8 × Q3	-2.37858**	-3.22243***	-3.17653***	-3.04895***
	(1.08309)	(.83058)	(.86108)	(.84766)
Age 9-11 × Q3	-.7166	-1.62652**	-1.66321**	-1.57121*
	(.86078)	(.74288)	(.77217)	(.78362)
Age 0-2 × Q4	-.81935	-.90589	-.73045	-.66122
	(.81531)	(.62924)	(.63442)	(.63394)
Age 3-8 × Q4	-2.10989*	-2.66304***	-2.61597***	-2.51915***
	(1.07154)	(.83542)	(.87257)	(.86901)
Age 9-11 × Q4	-.03103	-.7286	-.81983	-.77724
	(.98229)	(1.0209)	(1.05106)	(1.05178)
Observations	1471443	1471443	1471443	1471443
R-squared	.2312	.23138	.23188	.23189
Mean DV	916.384	916.384	916.384	916.384
City FE	✓	✓	✓	✓
Year FE	✓			
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table C-3 - Replicating the Main Results Using Influenza Rate as a Linear Exposure Measure

	Outcome: Age-at-Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q2	-2.55409*	-2.49995*	-2.50449*	-2.48285*
	(1.25258)	(1.32273)	(1.30902)	(1.30917)
Age 3-8 × Q2	-4.27654***	-4.43532***	-4.47987***	-4.46518***
	(.65065)	(.6003)	(.65901)	(.64788)
Age 9-11 × Q2	-2.88345***	-3.67107***	-3.78055***	-3.81784***
	(.46741)	(.73062)	(.78032)	(.7607)
Age 0-2 × Q3	-.37241	-.47589	-.47986	-.46129
	(.7711)	(.91031)	(.89853)	(.88021)
Age 3-8 × Q3	-1.82131**	-2.81915***	-2.90976***	-2.78002***
	(.80444)	(.84924)	(.89934)	(.87456)
Age 9-11 × Q3	-.13904	-1.22447	-1.39385	-1.30573
	(.55111)	(.80468)	(.85859)	(.86531)
Age 0-2 × Q4	-.99951	-.99138	-1.00842	-.97536
	(.65569)	(.81838)	(.836)	(.84547)
Age 3-8 × Q4	-2.31249**	-2.80222**	-2.95642**	-2.86714**
	(.92743)	(.93859)	(.98036)	(.98701)
Age 9-11 × Q4	-.26035	-.93393	-1.22774	-1.22562
	(1.00786)	(1.12809)	(1.16459)	(1.16069)
Observations	781110	781110	781110	781110
R-squared	.05716	.0573	.05792	.05792
Mean DV	926.789	926.789	926.789	926.789
City FE	✓	✓	✓	✓
Year FE	✓			
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix D

The 1918 influenza pandemic started from Spain, hence so-called Spanish flu. World War I veterans were among the first victims and also the main channel for the spread of the flu. Therefore, it is possible that returning veterans carried the virus back to their home, resulting in higher probability of infection among their families. In thse 1930 census, the veteran status variable reports whether individuals served the military during World War I. We link these records to parental records in 1940 census and use this final sample to implement some heterogeneity analyses based on father veteran status. The results are reported in Appendix Table D-1. In column 1, report the results for the full linked sample. In column 2, we restrict the sample to veteran fathers and in column 3, we focus on non-veteran subsample. We observed larger though insignificant coefficients for age at exposure of 0-2 and 9-11 for veterans' subsample but smaller coefficients for age at exposure of 3-8. However, among nonveterans, we observed significant and large effects for age of exposure of 3-8. Although the numbers of column 2 may suggest higher dosage of exposure to the disease on those aged 0-2 and 9-11, we do not observe the overall pattern of the full sample of column 1 or the main results.

Appendix Table D-1 - Heterogeneity in the Results based on Father's Veteran Status during World War I

	Outcome: Age at Death (Months)		
	Linked with 1930 Sample	Father WWI Veteran	Non-WWI Veteran
	(1)	(2)	(3)
Age 0-2 × Q 2-4	-.06205	-3.19119	.07991
	(1.14928)	(8.21546)	(1.28715)
Age 3-8 × Q 2-4	-2.1207**	.22669	-2.07547**
	(.99551)	(3.09984)	(.93174)
Age 9-11 × Q 2-4	-1.87562***	-1.03302	-1.78516***
	(.42995)	(4.97266)	(.48114)
Observations	824922	49787	775122
R-squared	.23242	.12742	.22779
Mean DV	915.519	863.926	918.831
City FE	✓	✓	✓
Census-division-year FE	✓	✓	✓
Controls	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix E

Appendix Table E-1 - Replicating the Main Results Using the Extended DMF Covering the Years 1975 – 2023

	Outcome: Age at Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q 2-4	-.95773	-1.02311	-.83914	-.84061
	(1.01886)	(1.1462)	(1.15994)	(1.17107)
Age 3-8 × Q 2-4	-2.52607**	-3.20297**	-3.16162**	-3.17681**
	(1.18494)	(1.37731)	(1.44936)	(1.44928)
Age 9-11 × Q 2-4	-.57893	-1.78339	-1.89632	-1.91138
	(.99079)	(1.45654)	(1.53143)	(1.52814)
Observations	1570412	1570412	1570412	1570412
R-squared	.14457	.14508	.14601	.14601
Mean DV	925.305	925.305	925.305	925.305
City FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix F

Appendix Table F-1 - Replicating the Main Results in Subsamples of Missing and Non-Missing Parents To Examine The Role Of Selection

	Outcome: Age at Death (Months), Subsamples:				
	Father SEI In 1940 Missing	Father SEI In 1940 Non-Missing	Mother Education In 1940 Missing	Mother Education In 1940 Non-Missing	Father SEI In Historical Censuses Missing
	(1)	(2)	(3)	(4)	(5)
Age 0-2 × Q 2-4	-.14866	-1.15299	-.98226	-.60994	.47559
	(.55641)	(1.39905)	(.63161)	(1.10542)	(1.76261)
Age 3-8 × Q 2-4	-2.84916***	-2.94592**	-2.9128***	-3.04859**	-2.11751**
	(.68)	(1.18047)	(.56416)	(1.17016)	(.86421)
Age 9-11 × Q 2-4	-1.58279**	-1.61207	-1.72536***	-1.56879	-.8396
	(.5806)	(1.07086)	(.55411)	(1.63469)	(1.0344)
Observations	851895	619548	758350	713093	201922
R-squared	.19612	.15256	.17912	.16386	.23164
Mean DV	940.481	883.249	945.026	885.924	927.345
	Father SEI In Historical Censuses Non-Missing	Father Literacy In Historical Censuses Missing	Father Literacy In Historical Censuses Non-Missing	Mother Literacy In Historical Censuses Missing	Mother Literacy In Historical Censuses Non-Missing
	(6)	(7)	(8)	(9)	(10)
Age 0-2 × Q 2-4	-.95145	.80793	-1.00707	.47691	-.95255
	(.96371)	(2.29124)	(.94293)	(1.95799)	(.91711)
Age 3-8 × Q 2-4	-2.91775***	-2.54685**	-2.85783***	-2.92032***	-2.77278***
	(.79956)	(1.07748)	(.80138)	(.81976)	(.80128)
Age 9-11 × Q 2-4	-1.65717***	-.64928	-1.71519**	-.59946	-1.65191***
	(.58154)	(.99843)	(.63075)	(1.49295)	(.57522)
Observations	1269521	223688	1247755	188118	1283325
R-squared	.23094	.24969	.22909	.25056	.22956
Mean DV	914.640	915.950	916.462	917.355	916.241
City FE	✓	✓	✓	✓	✓
Census-division-year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table F-2 - Examining the Robustness to Including Flexible Parental Controls from Historical Censuses

	Outcome: Age at Death (Months)	
	(1)	(2)
Age 0-2 × Q 2-4	-.82204	-.89441
	(.8843)	(.90506)
Age 3-8 × Q 2-4	-2.81696***	-2.94386***
	(.77822)	(.77298)
Age 9-11 × Q 2-4	-1.52625**	-1.71815**
	(.64785)	(.63972)
Observations	1471443	1471438
R-squared	.23201	.23559
Mean DV	916.384	916.384
City FE	✓	✓
Census-division-year FE	✓	✓
Controls	✓	✓
Parental controls from historical censuses	✓	✓
City FE × Parental Controls		✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
Parental controls in historical censuses include maternal literacy dummies, maternal literacy dummies, paternal socioeconomic index dummies, and missing indicators for missing values of these variables.
The regression of column 2 includes a full interaction of all dummy variables for parental controls in historical censuses as well as parental controls in 1940 census with city fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

Appendix G

Appendix Table G-1 - Including Mother Fixed Effects

	Outcome: Age at Death (Months)		
	(1)	(2)	(3)
Age 0-2 × Q 2-4	-1.52801	-2.354	-2.13302
	(1.80135)	(1.91058)	(1.9136)
Age 3-8 × Q 2-4	-3.33073	-3.31944	-3.03509
	(2.03511)	(2.15814)	(2.16826)
Age 9-11 × Q 2-4	-2.14362	-1.78817	-1.80976
	(3.94158)	(4.18484)	(4.19102)
Observations	177751	177745	177745
R-squared	.56574	.5666	.56663
Mean DV	885.950	885.945	885.945
City FE	✓	✓	✓
Year FE	✓	✓	✓
Census-division-year FE		✓	✓
City Controls			✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix H

Appendix Table H-1 - Examining Robustness across Alternative Control Cohorts

	Outcome: Age at Death (Months), Cohorts:				
	1900-1926	1902-1926	1904-1926	1906-1926	1908-1926
	(1)	(2)	(3)	(4)	(5)
Age 0-2 × Q 2-4	-.80729	-.8402	-.79994	-.98523	-1.03616
	(.86909)	(.88155)	(.88933)	(.90028)	(.93758)
Age 3-8 × Q 2-4	-2.83631***	-2.85465***	-2.80198***	-2.97881***	-3.04481***
	(.76434)	(.79158)	(.81527)	(.83785)	(.90547)
Age 9-11 × Q 2-4	-1.60063**	-1.60799**	-1.55434**	-1.72666**	-2.10911**
	(.62581)	(.65906)	(.68141)	(.72564)	(.8133)
Observations	1429037	1379615	1321218	1250997	1166676
R-squared	.21412	.1969	.18034	.16395	.14778
Mean DV	913.395	910.147	906.567	902.481	897.772
	1910-1926	1898-1918	1898-1920	1898-1922	1898-1924
	(6)	(7)	(8)	(9)	(10)
Age 0-2 × Q 2-4	-1.04016	-.18843	-.72123	-.7927	-.80732
	(.93734)	(.896)	(.8811)	(.85108)	(.87171)
Age 3-8 × Q 2-4	-3.03609***	-2.30289***	-2.845***	-2.86863***	-2.86017***
	(.90804)	(.70204)	(.69273)	(.67941)	(.74215)
Age 9-11 × Q 2-4		-1.07389*	-1.61029***	-1.63574***	-1.62781**
		(.53575)	(.55711)	(.54679)	(.59308)
Observations	1065679	973623	1104838	1235145	1356964
R-squared	.13006	.11898	.14103	.16885	.19907
Mean DV	892.083	946.774	939.338	931.537	923.930
City FE	✓	✓	✓	✓	✓
Census-division-year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix I

Appendix Table I-1 - Restricting to Cohorts of 1905-1920 and Death Ages of 70-85

	Outcome: Age at Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q 2-4	-.56257***	-.18765	-.1587	-.1597
	(.19025)	(.27703)	(.29743)	(.36937)
Age 3-8 × Q 2-4	-.9804**	-1.16263***	-1.16687***	-1.13307**
	(.37692)	(.3745)	(.38279)	(.44205)
Age 9-11 × Q 2-4	.06472	.06312	.02913	.06224
	(.16421)	(.31263)	(.31823)	(.34237)
Observations	569064	569064	569064	569064
R-squared	.00334	.00359	.00383	.00383
Mean DV	936.995	936.995	936.995	936.995
City FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table I-2 - Restricting to Cohorts of 1910-1920 and Death Ages of 65-85

	Outcome: Age at Death (Months)			
	(1)	(2)	(3)	(4)
Age 0-2 × Q 2-4	-.44255***	-.30818***	-.29432*	-.28311
	(.06061)	(.04507)	(.14416)	(.16844)
Age 3-8 × Q 2-4	-1.57691***	-2.11908***	-2.15533***	-2.11297***
	(.48653)	(.34177)	(.33184)	(.33354)
Observations	524897	524897	524897	524897
R-squared	.00314	.00331	.0038	.0038
Mean DV	911.971	911.971	911.971	911.971
City FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Census-division-year FE		✓	✓	✓
Family Controls			✓	✓
City Controls				✓

Notes. Standard errors, clustered on city and year, are in parentheses. Controls include race dummies, father's socioeconomic status dummies, maternal education dummies, county-level share of married, average socioeconomic score, and literacy rate.
*** p<0.01, ** p<0.05, * p<0.1