Unhooking the Past: Early-life Exposure to Hookworm Eradication and Later-life Longevity*

Hamid Noghanibehambari†

Jason Fletcher‡

Abstract

This study examines the long-term effects of the Rockefeller Sanitary Commission's (RSC) hookworm eradication campaign, initiated in the American South in the 1910s, on old-age longevity. Utilizing Social Security Administration death records linked to the 1940 full-count census, we employ a difference-in-differences approach to examine the effects of early-life exposure to the eradication campaign on later-life outcomes. We find that individuals exposed to the RSC campaign during in-utero and early-life experience an increase of 1.3 months in longevity. The effects are substantially larger among nonwhites, children of illiterate mothers, and those born in urban areas. Moreover, we provide evidence of dynamic complementarity in the effects of hookworm eradication on longevity, with larger effects observed in counties exposed to the Rosenwald school construction movement and in states with more stringent child labor laws. Using the 1940 census and World War II enlistment data, we provide suggestive evidence of improvements in educational attainment, income, and cognitive ability as possible pathways. Our findings contribute to the literature on the lasting effects of early-life public health interventions and underscore the importance of such programs in addressing present-day global health challenges.

Keywords: Mortality, Longevity, Hookworm, Early-Life, Disease Burden

JEL Codes: I18, J18, H75, N51, N52

---

* 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). We thank Joshua Goldstein for helpful comments.

† College of Business, Austin Peay State University, Marion St, Clarksville, TN 37040, USA
Email: noghanih@apsu.edu, Phone: +1-806-620-1812

‡ La Follette School of Public Affairs, University of Wisconsin-Madison, 1225 Observatory Drive, Madison, WI 53706-1211, USA
Email: jason.fletcher@wisc.edu

1. Introduction

The 20^{th} century witnessed remarkable improvements in human health, largely driven by advancements in public health, medical technology, and disease control. The burden of infectious diseases was gradually reduced through widespread initiatives and sanitation efforts, including vaccination campaigns, water treatment, increased public awareness, and targeted eradication programs (Armstrong et al., 1999). These efforts led to significant improvements in overall population health. Recently, a growing body of literature has begun to evaluate the long-term impacts of these public health interventions from the early 20^{th} century, providing evidence of their lasting benefits. These studies suggest that early-life and childhood exposure to these interventions are associated with improved human capital, labor market outcomes, cognitive ability, and mortality later in life (Anderson et al., 2019, 2022; Beach et al., 2016; Case et al., 2002, 2005; Case & Paxson, 2009; Cutler & Miller, 2005, 2022; Noghanibehambari & Fletcher, 2022).

In this paper, we focus on a large-scale targeted eradication program in the US South and explore its lasting impacts on individuals' longevity in old age. Specifically, we examine the hookworm eradication campaign that was initiated by the Rockefeller Sanitary Commission (RSC) in 1910. The RSC provided substantial funding, donating over $1 million to support the hookworm eradication campaign, facilitating widespread treatment, education, and infrastructure improvements in the US South.

Hookworm disease is caused by the parasitic worms *Necator americanus* and *Ancylostoma duodenale*, which typically enter the body through the skin, usually the feet, and reside in the intestines, where they feed on blood. This results in chronic anemia, malnutrition, and impaired physical and cognitive development, particularly in children. However, before

1910, there was limited understanding of the disease and its effects. The realization that many health problems among Southerners could be traced to hookworm disease was a key factor motivating the RSC's eradication campaign.

A narrow strand of research examines the effects of this campaign on the social and economic outcomes of affected areas in the short term and long term. For instance, Bleakley (2007) shows that cohorts who were exposed to hookworm eradication during childhood have higher education and literacy during adulthood. These individuals revealed higher earnings in 1940 and enjoyed higher returns to schooling for their earnings. Bleakley & Lange (2009) provide evidence that the sharp increase in the human capital of children due to hookworm eradication was accompanied by reductions in fertility, suggesting a quality-quantity trade-off in parental fertility decisions. Nonetheless, fewer studies have explored the longer-run impacts of this large-scale eradication program on health and mortality outcomes. The current study fills this gap in the literature.

We employ data from the Social Security Administration death records linked to the 1940 full-count census and implement a difference-in-difference estimation strategy to compare longevity outcomes of individuals who were born in different years relative to the starting year of the campaign in counties with differential intensity of treatment imposed by the RSC initiative. We find intent-to-treat effects of exposure during in-utero and early-life due to the eradication campaign of about a 1.3-month increase in longevity. We find considerably larger impacts among nonwhites, children of illiterate mothers, and those born in urban areas. Supplementary analyses using the 1940 census and World War II enlistment data provide suggestive evidence of improvements in education, wage income, and cognitive ability. We do not find discernible impacts of exposure during childhood despite the findings of previous literature on the improvements in their educational outcomes and the potential role of education on health and longevity. Further, we show that our findings do not pick up on differential characteristics of treated cohorts, do not reflect a trend in longevity among older children and adults, and are robust to a wide array of alternative specifications, subsamples, functional forms, and estimation strategies.

Moreover, we provide evidence of dynamic complementarity in the effects of hookworm eradication on longevity. Our findings show that the effects are more pronounced in counties that were exposed to the Rosenwald school construction movement. Such complementarity is our specifically larger among nonwhites, the target population of the Rosenwald school construction projects. Additionally, states with more stringent child labor laws also experienced larger improvements in longevity outcomes, suggesting that educational resources and policies played a critical role in strengthening the benefits of hookworm eradication.

This study makes two important contributions to the literature. First, it adds to the small literature that evaluates the impacts of the hookworm disease on outcomes, and more specifically, the RSC hookworm eradication campaign (Bleakley, 2007; Bleakley & Lange, 2009; Henderson, 2018; Ness et al., 2020; Sakti et al., 1999). Our study is the first to explore the benefits of the program on later-life longevity outcomes. Although longevity and mortality are extreme proxies of health, they are more accurately measured than many subjective measures of health. Moreover, studies show strong correlations between longevity and other economic outcomes (Blackburn & Cipriani, 2002).

Second, our study adds to the literature that examines the role of early-life disease environment on later-life health and mortality outcomes. Several studies of this literature proxy disease environment with county/state-level infant mortality rates and document correlations with later-life cognition, health, and mortality (Almond et al., 2012; Bozzoli et al., 2009; Case & Paxson, 2009; Noghanibehambari & Fletcher, 2023d). Other studies in this literature exploit natural experiment settings. Specifically, they examine the impacts of outbreaks, such as the polio pandemic of 1916 and the Spanish flu pandemic of 1918-1919, on later-life education, disability, and mortality (Almond, 2006; Almond & Mazumder, 2005; Meyers & Thomasson, 2021; Noghanibehambari & Fletcher, 2023e). However, the evidence of the long-term impacts of disease eradication and public health campaigns of the early 20th century is limited. Our study contributes to this line of research. Third, our study also contributes to the recently growing literature that evaluates the impacts of early-life conditions and exposures on later-life mortality and longevity outcomes (Aizer et al., 2016; Fletcher & Noghanibehambari, 2024; Lindeboom et al., 2010; Noghanibehambari et al., 2024; Noghanibehambari & Fletcher, 2023b, 2023a, 2024b; Van Den Berg et al., 2006; Vu et al., 2023).

Although several diseases of the early 20th century have been eradicated, many developing countries remain burdened by a variety of communicable and infectious diseases (Alirol et al., 2011; Boutayeb, 2010). These countries are experiencing similar social and economic developmental stages to those of the US in the early 20th century. Therefore, the results of this paper have policy implications for policymakers in these countries, as well as for international advocates and initiatives aiming to improve population health outcomes and eradicate diseases. Recent pandemics, such as the Ebola virus outbreak, the Middle East Respiratory Syndrome outbreak, and the COVID-19 pandemic, suggest that the spread of infectious diseases is not confined to developing countries in a globalized world (Baker et al., 2021). Therefore, understanding the long-term impacts of early-life disease environments remains highly relevant today.

The rest of the paper is organized as follows. Section 2 introduces data sources and the process of sample selection. Section 3 discusses the econometric method. Section 4 reviews the results. We conclude the paper in section 5.

2. Data and Sample

The primary source of data comes from Death Master Files (henceforth DMF) of the Social Security Administration. The DMF data contains deaths of male individuals who died between the years 1975-2005.[4] This data is extracted from the Censoc project (Breen & Osborne, 2022; Goldstein et al., 2021). The Censoc-DMF extract can be linked to the full count 1940 census at the individual level, extracted from Ruggles et al. (2020). This data linking brings two important advantages that are essential in our setting. First, it enables us to extract and infer the country of birth. Specifically, we use cross-census data linking rules provided by Price et al. (2021) to search for individuals in historical censuses 1880-1930 and observe their county of residence in early life.[5] This is an important identifier since the RSC treatment is at the county level. Second, we are also able to observe family covariates in historical censuses, which allows us to implement balancing tests and examine heterogeneity analyses. In addition, initial DMF data contains millions of observations. After sample selections described below, we still possess a large sample size and ample statistical power to detect effects. The combination of a large sample size, family information in early-life, and inferred county of birth is rare and unprecedented in alternative datasets of mortality studies.

[4] In Appendix E, we examine the effects among females using an alternative data source. We find larger effect sizes among female individuals.
[5] We provide details on the process of interfering the county of birth in Appendix A.

The information on the RSC campaign and the percentage of treated children in each county is extracted from Henderson (2018). We link this data with the DMF data based on the county of birth. We restrict the sample individuals born in the US. We further restrict the sample to individuals born between 1880 and 1940 to have many cohorts born after the eradication program and many cohorts during childhood and early adulthood. We further limit the sample to individuals born in counties where the campaign operated and counties that received RSC treatment.[6,7]

For the mechanism section, we employ data from the World War II enlistment data extracted from Goldstein et al. (2023). We focus on the Army General Classification Test (AGCT) score. The AGCT was a standardized test used by the United States Army to assess the mental and cognitive abilities and aptitudes of recruits and determine their suitability for various military occupations. We implement data merging and sample selections similar to those of the DMF data.

The top panel of Figure 1 depicts the geographic distribution of the RSC treatment across counties in the final sample. The bottom panel of this figure shows the geographic distribution of age-at-death based on the county of birth across counties in the final sample. Summary statistics of the final sample are reported in Table 1 for counties below the median and above the median percentage of campaign-induced treatment. The average age-at-death in low-treatment and high-treatment counties is 890.3 and 885.9 months, respectively. Roughly 11.3% of individuals in highly treated counties receive RSC treatment. As we discuss in section 3, the variable exposure measures the distance (in years) between birth cohorts and the year

[6] In Appendix G, we show the robustness of the results including nontreated counties in the final sample.
[7] These include all counties in the following states: Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia.

the campaign started operating, i.e., 1910. On average, exposure is 5.5 and 5.9 years for individuals in low and high-treated counties. Low-treament counties have slightly higher occupational income scores over the sample period compared with high-treatment counties, 18.5 versus 17.3. In the 1940 census, we observe a marginally lower share of low-educated individuals in high-treatment counties compared with low-treatment counties, 0.82 versus 0.83.

In the World War II enlistment data, individuals from high-treatment counties reveal an average AGCT score of 76.5, while individuals from low-treatment counties have an average score of 79.5.

3. Econometric Method

Our identification strategy is based on variations induced by the RSC campaign treatments and cross cohort age differences in exposure to those treatments. The treatment intensity highly correlates with pretreatment hookworm infection rates, which largely depend on the climatic and ecological conditions of counties (Bleakley, 2007; Bleakley & Lange, 2009). Further, the treatment timing was based on medical innovations that discovered the relevance of hookworms to various symptoms among the affected populations. In addition, the campaign was funded by a one-time donation from John Rockefeller, which makes the campaign reasonably uncorrelated with other cross-cohort characteristics and the differential evolution of their longevity in decades to come. Specifically, we operationalize these variations through the following difference-in-difference specification:

$$y_{icsb} = \alpha_0 + \sum_{j \neq -1} \beta_j Treat_{cs} \times I(b - 1910 = j) + \alpha_1 X_i + \alpha_2 Z_{csb} + \xi_{sb} + \zeta_c \times T_b + \varepsilon_{icsb} \tag{1}$$

Where y is age-at-death for individual i born in county c state s and year b. The variable $Treat$ is the percentage treatment in the county as a result of campaign efforts.[8] We standardize this variable with respect to the sample mean and standard deviation to ease interpretations. The parameter $I(.)$ is a unit function that turns on if the inside argument is true. The interaction of $Treat$ and cohort-distance dummies (i.e., $I(.)$) generate difference-in-difference estimates. Therefore, coefficients β can be interpreted as changes in longevity for cohorts born in different years relative to the campaign initiation (i.e., 1910) and born in counties with different shares of receiving treatment.

In X, we include as individual controls race and ethnicity dummies. We also include maternal literacy dummies, paternal socioeconomic status dummies, and missing indicators for missing values of these variables as parental controls. In Z, we include county characteristics extracted from historical censuses 1880-1940 and interpolated for inter-decennial years. These covariates include the share of homeowners, the share of children under age 5, the share of immigrants, the share of literate people, the share of employed, the share of employment in different industries, the share of married individuals, and average occupational income score. We also include a variable that indicates cohort distance from the county-specific year of the Boll Weevil arrival, extracted from Baker et al. (2020).[9]

[8] In Appendix F, we show that the effects of treatment and longevity are primarily concentrated in counties that revealed higher pre-eradication hookworm rate.
[9] Several studies point to the lasting impacts of the Boll Weevil on reducing child labor and improving educational outcomes (Baker, 2015; R. B. Baker et al., 2020; Noghanibehambari & Fletcher, 2024a).

The parameter ξ represents birth state by birth year fixed effects. Therefore, we exploit variations across individuals born in the same state in the same year. The parameter ζ represents county fixed effects, which absorb all time-invariant unobservable characteristics of local areas that influence population health and longevity. In our preferred specification, we also include a county-specific linear trend to account for all county characteristics that evolve linearly across cohorts. Since campaign efforts concentrated not only on treatment but also on public health education, and both treatment and education have spillover effects across other individuals, it is reasonable to assume larger benefits of the program in more populated counties. Therefore, we weight regressions using the mean county population. Finally, ε is a disturbance term. We cluster standard errors at the birth county and birth year levels to account for both serial and spatial correlations in error terms.

To summarize the coefficient sets of Equation 1, we group positive event time coefficients into an exposure measure (i.e., born after 1910, β_j $j > 0$) using the following specification:

$$y_{icsb} = \alpha_0 + \beta Treat_{cs} \times Exposure_b + \alpha_1 X_i + \alpha_2 Z_{csb} + \xi_{sb} + \zeta_c \times T_b + \varepsilon_{icsb} \qquad (2)$$

All covariates and fixed effects in this equation are similar to those of Equation 1.

4. Results

4.1. Main Results

The main results of Equation 1 are illustrated in Figure 2. Across counties with high versus low RSC-induced treatment, we do not observe discernible differences in longevity across different childhood ages compared with those who were born in 1910. However, the respective coefficients rise significantly in magnitude for those born post-1910. The overall trend in point estimates of cohorts born between 1900-1910 (i.e., labeled [-10,0]) is slightly higher than cohorts born between 1880-1899 (i.e., labeled <-10).[10] This difference may imply that improvements in education and health of these children, as suggested by prior literature, are reflected in their old age longevity. However, we avoid highlighting this interpretation because those individuals born in the late 19th century must have lived roughly twice the life expectancy of their cohorts to appear in the DMF death window (i.e., 1975-2005), a fact that adds selection and makes comparisons of these cohorts with later cohorts challenging.

On the other hand, the nontrivial change in the coefficient sizes for those born post-1910 reveals the relevance of early-life disease environment for later-life longevity as opposed to exposure to and contraction of the hookworm disease in later childhood ages. It is also possible that treatment and accompanying educational efforts of the campaign improved the general health environment and health knowledge of the affected populations, and those changes benefited future infants more than current children. We should also note that although the campaign started in 1910, it was expanded over time until 1915 with potential staggered impacts across the following years. Therefore, it is quite possible that the effects appear with a lag, and what we observe for the early years of life occurs in childhood ages. Since we do not have information on the exact dates of treatments in each county, and it is challenging to disentangle contemporaneous effects from overtime spillovers, we are unable to distinguish between in-utero, early-life, and early childhood as the critical age at which the benefits of hookworm eradication are concentrated.

[10] In Appendix D, we summarize these values using three groups of age at exposure. The results are consistent with our interpretations in this section.

The results of Equation 2 are reported in Table 2. Using the full sample and the fully parameterized specification of column 1, we find that a one standard deviation increase in treatment share for fully exposed cohorts is associated with 1.3 months higher longevity.

In columns 2-3, we examine the heterogeneity in the effects by race. The coefficient for nonwhites is roughly 1.7 times as large as that of whites. This is expected, given that hookworm disease primarily spreads through unsanitary environments, and the available sanitation infrastructure for Blacks was lower than that of whites (Elman et al., 2019).

In columns 4-5, we examine heterogeneity by maternal literacy status. The point estimate among those born to illiterate mothers is roughly 1.5 times that of literate mothers. This fact might reflect the effectiveness of the public health education efforts due to the program. Another channel to explain the heterogenous coefficients is the complementarity influence of the program in mitigating the role of other adversities experienced by children of low-educated parents in determining their long-run health and longevity (Cunha & Heckman, 2007; Huebener, 2019).

In columns 6-7, we examine heterogeneity by birthplace urban/non-urban status. A priori, we expect the program's impacts to reveal larger impacts in more populated areas due to easier spread of spillovers, especially given the multifaceted initiatives taken by the campaign. Indeed, we observe that the benefits are largely concentrated among individuals born in urban areas.

In interpreting the magnitudes, we consider two approaches. First, we emphasize that the observed point estimates reveal intent-to-treat effects across the whole population. The substantial heterogeneity observed in this table confirms that treatment-on-treated impacts could be several times larger. Second, we compare these estimates with similar early-life shocks documented by other studies using similar data in similar settings. For instance, Noghanibehambari & Fletcher (2023c) examine the impacts of early-life exposure to the Dust Bowl of the 1930s, a large-scale environmental catastrophe with lasting impacts on agricultural income, on long-term male longevity and document a reduction of roughly 1 month. The observed benefit of hookworm eradication is roughly 30% larger than the adverse effects of the Dust Bowl. Noghanibehambari & Fletcher (2024a) explore the effects of the arrival of the Boll Weevil in the American South during the early decades of the 20th century on children's later-life old-age longevity. They show that the infamous pest reduced the labor demand of children in the cotton industry and resulted in improvements in their education by roughly 0.5 years. Later in life, these children enjoyed approximately 3 months of higher longevity. The coefficients of Table 2 are between 43-110% of the effects of the Boll Weevil and associated reductions in child labor. Vu et al. (2023) examine the effects of prenatal exposure to the incidence of lynching among African Americans in the early decades of the 20th century in the American South on males' later-life longevity. They document a reduction of roughly 3.7 months. Given the disastrous influence of lynching among the affected subpopulation of black individuals, the coefficient of column 3 of Table 2 points to a relatively meaningful and large impact.

4.2. Heterogeneity by Educational Resources and Policies

As we discuss in section 4.4 and shown by Bleakley (2007), improvements in human capital and adulthood labor market outcomes are potential pathways. However, such channels may dynamically interact with available educational resources and other schooling and child labor policies to influence long-run health effects driven by hookworm eradication (Cunha & Heckman, 2007; Johnson & Jackson, 2019). One important consideration for such dynamic complementarity, especially relevant during the Jim Crow South, is the Rosenwald school construction. The Rosenwald movement was funded by philanthropist Julius Rosenwald and aimed to improve education for African American children in the rural South. Studies suggest that improvements in school resources during the early decades of the 20th century under the Rosenwald initiative improved children's education, income, migration, and long-term health outcomes (Aaronson et al., 2021; Aaronson & Mazumder, 2011; Carruthers & Wanamaker, 2013, 2017b, 2017a; Eriksson, 2018). We use data from Aaronson et al., (2021) on county-level dates of Rosenwald school construction and generate the dummy variable indicating exposure to Rosenwald school in the county of birth during the ages 5-17. We then split the sample based on exposure to Rosenwald schools and replicate the main results in three panels of Table 3 for the full sample, subsample of nonwhites, and subsample of whites, respectively. We observe higher effects of exposure to hookworm eradication and longevity of individuals who are also exposed to Rosenwald. The estimates point to a sharp difference between exposed and unexposed nonwhites to Rosenwald. As reported in columns 1-2 of panel B, we observe an increase of about 4 months for exposure to hookworm eradication among nonwhites who were also exposed to the Rosenwald school construction movement. Among nonwhites in non-Rosenwald counties, we observe an insignificant change of 0.6 months.

A prominent line of research examines the effects of schooling policies and child labor laws on educational outcomes, adulthood labor market outcomes, and various health outcomes later in life (Card, 1999; Fletcher, 2015; Lleras-Muney, 2005; Oreopoulos et al., 2006). To investigate the complementarity between these policies and exposure to hookworm eradication, we extract child labor laws from Acemoglu & Angrist (2000). Following their method, we construct a child labor law index and split the sample based on the stringency and rigidity of child labor policies. We replicate the main results for the subsamples based on the stringency of child labor laws. These results are reported in columns 3-4 and three panels of Table 3. In the full sample, we observe considerably larger effects driven from the subsample of individuals born in states with stricter child labor laws. Among nonwhites, however, we observe fairly similar coefficients across subsamples. In Panel C, for whites, we observe a coefficient that is indistinguishable from zero in the subsample of states with weak child labor laws. In this panel, the effects of hookworm eradication are primarily concentrated among individuals in states with more stringent child labor policies.

4.3. Robustness Checks

In this section, we discuss several sources of selection and endogeneity that raise concerns regarding the interpretability of the results. The first concern relates to changes in fertility patterns across different sociodemographic groups, differential selections caused by infant mortality, and selective survival into adulthood and old age. These selections could confound our estimates if characteristics of selected individuals in the final sample that correlate with their longevity also correlate with their likelihood of receiving the treatment. We empirically explore this concern using a series of balancing tests to examine the correlation between observable characteristics and exposure measures of Equation 2, conditional on fixed effects and trends. We report these results in Appendix B. Two facts arise from this analysis that mitigate such selection concerns. First, the estimated correlations are infinitesimal, allowing little space for any confounding influences. Second, the possible direction of endogeneity implied by these correlations points to an underestimation of true effects in our main results. For instance, we observe small reductions in the share of whites as a result of exposure. Since white people have, on average, higher longevity due to unobservable factors, the underrepresentation of whites pushes down the estimated benefits of the program.

The second concern is the potential endogeneity of DMF-1940-Census data linking. For instance, there is evidence that white individuals and people of higher socioeconomic status are more likely to appear in the linked DMF data who have, on average, higher longevity (Balia & Jones, 2008; Nandi et al., 2014). In that case, the results could pick up on the overrepresentation of whites due to the data merging procedure. We empirically test for this concern using the original population of male individuals in the final sample counties born between 1880-1940 observed in the full-count 1940 census. We then merge this data with our final sample and examine the association between exposure measures in Equation 2 and successful merging. These results are reported in Appendix A. The estimated correlations are economically and statistically indistinguishable from zero.

The third concern relates to the sensitivity of the results to alternative models. In Appendix C, we investigate robustness checks. First, we show the results are robust across more parsimonious specifications. Second, we show that the results are robust to adding more covariates, including interacting county fixed effects with parental dummies, adding county of death fixed effects, interacting birth state fixed effects with the 1940 state of residence fixed effects, and adding birth months and death months fixed effects. Moreover, we examine the functional form by replacing the outcome with the log of age-at-death, and indicators of longevity beyond ages 70, 75, and 80. We observe larger impacts for older ages. Specifically, among fully exposed individuals, a one standard deviation rise in treatment share is associated with increases in longevity beyond ages 70, 75, and 80 of approximately 0.5%, 0.8%, and 1.3%, with respect to the mean of the outcomes. We also show that the estimates are almost identical if we employ Sun & Abraham (2021)'s difference-in-difference estimator, suggesting that the negative weighting of ordinary least square estimators is not likely a confounding factor.

4.4. Mechanisms

The multifaceted RSC campaign can affect infants' health capital development in several ways. The campaign's educational efforts regarding increasing public knowledge about sanitation and health, as well as reforming sanitation infrastructures, could reduce disease burden during the prenatal development period. Studies in several disciplines provide evidence of the role of in-utero and early-life health conditions and disease burden on infants' health outcomes (Yang et al., 2016). For instance, the anemia caused by hookworms feeding on maternal blood reduces hemoglobin levels, impairing the oxygen-carrying capacity of the blood. Oxygen is critical for fetal development, particularly for brain growth and neural differentiation (Xiong et al., 2000). These channels of impact can influence the initial health capital of infants and change the trajectory of their developmental outcomes.

Further, direct exposure to the disease during the early years of life may trigger an immune response, causing the body to produce high levels of immune cells and inflammatory markers. This hyperactive immune system can lead to chronic low-grade inflammation, often called *inflammaging*. Over time, it increases the risks of certain diseases such as cardiovascular diseases, diabetes, and certain cancers (Fulop et al., 2018). Empirical studies suggest that early-life disease burden and disease exposure are associated with later-life health and mortality outcomes (Blackwell et al., 2001; Currie et al., 2010; Smith, 2009).

As documented by Bleakley (2007), counties that benefited more from the program experienced higher economic growth and productivity in the following years. These benefits can then be translated into higher financial resources among families with their specific pathways to affect infants' and children's health outcomes. Several studies document an association between early-life economic conditions in later-life health and mortality (Noghanibehambari et al., 2024; Schmitz & Duque, 2022; Van Den Berg et al., 2006).

The initial health endowment of infants may affect their developmental outcomes and can be detected in their cognitive ability, human capital, and later-life labor market outcomes (Black et al., 2007; Fletcher, 2011; Maruyama & Heinesen, 2020). Several strands of research examine the association between these mediatory factors and later-life mortality (Chetty et al., 2016; Fletcher, 2015; Fletcher & Noghanibehambari, 2023; Galama et al., 2018; Lleras-Muney, 2005; Lleras-Muney et al., 2022; Noghanibehambari & Fletcher, 2024c).

We can empirically test these pathways using available data from the 1940 census and the World War II enlistment records. For both samples, we implement similar sample selections and econometric methods as the main results of the paper. We further restrict the sample to individuals who are at least 24 years old to ensure they have completed their education. We report these results in Table 4. Although we do not find significant effects on years of schooling (column 1), we observe reductions in the share of individuals with less than 12 years of schooling (column 2). Further, we find increases in individuals with 12 years of education. However, the point estimate is small and statistically insignificant (column 3). Further, we observe significant increases in college education (column 4). Among fully exposed individuals, a one standard deviation rise in treatment share is associated with an 16 basis-point increase in the probability of any college education, equivalent to roughly 1.7% change with respect to the mean of the outcome. Fletcher & Noghanibehambari (2023) examine the effects of college openings on college education and longevity. They document that having any college education is associated with 1.3-2.7 years of higher longevity. Using their estimates and the calculated percentage change of column 3, we show that the longevity gains through improvements in educational outcomes due to exposure to the eradication program are around 0.26-0.55 months, roughly 20-42% of the observed reduced form effect of Table 2.

Further, we observe a 2.7% increase in wage income (column 5). Finally, we observe significant increases in the AGCT score of World War II enlistees, a proxy for aptitude and cognitive ability. Among fully exposed individuals, a one standard deviation increase in treatment share is associated with 1.3 units higher AGCT score, off a mean of 73.5.

The positive, small, and statistically insignificant effect on years of schooling is consistent with the effects reported by Bleakley (2007). He further uses prior censuses and shows significant improvements in literacy and school enrollment. In addition, he reports improvements in income. His estimates of income are quite comparable in magnitude to those reported in column 5 of Table 4.

5. Conclusion

Although many infectious diseases were eradicated in the 20th century, they continue to play a critical role in shaping health and economic outcomes in today's globalized economy (Baker et al., 2021). Increased global connectivity through trade, travel, and migration has increased risks of disease burden and cross-border disease transmission. Policy interventions remain instrumental but costly and challenging. Policy debates usually underscore the balance between immediate costs and benefits. However, there is growing evidence of the long-term effects of exposure to disease burden and intervention programs, especially during critical ages of development and among vulnerable populations. This study added to this policy debate by providing evidence of the long-term effects of a large-scale disease eradication program.

We focused on the case of hookworm eradication in the American South during the 1910s as a result of the Rockefeller Sanitation Commission campaign. We investigated the effects of early-life exposure to the RSC program on later-life male longevity. We found that full exposure to the eradication campaign resulted in an increase of 1 month in longevity. The benefits of the program were not uniformly distributed. We found larger impacts among nonwhite individuals, children of illiterate mothers, and those born in populated urban areas. Using the 1940 census and World War II enlistment data, we found significant increases in measures of educational attainment, income, and cognitive ability as pathways between early-life exposure to the program and later-life longevity. Moreover, we provide evidence of dynamic complementarity in the effects of hookworm eradication on longevity. The effects are more pronounced in counties exposed to the Rosenwald school construction movement and in states with stricter child labor laws, indicating that educational resources and policies amplified the benefits of eradication efforts. Such complementarity is specifically larger among nonwhites, the target population of the Rosenwald school construction projects.

To understand the magnitude of the findings, we calculate the equivalent monetary value of the gains in longevity using Value of Statistical Life (VSL) estimates. Studies suggest a VSL of $10 million (in 2020 dollars) for US residents (Kniesner & Viscusi, 2019). In the final sample, the average longevity is 888.2 months. To put it simply, each month corresponds to $11.26 thousand. Since the exposure measure in our main results is based on standardized treatment, it is safe to assume that half the individuals (above median treatment share) receive a value of one and the other half (below median treatment share) receive a value of zero.

Further, roughly 66% of the individuals in the final sample are born post-1910. Therefore, we calculate gains in longevity for exposed individuals born in above-median treatment share counties of about 382.6 thousand months. Using this number and the monetary equivalent calculated above, we estimate roughly $4.3 billion gained from the program by increasing longevity. We should note that the program's cost was $27 million (in 2020 dollars).

References

Aaronson, D., & Mazumder, B. (2011). The impact of Rosenwald schools on black achievement. *Journal of Political Economy*, *119*(5), 821–888. https://doi.org/10.1086/662962/SUPPL_FILE/2010282DATA.ZIP

Aaronson, D., Mazumder, B., Sanders, S. G., & Taylor, E. J. (2021). Estimating the effect of school quality on mortality in the presence of migration: Evidence from the jim crow south. *Journal of Labor Economics*, *39*(2), 527–558. https://doi.org/10.1086/709783/SUPPL_FILE/17462DATA.ZIP

Acemoglu, D., & Angrist, J. (2000). How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual*, *15*, 9–59. https://doi.org/10.1086/654403

Aizer, A., Early, N., Eli, S., Imbens, G., Lee, K., Lleras-Muney, A., & Strand, A. (2024). The Lifetime Impacts of the New Deal's Youth Employment Program. *The Quarterly Journal of Economics*, 1–57. https://doi.org/10.1093/QJE/QJAE016

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

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases*, *11*(2), 131–141. https://doi.org/10.1016/S1473-3099(10)70223-1/ASSET/9EB58DC3-CE43-4CE6-B726-9DD3A6FAA12F/MAIN.ASSETS/GR3.JPG

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., & Herrmann, M. (2012). From infant to mother: Early disease environment and future maternal health. *Labour Economics*, *19*(4), 475–483. https://doi.org/10.1016/J.LABECO.2012.05.015

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

Anderson, D. M., Charles, K. K., McKelligott, M., & Rees, D. I. (2022). Estimating the Effects of Milk Inspections on Infant and Child Mortality, 1880−1910. *AEA Papers and Proceedings*, *112*, 188–192. https://doi.org/10.1257/PANDP.20221066

Anderson, D. M., Charles, K. K., Olivares, C. L. H., & Rees, D. I. (2019). Was the First Public Health Campaign Successful? *American Economic Journal: Applied Economics*, *11*(2), 143–175. https://doi.org/10.1257/APP.20170411

Armstrong, G. L., Conn, L. A., & Pinner, R. W. (1999). Trends in Infectious Disease Mortality in the United States During the 20th Century. *JAMA*, *281*(1), 61–66. https://doi.org/10.1001/JAMA.281.1.61

Baker, R. B. (2015). From the Field to the Classroom: The Boll Weevil's Impact on Education in Rural Georgia. *The Journal of Economic History*, *75*(4), 1128–1160. https://doi.org/10.1017/S0022050715001515

Baker, R. B., Blanchette, J., & Eriksson, K. (2020). Long-Run Impacts of Agricultural Shocks on Educational Attainment: Evidence from the Boll Weevil. *The Journal of Economic History*, *80*(1), 136–174. https://doi.org/10.1017/S0022050719000779

Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L. F., Wesolowski, A., & Metcalf, C. J. E. (2021). Infectious disease in an era of global change. *Nature Reviews Microbiology 2021 20:4*, *20*(4), 193–205. https://doi.org/10.1038/s41579-021-00639-z

Balia, S., & Jones, A. M. (2008). Mortality, lifestyle and socio-economic status. *Journal of Health Economics*, *27*(1), 1–26. https://doi.org/10.1016/J.JHEALECO.2007.03.001

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

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, *122*(1), 409–439. https://doi.org/10.1162/qjec.122.1.409

Blackburn, K., & Cipriani, G. Pietro. (2002). A model of longevity, fertility and growth. *Journal of Economic Dynamics and Control*, *26*(2), 187–204. https://doi.org/10.1016/S0165-1889(00)00022-1

Blackwell, D. L., Hayward, M. D., & Crimmins, E. M. (2001). Does childhood health affect chronic morbidity in later life? *Social Science & Medicine*, *52*(8), 1269–1284. https://doi.org/10.1016/S0277-9536(00)00230-6

Bleakley, H. (2007). Disease and Development: Evidence from Hookworm Eradication in the American South. *The Quarterly Journal of Economics*, *122*(1), 73–117. https://doi.org/10.1162/QJEC.121.1.73

Bleakley, H., & Lange, F. (2009). Chronic Disease Burden and the Interaction of Education, Fertility, and Growth. *The Review of Economics and Statistics*, *91*(1), 52–65. https://doi.org/10.1162/REST.91.1.52

Boutayeb, A. (2010). The Burden of Communicable and Non-Communicable Diseases in Developing Countries. *Handbook of Disease Burdens and Quality of Life Measures*, 531–546. https://doi.org/10.1007/978-0-387-78665-0_32

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

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

Card, D. (1999). The Causal Effect of Education on Earnings. *Handbook of Labor Economics*, *3*(1), 1801–1863. https://doi.org/10.1016/S1573-4463(99)03011-4

Carruthers, C. K., & Wanamaker, M. H. (2013). Closing the gap? The effect of private philanthropy on the provision of African-American schooling in the U.S. south. *Journal of Public Economics*, *101*(1), 53–67. https://doi.org/10.1016/J.JPUBECO.2013.02.003

Carruthers, C. K., & Wanamaker, M. H. (2017a). Returns to school resources in the Jim Crow South. *Explorations in Economic History*, *64*, 104–110.

https://doi.org/10.1016/J.EEH.2017.02.004

Carruthers, C. K., & Wanamaker, M. H. (2017b). Separate and unequal in the labor market: Human capital and the jim crow wage gap. *Journal of Labor Economics*, *35*(3), 655–696. https://doi.org/10.1086/690944/SUPPL_FILE/15175DATA.ZIP

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, *24*(2), 365–389. https://doi.org/10.1016/J.JHEALECO.2004.09.008

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

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

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47. https://doi.org/10.1257/AER.97.2.31

Currie, J., Stabile, M., Manivong, P., & Roos, L. L. (2010). Child Health and Young Adult Outcomes. *Journal of Human Resources*, *45*(3), 517–548. https://doi.org/10.3368/JHR.45.3.517

Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography 2005 42:1*, *42*(1), 1–22. https://doi.org/10.1353/DEM.2005.0002

Cutler, D., & Miller, G. (2022). Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality: Comment. *American Economic Journal: Applied Economics*, *14*(2), 158–165. https://doi.org/10.1257/APP.20190711

Elman, C., McGuire, R. A., & London, A. S. (2019). Disease, plantation development, and race-related differences in fertility in the early Twentieth-Century American South. *American Journal of Sociology*, *124*(5), 1327–1371. https://doi.org/10.1086/702008/ASSET/IMAGES/LARGE/FG4.JPEG

Eriksson, K. (2018). Education and Incarceration in the Jim Crow South: Evidence from Rosenwald Schools. *Journal of Human Resources*, *55*(1), 43–75. https://doi.org/10.3368/JHR.55.2.0816.8142R

Fletcher, J. M. (2011). The medium term schooling and health effects of low birth weight: Evidence from siblings. *Economics of Education Review*, *30*(3), 517–527. https://doi.org/10.1016/j.econedurev.2010.12.012

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. (2023). The effects of education on mortality: Evidence using college expansions. *Health Economics*. https://doi.org/10.1002/HEC.4787

Fletcher, J., & Noghanibehambari, H. (2024). The siren song of cicadas: Early-life pesticide exposure and later-life male mortality. *Journal of Environmental Economics and Management*, *123*, 102903. https://doi.org/10.1016/J.JEEM.2023.102903

Fulop, T., Witkowski, J. M., Olivieri, F., & Larbi, A. (2018). The integration of inflammaging in age-related diseases. *Seminars in Immunology*, *40*, 17–35. https://doi.org/10.1016/J.SMIM.2018.09.003

Galama, T., Lleras-Muney, A., & Kippersluis, H. van. (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. *Oxford Research Encyclopedia of Economics and Finance*. https://doi.org/10.1093/ACREFORE/9780190625979.013.7

Goldstein, J. R., Alexander, M., Breen, C., Miranda González, A., Menares, F., Osborne, M., Snyder, M., & Yildirim, U. (2021). Censoc Project. In *CenSoc Mortality File: Version 2.0. Berkeley: University of California*. https://censoc.berkeley.edu/data/

Goldstein, J. R., Breen, C., Alexander, M., Miranda González, A., Menares, F., Osborne, M., Snyder, M., Yildirim, U., & Wickle, A. (2023). *CenSoc Army Enlistment Records*. https://doi.org/10.7910/DVN/ZFVVNA

Henderson, J. A. (2018). Hookworm Eradication as a Natural Experiment for Schooling and Voting in the American South. *Political Behavior*, *40*(2), 467–494. https://doi.org/10.1007/S11109-017-9408-6/FIGURES/4

Huebener, M. (2019). Life expectancy and parental education. *Social Science & Medicine*, *232*, 351–365. https://doi.org/10.1016/J.SOCSCIMED.2019.04.034

Johnson, R., & Jackson, K. (2019). Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending. *American Economic Journal: Economic Policy*, *11*(4), 310–349. https://doi.org/10.1257/POL.20180510

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

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

Lleras-Muney, A., Price, J., & Yue, D. (2022). The association between educational attainment and longevity using individual-level data from the 1940 census. *Journal of Health Economics*, *84*, 102649. https://doi.org/10.1016/J.JHEALECO.2022.102649

Maruyama, S., & Heinesen, E. (2020). Another look at returns to birthweight. *Journal of Health Economics*, *70*, 102269. https://doi.org/10.1016/j.jhealeco.2019.102269

Meyers, K., & Thomasson, M. A. (2021). Can pandemics affect educational attainment? Evidence from the polio epidemic of 1916. *Cliometrica*, *15*(2), 231–265. https://doi.org/10.1007/S11698-020-00212-3/TABLES/14

Nandi, A., Glymour, M. M., & Subramanian, S. V. (2014). Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. *Epidemiology*, *25*(2), 170–177. https://doi.org/10.1097/EDE.0000000000000038

Ness, T. E., Agrawal, V., Bedard, K., Ouellette, L., Erickson, T. A., Hotez, P., & Weatherhead, J. E. (2020). Maternal Hookworm Infection and Its Effects on Maternal Health: A Systematic Review and Meta-Analysis. *The American Journal of Tropical Medicine and Hygiene*, *103*(5), 1958. https://doi.org/10.4269/AJTMH.20-0503

Noghanibehambari, H., & Engelman, M. (2022). Social insurance programs and later-life mortality: Evidence from new deal relief spending. *Journal of Health Economics*, *86*. https://doi.org/10.1016/J.JHEALECO.2022.102690

Noghanibehambari, H., & Fletcher, J. (2023a). In utero and childhood exposure to alcohol and old age mortality: Evidence from the temperance movement in the US. *Economics & Human Biology*, *50*, 101276. https://doi.org/10.1016/J.EHB.2023.101276

Noghanibehambari, H., & Fletcher, J. (2023b). Long-Term Health Benefits of Occupational Licensing: Evidence from Midwifery Laws. *Journal of Health Economics*, *92*, 102807. https://doi.org/10.1016/J.JHEALECO.2023.102807

Noghanibehambari, H., & Fletcher, J. (2024a). A Blessing in Disguise: The Long-Term Effects of Childhood Exposure to the boll weevil on Old-Age Longevity. *Working Paper*.

Noghanibehambari, H., & Fletcher, J. (2024b). Dust to Feed, Dust to Gray: The Effect of in Utero Exposure to the Dust Bowl on Old-Age Longevity. *Demography*. https://doi.org/10.1215/00703370-11140760

Noghanibehambari, H., & Fletcher, J. (2024c). Unequal before death: The effect of paternal education on children’s old-age mortality in the United States. *Population Studies*. https://doi.org/10.1080/00324728.2023.2284766

Noghanibehambari, H., & Fletcher, J. M. (2022). Water is Life, Clean Water is Health: The Effect of Early-Life Exposure to the City-Wide Water Filtration on Old-Age Mortality. *Working Paper*.

Noghanibehambari, H., & Fletcher, J. M. (2023c). Dust to Feed, Dust to Grey: The Effect of In-Utero Exposure to the Dust Bowl on Old-Age Longevity. *Demography*. https://doi.org/10.3386/W30531

Noghanibehambari, H., & Fletcher, J. M. (2023d). The Long Shadow of the Past: Early-Life Disease Environment and Later-Life Mortality. *SSRN Electronic Journal*.

https://doi.org/10.2139/SSRN.4597720

Noghanibehambari, H., & Fletcher, J. M. (2023e). Walking with Feet Tied to the Past: Childhood Exposure to the 1918 Influenza Pandemic and Later-Life Male Mortality. *Working Paper*.

Noghanibehambari, H., Fletcher, J., Schmitz, L., Duque, V., & Gawai, V. (2024). Early-life economic conditions and old-age male mortality: evidence from historical county-level bank deposit data. *Journal of Population Economics*, *37*(1), 1–33. https://doi.org/10.1007/S00148-024-01007-W/TABLES/7

Oreopoulos, P., Page, M. E., & Stevens, A. H. (2006). The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, *24*(4), 729–760. https://doi.org/10.1086/506484/ASSET/IMAGES/LARGE/FG2.JPEG

Price, J., Buckles, K., Van Leeuwen, J., & Riley, I. (2021). Combining family history and machine learning to link historical records: The Census Tree data set. *Explorations in Economic History*, *80*, 101391. https://doi.org/10.1016/J.EEH.2021.101391

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

Sakti, H., Nokes, C., Hertanto, W. S., Hendratno, S., Hall, A., Bundy, D. A. P., & Satoto. (1999). Evidence for an association between hookworm infection and cognitive function in Indonesian school children. *Tropical Medicine & International Health*, *4*(5), 322–334. https://doi.org/10.1046/J.1365-3156.1999.00410.X

Schmitz, L. L., & Duque, V. (2022). In utero exposure to the Great Depression is reflected in late-life epigenetic aging signatures. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(46), e2208530119. https://doi.org/10.1073/PNAS.2208530119/SUPPL_FILE/PNAS.2208530119.SAPP01.PDF

Smith, J. P. (2009). 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

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

Vaiserman, A. (2021). Season-of-birth phenomenon in health and longevity: epidemiologic evidence and mechanistic considerations. *Journal of Developmental Origins of Health and Disease*, *12*(6), 849–858. https://doi.org/10.1017/S2040174420001221

Van Den Berg, G. J., Lindeboom, M., Portrait, F., Berg, G. J. Van Den, Lindeboom, M., Portrait, F., den Berg, G. J., Lindeboom, M., & Portrait, F. (2006). Economic Conditions Early in Life and Individual Mortality. *American Economic Review*, *96*(1), 290–302. https://doi.org/10.1257/000282806776157740

Vu, H., Noghanibehambari, H., Fletcher, J., & Green, T. (2023). Prenatal Exposure to Racial Violence and Later Life Mortality among Males: Evidence from Lynching. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.4613372

Xiong, X., Buekens, P., Alexander, S., Demianczuk, N., & Wollast, E. (2000). Anemia during pregnancy and birth outcome: A meta-analysis. *American Journal of Perinatology*, *17*(3), 137–146. https://doi.org/10.1055/S-2000-9508/ID/35/BIB

Yang, I., Corwin, E. J., Brennan, P. A., Jordan, S., Murphy, J. R., & Dunlop, A. (2016). The Infant Microbiome: Implications for Infant Health and Neurocognitive Development. *Nursing Research*, *65*(1), 76. https://doi.org/10.1097/NNR.0000000000000133

Tables

Table 1 - Summary Statistics

	Below Median Treatment Share		***Above Median Treatment Share***	
	Mean	Std. Dev.	Mean	Std. Dev.
DMF Data:				
Death Age Month	890.3695	136.9175	885.9761	136.9532
Log Age-at-death	6.7789	.1628	6.7739	.1634
Age-at-death > 70	.6221	.4849	.61	.4878
Age-at-death > 75	.461	.4985	.4473	.4972
Age-at-death > 80	.2976	.4572	.2853	.4515
Birth Year	1915.5397	12.592	1915.9017	12.5153
Death Year	1989.7464	8.8995	1989.7407	8.8773
White	.7417	.4377	.7706	.4205
Nonwhite	.2573	.4372	.2266	.4186
Share Treated	.01	.0093	.1126	.0894
Share Treated (STD)	-.6313	.1158	.6425	1.1097
Exposure	5.5397	12.592	5.9017	12.5153
Exposure × Treatment Share (STD)	-.4127	.314	.4217	.9464
Father SEI	20.7499	17.6819	18.1935	15.2941
Father SEI Missing	.1763	.3811	.1455	.3526
Mother Literate	.6288	.4831	.6394	.4802
Mother Literacy Missing	.2963	.4566	.2712	.4446
County Covariates:				
Share Of Homeownership	.4152	.1385	.4849	.1376
sShare Under Five Children	.5653	.1688	.6915	.1429
Share of Immigrants	.0219	.0372	.0064	.0118
Average Occupational Income Score	18.1571	3.0836	17.3762	2.2798
Share of Literate	.787	.1762	.7731	.163
Share of Employed	.5823	.0937	.5733	.0936
Share Employed in Manufacturing	.1049	.0953	.1087	.0965
Share Employed in Mining	.013	.0463	.0223	.0726
Share Employed in Construction	.0342	.0238	.0272	.0192
Share Employed in Transportation	.0488	.0437	.0346	.0301
Share Employed in Farming	.2293	.1223	.261	.0937
Share of Married	.6058	.0364	.6139	.0324
Exposure To Boll Weevil	.4724	.4992	.3703	.4829
Observations	446,817		445,023	
Mechanism Sample from the 1940 Census:				
Education < 12 Years	.8165	.3871	.8334	.3726
Education = 12 Years	.0967	.2956	.0879	.2832
Education > 12 Years	.0868	.2815	.0787	.2693
Wage and Salary Income	624.7487	796.9377	566.4442	738.3595
Observations	1,307,119		1,295,185	
Mechanism Sample from the WWII Enlistment Data:				
AGCT Score	76.5796	38.2083	79.5498	36.4589
Observations	76,323		84,236	

Notes. STD stands for standardized values. The DMF data covers birth cohorts of 1880-1940. The data for the mechanics and samples cover birth cohorts of 1880-1915.

Table 2 - The Impacts of Early-life Exposure to the Hookworm Eradication and Later-life Longevity

	Outcome: Age-at-Death (Months), subsamples:						
	Full Sample	Whites	Nonwhites	Literate Mother	Illiterate Mother	Born in Urban Areas	Born in Non-Urban Areas
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure × Treatment Share (STD)	1.25503***	1.27166**	2.14623**	.9891*	1.47023*	3.33672**	.87533*
	(.42828)	(.49557)	(.93858)	(.54934)	(.81422)	(1.44974)	(.45863)
Observations	891840	674339	215829	565509	326316	151887	739946
R-squared	.52321	.51586	.54788	.39074	.55742	.50174	.53549
Mean DV	860.862	864.330	848.683	896.356	810.891	850.655	866.831

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, a county-specific linear trend in birth year, and covariates. Individual and parental controls include dummies for individual race and dummies for paternal socioeconomic index and maternal literacy. County controls include dummies for individual race, county-level homeownership rate, the proportion of children under five, the proportion of immigrants, the literacy rate, the employment rate, the share employed in the manufacturing sector, the share employed in the construction sector, the share employed in the transportation sector, the share employed in the farming sector, the proportion of married individuals, exposure to the Boll Weevil, and the average occupational income score. Regressions are weighted by the mean county-level population. The data covers males born between 1880 and 1940 who died between 1975 and 2005.

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

Table 3 - Heterogeneity in the Effects Based on Exposure to County-Level Rosenwald School Construction and State-Level Child Labor Laws

	Outcome: Age-at-Death (Months)			
	Rosenwald Exposure > 0	Rosenwald Exposure = 0	Child Labor ≥ 7	Child Labor < 7
	(1)	(2)	(3)	(4)
Panel A. Full Sample:				
Exposure × Treatment Share (STD)	2.3391***	1.41202**	1.74758*	.63365
	(.81222)	(.69416)	(.96062)	(1.66411)
Observations	268213	623627	536585	240306
R-squared	.09284	.53513	.43994	.35301
Mean DV	935.141	830.732	828.034	894.377
Panel B. Nonwhites:	4.00606**	.59581	3.39122*	3.50542
Exposure × Treatment Share (STD)	(1.74135)	(1.6865)	(2.04841)	(3.56992)
	67165	150326	129538	59310
Observations	.11247	.55208	.42604	.33995
R-squared	930.706	816.544	803.906	896.818
Mean DV				
Panel C. Whites:				
Exposure × Treatment Share (STD)	2.29568**	1.76524**	1.64717	.01937
	(.95124)	(.77391)	(1.11488)	(1.89394)
Observations	201047	473292	407043	180989
R-squared	.09071	.53005	.43946	.3602
Mean DV	936.356	834.808	834.545	893.623

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.

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

Table 4 - Exploring Mechanisms Channels Using the 1940 Census and the World War II Enlistment Data

	Outcomes:					
	Years of Schooling	Education < 12 Years	Education = 12 Years	Education > 12 Years	Log Wage Income	AGCT Score
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure × Treatment Share (STD)	.0005184	-.001966**	.0003331	.0016329**	.0270171**	1.0384823*
	(.0094082)	(.0009643)	(.0007536)	(.0007264)	(.0126609)	(.6036362)
Observations	2602304	2602304	2602304	2602304	2399128	160519
R-squared	.257693	.128804	.0588394	.0675331	.0855907	.4274437
Mean DV	7.486	0.798	0.107	0.095	3.429	75.292

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, a county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, county-level homeownership rate, the proportion of children under five, the proportion of immigrants, the literacy rate, the employment rate, the share employed in the manufacturing sector, the share employed in the construction sector, the share employed in the transportation sector, the share employed in the farming sector, the proportion of married individuals, exposure to the Boll Weevil, and the average occupational income score. Regressions are weighted by the mean county-level population. The data covers males born between 1880 and 1915.

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

Figures

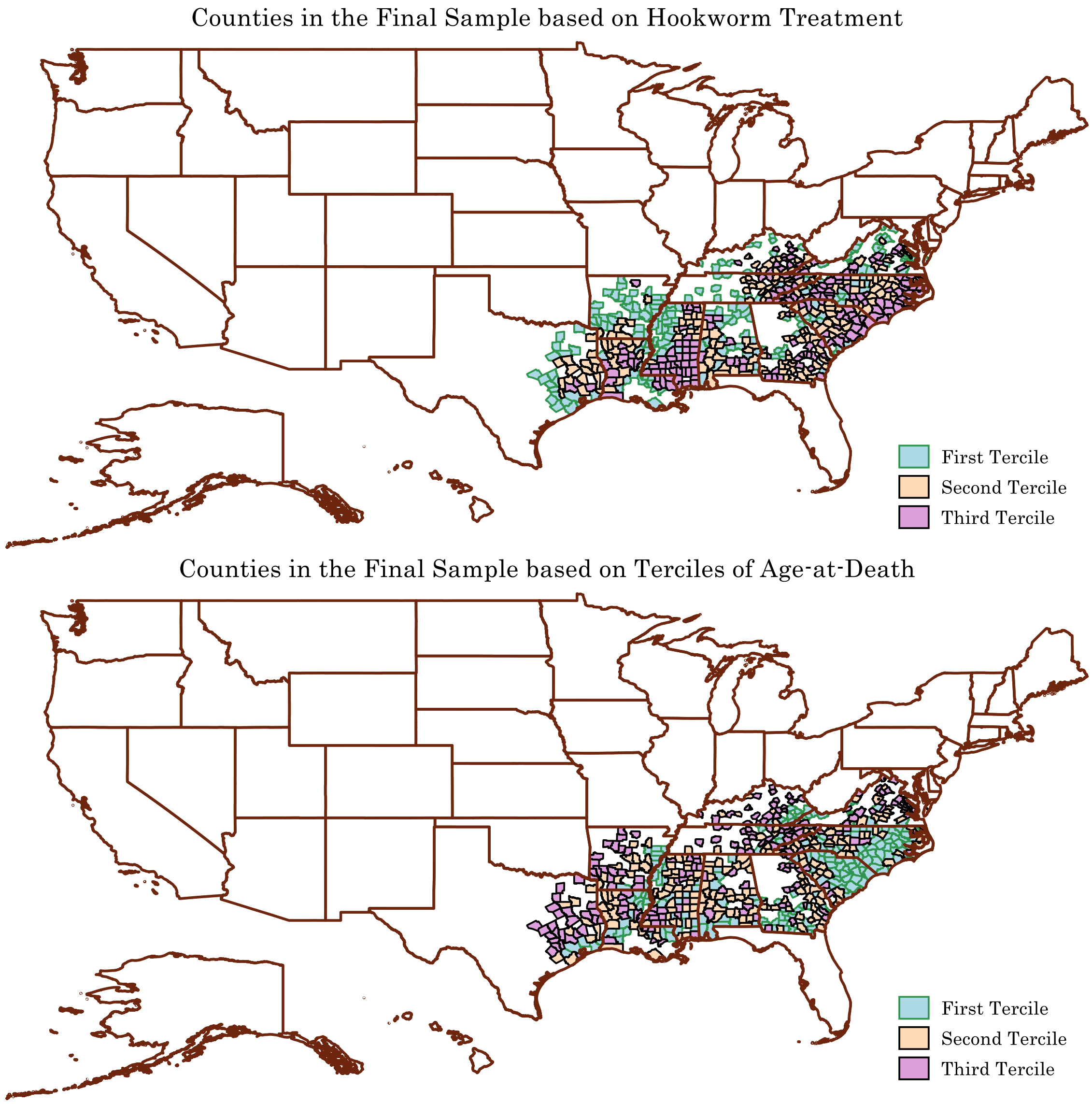

Figure 1 - Geographic Distribution of Treatment Exposure and Age-at-Death across Counties in the Final Sample

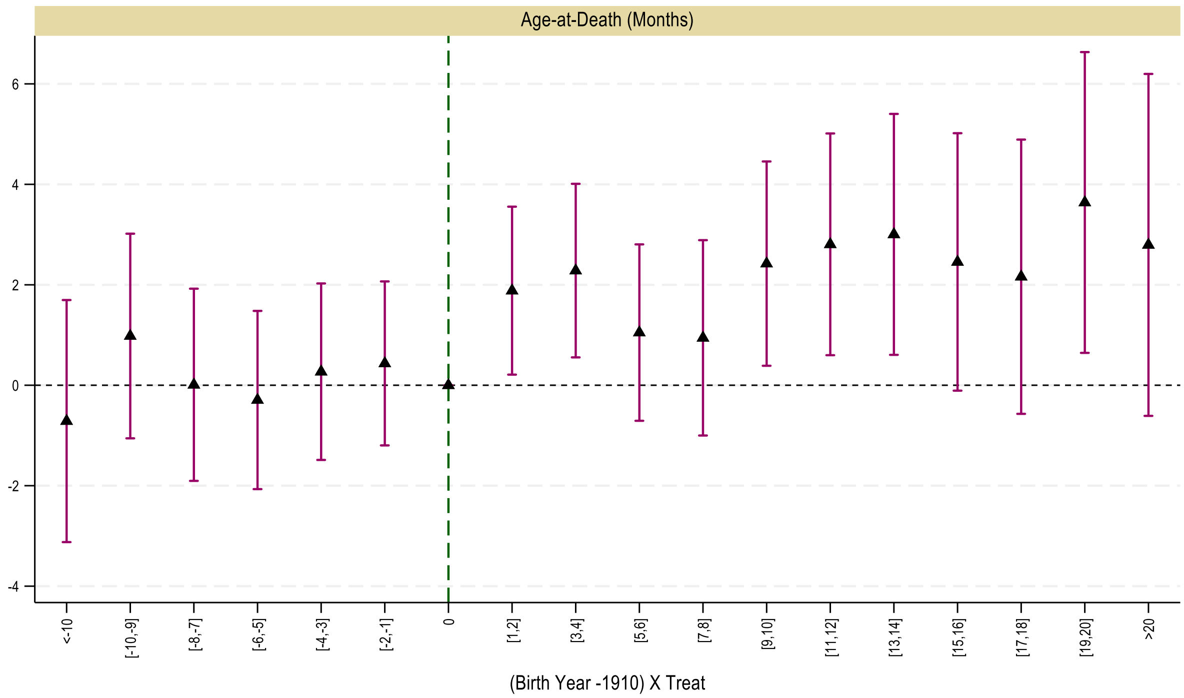

Figure 2 - Event Study Analyses to Examine Changes in Longevity across Different Cohorts with Different Treatment Exposure

Notes. Point estimates and 95% confidence intervals are reported. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, a county-specific linear trend in birth year, and covariates. Individual and parental controls include dummies for individual race and dummies for paternal socioeconomic index and maternal literacy. County controls include dummies for individual race, county-level homeownership rate, the proportion of children under five, the proportion of immigrants, the literacy rate, the employment rate, the share employed in the manufacturing sector, the share employed in the construction sector, the share employed in the transportation sector, the share employed in the farming sector, the proportion of married individuals, exposure to the Boll Weevil, and the average occupational income score. Regressions are weighted by the mean county-level population. The data covers males born between 1880 and 1940 who died between 1975 and 2005.

Appendix A

A.1 - Inferring County of Birth

The process for inferring the county of residence during early-life is as follows. First, a portion of the DMF data is also recorded in the Numerical Identification (Numident) death records. Although the Numident data covers a shorter period (1988-2005), it includes the county of birth as reported in individuals' Social Security records. For observations present in both the DMF and Numident datasets, we assign the county of birth based on the Numident data, covering roughly 49% of the observations.

Second, we apply cross-census linking rules from Price et al. (2021) to connect 1940 census records to full-count historical censuses from 1900-1930. Approximately 4.6% of observations are linked to the 1900 census, where the county of residence in 1900 is used as the early-life county. Similarly, 8.6%, 11.7%, and 13.4% of observations are linked to the 1910, 1920, and 1930 censuses, respectively, with the county of residence in each census used as the early-life county. For the 1940 census, which includes information on the county of residence five years prior (1935), we assign the county of residence in 1935 for individuals who had migrated within the past five years. This method accounts for 1.5% of observations. Finally, for all unlinked cases, we use the county of residence in 1940 as a proxy for early-life, covering about 10.5% of the observations.

A.2-Comparing with the Original Population

Appendix Table A-1 reports a summary of selected statistics in the final sample versus the original population of the 1940 census. The final sample contains relatively older individuals, a higher share of whites, and a lower share of nonwhite people. However, in terms of family characteristics and county covariates, both samples reveal quite similar features.

To examine whether differences in characteristics of the linked DMF data confound the estimates of the main results of the paper, we merge the final sample and the original population. We then generate a successful merging indicator and investigate whether the exposure measure of Equation 2 is correlated with this indicator, conditional on fixed effects and trends. These results are reported in Appendix Table A-2. Exposure is associated with a lower likelihood of being in the final sample in the full sample, the subsample of whites, and the subsample of nonwhites. However, the point estimates are quite small in magnitude. For instance, in the full sample, the coefficient implies a change of roughly 1.9% with respect to the mean of the outcome. This is too small economically to raise concerns regarding endogenous selection into the final sample from the original population.

Appendix Table A-1 - Comparing Characteristics of the Original Population and the Final Sample

	Original population		***Final sample***	
	Mean	Std. Dev.	Mean	Std. Dev.
Birth Year	1917.7274	14.00272	1915.7204	12.55508
White	.63824	.48051	.75612	.42942
Nonwhite	.35988	.47996	.24201	.4283
Exposure × Treatment Share (STD)	.00124	.81877	.01654	.82546
Mother Education < High School	.36928	.48261	.362	.48058
Mother Education = High School	.11149	.31474	.09592	.29448
Mother Education College-More	.02298	.14985	.02006	.1402
Mother Education Missing	.49624	.49999	.52203	.49951
Father Socioeconomic Index	18.63932	17.15047	19.316	17.51443
Father Socioeconomic Index Missing	.54384	.49807	.56613	.49561
County Covariates:				
Share Of Homeownership	.41159	.12905	.44961	.1424
Share Under Five Children	.59473	.16689	.6279	.16875
Share of Immigrants	.01277	.02649	.01421	.02877
Average Occupational Income Score	17.59521	2.67395	17.77026	2.7451
Share of Literate	.71824	.2253	.78005	.16992
Share of Employed	.5933	.0977	.57795	.09378
Share Employed in Manufacturing	.10886	.09493	.10691	.09599
Share Employed in Mining	.00872	.03632	.01766	.06093
Share Employed in Construction	.03204	.0232	.03075	.02191
Share Employed in Transportation	.03883	.03627	.04173	.0383
Share Employed in Farming	.24421	.10832	.24502	.11027
Share of Married	.60908	.03336	.60982	.03471
Exposure To Boll Weevil	.56395	.49589	.42196	.49387
Observations	6,681,150		891,840	

Notes. STD stands for standardized values. Mother education and father socioeconomic status variables are based on the 1940 census. The samples cover birth cohorts of 1880-1940. The samples are restricted to male individuals only.

Appendix Table A-2 - Exploring the Association between Early-life Exposure to Hookworm Eradication and Successful Merging between 1940 Census and DMF

	Outcome: Successful Merging Indicator		
	Full sample	Whites	Nonwhites
	(1)	(2)	(3)
Age at Exposure ≤ 0 × Treatment Share	-.00263***	-.00418***	-.00161*
(STD)	(.00062)	(.00081)	(.0009)
Observations	6681150	4264167	2416983
R-squared	.2798	.32086	.12258
Mean DV	0.133	0.155	0.088

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal education (extracted from the 1940 census), county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.

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

Appendix B

One concern in interpreting the results of the paper is the endogenous changes in the sample sociodemographic composition as a result of endogenous changes in fertility, selective infant mortality, and selective survival into adulthood. We empirically test this concern by regressing several observable characteristics of the final sample and exposure measures of Equation 2, conditional on fixed effects and trends. These results are reported in Appendix Table B-1. Exposure is associated with a reduction in the probability of being white, a reduction in the father's socioeconomic index, and a reduction in the likelihood of a literate mother. However, the point estimates are quite small in magnitude. For instance, the change in the share of white implied by the coefficient of column 1 is 0.7% with respect to the mean. Further, white people and children of higher socioeconomic status families and higher educated parents have, on average, higher longevity. Underrepresentation of these individuals in the final sample due to higher exposure to the program may add a downward bias, if anything, in our main results.

Appendix Table B-1 - Balancing Tests: Early-life Exposure to the Hookworm Eradication and Sociodemographic Composition of the Final Sample

	Outcomes:				
	White	Father Socioeconomic Index	Father Socioeconomic Index Missing	Mother Literate	Mother Literacy Missing
	(1)	(2)	(3)	(4)	(5)
Exposure × Treatment Share (STD)	-.00553**	-.31454**	.01175***	-.00428	.00578**
	(.00276)	(.13748)	(.00236)	(.00271)	(.00239)
Observations	891840	748316	891840	891840	891840
R-squared	.16709	.1077	.04997	.39036	.45918
Mean DV	0.779	23.707	0.168	0.585	0.355

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, and a county-specific linear trend in birth year. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.
*** p<0.01, ** p<0.05, * p<0.1

Appendix C

In this appendix, we examine the robustness of the results across alternative specifications. Appendix Table C-1 shows the estimates across more parsimonious specifications. In column 1, we include only county and year fixed effects. We add state-by-year fixed effects in column 2. Comparing the point estimate of column 2 versus column 1 reveals the relevance of accounting for other time-varying changes at the state level. In column 3, we add county trend. The estimated coefficient remains robust. In column 4, we include family and county covariates. The point estimate is quite comparable to that of column 3. Overall, the estimated coefficients reveal a robust pattern across these specifications.

In Appendix Table C-2, we start by replicating the results of column 4 of Appendix Table C-1. In column 2, we allow for county fixed effects to have differential impacts on longevity across different parental characteristics. The point estimate is almost identical to that of column 1. In column 3, we control for contemporaneous place-specific determinants of mortality by including the county of death fixed effects. We should note that this variable is available for a much smaller portion of the sample. We observe an increase of 2.4 months in longevity, an effect that is statistically significant.

In column 4, we add birth state by the 1940 state of residence fixed effects to absorb the influence of early adulthood cross-state migration in the results. The estimated coefficient is similar to the main results.

There is evidence of the influence of the season of birth and the season of death on mortality and longevity outcomes (Vaiserman, 2021). We account for these confounding influences by including birth month and death month fixed effects. The results, reported in column 5, are identical to those of column 1.

In columns 6-9, we examine the robustness to functional forms. In column 6, we replace the outcome with log age-at-death. In columns 7-8, we replace the outcome with indicators of longevity beyond ages 70, 75, and 80, respectively. We observe significant coefficients across all outcomes. In addition, the effects appear to be larger as we look at older ages. Comparing the effects with the mean of the outcomes, exposure is associated with 0.7%, 1.3%, and 1.9% higher longevity beyond ages 70, 75, and 80, respectively.

While in the main results of the paper, we cluster standard errors at county and year levels, we show the robustness of standard errors to county-level clustering in column 10. The estimated coefficient remains significant at 5% level.

Several studies document the influence of the infamous Spanish flu of 1918-19 on long-term outcomes (Almond, 2006; Almond & Mazumder, 2005; Fletcher, 2018c, 2018b, 2018a). In column 11, we remove these cohorts and replicate the main results. We observe an increase of roughly 15% in the magnitude of the coefficient compared with that of column 1.

To partially account for the influence of the Great Depression and the accompanying changes in the social policy environment under the New Deal relief programs and their long-run influences on health outcomes (Aizer et al., 2024; Noghanibehambari & Engelman, 2022), we remove cohorts of 1929 – 1940 in column 12. This sample restriction removes in utero exposure to the 1930s environment, but other individuals could be affected during childhood, adolescence, and early adulthood. However, the fact that the coefficient of column 12 remains comparable to that of column 1 suggests little concern about the adverse effects of the 1930s to confound our estimates.

Appendix Table C-1 - Early-life Exposure to the Hookworm Eradication and Later-life Longevity

	Outcome: Age-at-Death (Months)			
	(1)	(2)	(3)	(4)
Exposure × Treatment Share (STD)	1.12728**	1.26667***	1.30422***	1.25503***
	(.50652)	(.42696)	(.42718)	(.42828)
Observations	891840	891840	891840	891840
R-squared	.52209	.52261	.5232	.52321
Mean DV	860.862	860.862	860.862	860.862
Birth County FE	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓
Birth State by Birth Year FE		✓	✓	✓
County Trend			✓	✓
Controls				✓

Standard errors, clustered on birth county and birth year, are in parentheses. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.

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

Appendix Table C-2 - Examining Robustness Checks across Alternative Specifications

	Column 1 of Table 2	Interacting county fixed effects with parental covariates	Adding death county fixed effects	Interacting birth state by 1940 state	Adding birth month and death month fixed effects	Outcome: log age-at-death
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure × Treatment Share	1.25503***	1.25536**	2.36588**	1.22058***	1.28283***	.0014***
(STD)	(.42828)	(.57382)	(.99931)	(.4292)	(.42797)	(.0005)
Observations	891840	891840	208133	891239	891840	891840
R-squared	.52321	.52343	.64873	.52574	.52391	.50079
Mean DV	860.862	860.862	873.195	860.785	860.862	6.744
	Outcome: age-at-death > 70	Outcome: age-at-death > 75	Outcome: age-at-death > 80	Clustering standard errors on county	Removing cohorts of 1918-1919	Removing cohorts of 1929-1940
	(7)	(8)	(9)	(10)	(11)	(12)
Exposure × Treatment Share	.00365**	.00474**	.00436**	1.25503**	1.4435***	1.18412**
(STD)	(.00182)	(.00201)	(.00193)	(.52611)	(.43877)	(.47344)
Observations	891840	891840	891840	891840	841077	737959
R-squared	.39038	.39807	.36375	.52321	.53812	.38819
Mean DV	0.533	0.371	0.224	860.862	859.941	904.059

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.
*** p<0.01, ** p<0.05, * p<0.1

Appendix D

Figure 2 shows the point estimates across different age groups who were born in counties with a differential share of treatment. In this appendix, we group these cohorts to summarize those point estimates. Specifically, we create one group for exposure at age 0, another for exposure at ages 1-5, and another for exposure at ages 6-10. We use individuals with age at exposure of above 10 as the contrast group. The results are reported in Appendix Table D-1. Similar to the pattern of Figure 2, we observe quite small and insignificant coefficients for childhood ages. The point estimate for age at exposure of [≤ 0] is similar to the main results of the paper.

Appendix Table D-1 - Examining the Effects across Different Ages at Exposure

	Outcome: Age-at-Death (Months)		
	(1)	(2)	(3)
Age at Exposure ≤ 0 × Treatment Share (STD)	1.25353***	1.27019**	1.50587**
	(.42828)	(.5169)	(.61077)
Age at exposure [1-5] × Treatment Share (STD)		.02515	.22205
		(.42384)	(.48187)
Age at exposure [6-10] × Treatment Share (STD)			.35407
			(.43843)
Observations	891840	891840	891840
R-squared	.52321	.52321	.52321
Mean DV	860.862	860.862	860.862

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.
*** p<0.01, ** p<0.05, * p<0.1

Appendix E

The DMF data contains deaths of male individuals only. To examine the impacts on female individuals, we employ Numident data extracted from the Censoc project (Goldstein et al., 2021). The disadvantage of the Numident data is that it covers a much narrower death window, 1988-2005. However, it contains deaths of both genders. We implement sample selections similar to the DMF data to examine the effects among females in the Numident data. These results are reported in Appendix Table E-1. In column 1, we report the results of the full sample of the DMF data. In column 2, we restrict the DMF data to death years 1988-2005. In column 3, we use male individuals of the Numident data (death years 1988-2005). We observe a smaller effect among male individuals of the Numident data than male individuals of the DMF data, although both coefficients are insignificant. We should note that the mean age-at-death in the Numident data is considerably lower than that of DMF data, which makes comparisons of columns 2 and 3 challenging. However, when we focus on female individuals in the Numident data, we observe a point estimate of 0.86. This estimate is statistically significant and larger than the male individuals in the Numident data. Cross-column and cross-data comparisons of this table provide suggestive evidence that the effects on females in the DMF death years of 1975-2005 could be approximately 2.6 months ($= 1.25 * \frac{0.86}{0.41}$).

Appendix Table E-1 - Replicating the Main Results Using Numident Data

	Outcome: Age-at-Death (Months)			
	DMF, Males, Death Years 1975-2005	DMF, Males, Death Years 1988-2005	Numident, Males, Death Years 1988-2005	Numident, Females, Death Years 1988-2005
	(1)	(2)	(3)	(4)
Exposure × Treatment Share (STD)	1.25503***	.56067	.40943	.86237***
	(.42828)	(.34199)	(.47217)	(.30189)
Observations	891840	506268	612336	699324
R-squared	.52321	.75649	.67446	.75525
Mean DV	860.862	891.711	865.840	916.201

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.

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

Appendix F

In this appendix, we examine the role of the pre-eradication prevalence of hookworm on the effects of treatments on longevity. To ease the interpretations, we use a dummy variable indicating treatment share above the median instead of a standardized value of the treatment share. Column 1 of Appendix Table F-1 suggests an identical coefficient size to the main results when we use this dummy variable. However, the estimated effect is statistically insignificant. In column 2, we interact this dummy variable with an indicator of pre-eradication hookworm rate being above the sample median. The point estimate suggests an increase in longevity of about 1.4 months. In column 3, we also include another indicator for pre-eradication hookworm rate below the median. Comparing the point estimates of interaction terms in this specification implies that the effects of treatment are concentrated in areas with higher pre-eradication hookworm prevalence.

Appendix Table F-1 - Exploring the Effects of Treatments Based on Pre-Eradication Hookworm Rate

	Outcome: Age-at-Death (Months)		
	(1)	(2)	(3)
Age at Exposure ≤ 0 × Treatment Share Above-Median	1.30485 (.82241)		
Age at Exposure ≤ 0 × Treatment Share Above-Median × Pre-Eradication Hookworm Rate Above-Median		1.40801* (.83073)	1.52548* (.87112)
Above-Median × Pre-Eradication Hookworm Rate Below-Median			.66556 (1.35418)
Observations	891840	891840	891840
R-squared	.5232	.5232	.5232
Mean DV	860.862	860.862	860.862

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.
*** p<0.01, ** p<0.05, * p<0.1

Appendix G

In this appendix, we include all other counties in the final sample and assign a treatment value of zero to individuals born in these counties. These results are reported in column 1 of Appendix Table G-1. We avoid including state-by-year fixed effects to allow cross-county comparisons across all states. We observe an increase in longevity of about 3.2 months. In column 2, we focus only on counties in the states that are present in the final sample of the paper. In this specification, similar to the main results of the paper, we include state-by-year fixed effects. We observe an almost identical coefficient to the main results of column 1 of Table 2.

Appendix Table G-1 - Exploring the Effects Including Other Counties Assuming Zero Treatment

	Outcome: Age-at-Death (Months)	
	All US Counties	Counties in the States of the Final Sample
	(1)	(2)
Age at Exposure ≤ 0 × Treatment Share	3.16443***	1.34495***
(STD)	(.73765)	(.48126)
Observations	7679358	1671323
R-squared	.49961	.51333
Mean DV	885.524	861.524

Standard errors, clustered on birth county and birth year, are in parentheses. Regressions include birth county fixed effects, birth year fixed effects interacted with birth state fixed effects, county-specific linear trend in birth year, and covariates. Controls include dummies for individual race, dummies for paternal socioeconomic index and maternal literacy, county level share of homeownership, share of under five children, share of immigrants, share of literate people, share of employed, share of employed in manufacturing sector, share of employed in construction sector, share of employed in transportation sector, share of employed in farming sector, share of married, exposure to Boll Weevil, and average occupational income score. Regressions are weighted using mean county level population. The data covers male individuals born between 1880-1940 and died between 1975-2005.

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