Early-life infectious disease exposure, the "hygiene hypothesis", and lifespan: evidence from hookworm *

Ralph I. Lawton

Harvard University

ralphlawton@hms.harvard.edu

December 9, 2024

Abstract

Exposure to infectious disease in early life may have long-term ramifications for health and mortality. This study leverages quasi-experimental variation from the Rockefeller Sanitary Commission's de-worming campaign in the early 20th century, combined with pre-campaign hookworm prevalence, to rigorously examine the impacts of childhood hookworm exposure on later-life morbidity and lifespan. Pre-intervention surveys find widespread hookworm exposure among children in the American South, but minimal prevalence among adults. I show exposure to de-worming before age five leads to 2.5 additional months of life in a sample of older-age mortality. Further, decreasing hookworm exposure is related to later-life declines in biomarkers for inflammation and skin-tested allergies, in contrast with the predictions of the "hygiene hypothesis". Placebo tests using health outcomes that should not be affected by de-worming do not show similar patterns.

*Acknowledgements: Many thanks to David Cutler, Ellen Meara, Lawrence Katz, Duncan Thomas, Claudia Goldin and Kirti Nath for helpful comments and discussion in the development of this draft. This work was supported by the National Institute of Aging (T32AG51108) and the National Institute of General Medical Sciences (T32GM144273).

1 Introduction

Over 400 million people still live with hookworm (Hotez et al., 2014). Seminal work finds human capital and short-term health benefits from mass de-worming (Bleakley, 2007; Miguel and Kremer, 2004). However, the value of large-scale de-worming continues to be debated (Croke et al., 2024; Taylor-Robinson et al., 2019). This study contributes direct evidence on the adult life expectancy and morbidity consequences of childhood de-worming.

I also use de-worming to test the pathways linking early life infectious disease exposure and adult life expectancy. While in-utero shocks have well-known consequences, relatively little is known about the long-term mortality and morbidity effects of childhood exposures (Almond and Currie, 2011). Reductions in early-life infectious diseases have been implicated in 20th-century declines in older-age mortality (Crimmins and Finch, 2006; Cutler and Meara, 2001; Cutler and Miller, 2005; Finch and Crimmins, 2004). However, causal evidence and mechanistic understanding remain sparse. This is in part because it is difficult to separate in-utero maternal health shocks from changes in child health, and because childhood infectious disease exposure may be endogenous to other socio-economic forces (Almond and Currie, 2011; Beach et al., 2022; Bengtsson and Lindström, 2003; Cohen et al., 2010; Mazumder et al., 2010; Myrskylä et al., 2013). [1] At the same time, the "hygiene hypothesis" suggests removing pathogens may not be uniformly beneficial (Strachan, 1989). Helminth reductions to the rise of allergies and auto-immune conditions, but with contested theoretical mechanisms and little long-term human evidence (Briggs et al., 2016; Loukas et al., 2016; Wammes et al., 2014).

This study's approach leverages a natural experiment to provide rigorous evidence on the role of childhood hookworm in long-term outcomes. Hookworm in the early 20th century United States was nearly-entirely a disease among children (Commission, 1914; Loukas et al., 2016). Hookworm also caused very little mortality, but potentially serious chronic morbidity: a substantial change in hookworm prevalence thus could provide compelling evidence on the impact of chronic inflammatory processes, and rule out potential in-utero mechanisms and adult exposures. In 1909, the Rockefeller Sanitary Commission was established, and began a mass de-worming campaign for children in the American southeast. While incorporated in 1909, significant de-worming did not occur until 1913 (Elman et al., 2014). Prior to 1910, 40% of children had hookworm infection in the American southeast, with substantial variation within and across states (Ettling, 2013). The distribution, from Rockefeller Archive data digitized by Roodman, 2018 is shown in figure 1. Areas with greater hookworm

[1] Another potential concern is that mortality selection in childhood may impact adult outcomes - studies from diseases with high mortality rates may in part be picking up changes in mortality selection over time (Caruso et al., 2005).

prevalence stood to benefit more from de-worming (Bleakley, 2007). Combined with rapid, unanticipated changes in hookworm prevalence due to the RSC campaign, this provides a rigorous setting to explore the long-term health effects of hookworm.

To identify the later-life impacts of hookworm on mortality and morbidity, I use an approach analogous to difference-in-differences that compares adults born in areas with more and less hookworm exposure, before and after the RSC campaign. The estimate captures the implied effect of each year of hookworm exposure. I examine the impacts of de-worming on lifespan in a sample of older adult deaths drawn from social security records. Importantly, I am able to examine variation across counties, and account for time-specific changes in each state. In a parallel set of analyses, I draw morbidity outcomes from nationally-representative samples that include biomarkers for immunologic processes such as erythrocyte sedimentation rate and allergy skin tests, other hookworm-related outcomes including BMI and hemoglobin, and a set of placebo health outcomes that should be unrelated to hookworm exposure.

2 Results

2.1 Study population

This study draws data from two populations. The primary mortality analyses draw data from the Berkeley Unified Numident Mortality Database (BUNMD) (Breen and Goldstein, 2022; Goldstein et al., 2023). These data clean and assemble individual-level death records, primarily from a 2013 transfer of Numident records from the Social Security Administration to the National Archives and Records Administration. Critically, these data include county of birth, as well as age and date of death, which form the foundation of my primary analyses. While the coverage is nearly universal among those dying over age 65 between 1988 and 2005, death records are recorded throughout the 20th century, and coverage increases significantly after 1975. I restrict to individuals born in areas of the American South where the RSC was active and where I have SEA-level hookworm prevalence data, and focus on those born between 1900-1935. This yields an analytic sample of 3,980,291 individuals, 47% of whom are male, and 27% are Black (table 1). The average lifespan in this sample is 77.6, older than expected for these birth cohorts, since the BUNMD over-represents deaths at the later half of the 20th century.[2] While it is unlikely that representation in the sample is related directly to the RSC campaign besides through selected mortality, especially since the birth cohorts represented in the BUNMD center on my cohorts of interest, it is important to consider that estimates are in a sample of older-age deaths. Selection into the sample on the basis of

[2]Given that men tend to live less long than women, minimal differences in our sample representation by sex suggest that endogenous selection on older-age mortality may not be particularly concerning in practice.

older ages of death would likely bias estimates toward 0. Nonetheless, I conduct additional analyses with census data to validate results.

My second sample draws from two waves of the National Health and Nutrition Examination Survey (NHANES), which were fielded in 1971-74, and 1976-80. Both surveys are nationally-representative probability samples of the United States, and both surveys include information about state of birth. Similar to my primary specification, I focus on those born from 1900-1935, for a sample of 18,377. Unlike the primary specification, I use state of birth to assign hookworm prevalence, and thus will make comparisons across states. I only restrict to individuals born in the contiguous United States. These morbidity outcomes are assessed in mid to late life, with an average age of 57.18. 46% are male, 15% are Black, 39% live in rural areas, and 47% finished high school. Results should be interpreted as morbidity in mid-to-late life. I test several biomarkers from this population. I will examine erythrocyte sedimentation rate (ESR), a biomarker for background inflammation; number of skin-tested allergies out of 8 skin pricks; body-mass-index, hemoglobin, and a series of placebo outcomes. Not every outcome was assessed in every wave or in all individuals - many were only evaluated in random sub-samples of the population.

2.2 Impacts of de-worming on mortality

I evaluate the impacts of de-worming before age 5 on lifespan in the BUNMD sample. I find one year of de-worming is linked to 0.042 additional life years (0.504 months), in a sample of older adult deaths from RSC-participant states (Table 2). This effect is qualitatively large - the full implied impact of 5 years of exposure is 0.21 additional years of life at older age, similar to the life expectancy gained by eliminating Alzheimers' disease, half of cerebrovascular disease in older age, or removing lead in-utero (Arias et al., 2013; Fletcher and Noghanibehambari, 2023).

The point estimates for females are larger than for males, though the difference between the two is not statistically significant. This is potentially consistent with the fact that women have conventionally benefited more than men from improvements in the infectious disease environment (Goldin and Lleras-Muney, 2019).

Two parallel event-study style analyses support the argument that changes in lifespan are driven by de-worming due to the RSC-campaign. In the left panel of figure 2, an event-study examines the change in the estimated relationship between baseline hookworm prevalence and longevity relative to the last birth years before the intervention began. In the 12 preceding years, there is no evidence that the outcomes for areas with more hookworm prevalence were evolving differently from those with less, supporting the validity of the parallel-trends assumption. After the RSC campaign, the coefficient changes positively, suggesting increases

in lifespan related to deworming that occur at the time of the campaign. On the right panel, a similar analysis is conducted, but the SEA fixed effects are excluded, so the hookworm prevalence-lifespan relationship can be estimated specifically within each birth cohort. A similar pattern is visible - there are no systematic trends related to hookworm prevalence in the pre-period, but in the post period the coefficient goes from negative to essentially 0, suggesting that the relationship between baseline hookworm prevalence and lifespan did not just become more positive, but in fact went from negative to zero in the birth cohorts after the campaign occurs.

Results are robust to a variety of alternative specifications and functional forms, including linear years of exposure before age 5, a binary pre/post specification, and a semi-parametric specification using dummies for tertiles of baseline hookworm prevalence (supplement table A1).

Another potential concern is that these data do not capture mortality that occurs earlier in life. While this is likely to make the results shown underestimates, I conduct parallel analyses of survival until 1960 using state data from the census, focusing on the same RSC-campaign states. I find similar findings to the lifespan analysis with finer-grained hookworm prevalence - de-worming significantly increases the probability of survival until 1960 (A2).

2.3 Impact of childhood de-worming on morbidity

Using data from the NHANES I and NHANES II with information on state of birth, I estimate a parallel model to the specification for longevity, using state hookworm prevalence at baseline in the contiguous United States. I focus on domains in which childhood exposure to hookworm may be persistent: health outcomes related to immunologic changes, and health outcomes related to nutritional status and anemia.

I find evidence for long-term immunologic changes related to de-worming (table 3. Estimating the effect of 1 year of de-worming on erythrocyte sedimentation rate, a biomarker for inflammation, finds each year is related to a reduction of 1.5 later in life, which is significant at a 10% size of test. This implies a reduction by 7.5 with full campaign exposure (approximately 3/4 of a standard deviation), which is slightly larger than the gap between people born before the campaign in the South versus the rest of the country pre-intervention. I further find significant reductions in the number of skim-tested allergies, particularly among women. The implied reduction in the number of allergies is substantial: five years of childhood de-worming exposure is estimated to reduce the expected number of allergies by nearly 0.75.

Beyond their immunologic impacts, hookworm also may cause decreases in BMI and hemoglobin, which may be persistent. I find evidence for long-term improvements. Exposure

to de-worming is linked to increased BMI and hemoglobin in adulthood. This appears to be true for both males and females. However, the benefits appear to be sub-clinical improvements in hemoglobin status. While the point estimates for anemia are negative, they are small and not statistically significant.

In the NHANES data, unlike the lifespan data, I can control for individual-specific socio-economic and geographic characteristics at the time of measurement. I include these in the estimates shown, but estimates are minimally-affected when a simpler specification without controls is used.

2.4 Placebo health tests

I also examine a series of placebo health outcomes from the NHANES. I first examine height. Height is not a perfect placebo outcome, as extreme calorie deprivation can reduce heights. However, most hookworm infections were not extreme, and height also reflects in-utero health, early life conditions, and characteristics of parents that may be changing in the background (Beach et al., 2022; Martorell, 1995; Strauss and Thomas, 1998). While point estimates are small and positive, I find no significant effect of exposure to the de-worming campaign and height. This is particularly re-assuring insofar as it indicates little in the way of in-utero changes or other early-life impacts.

I also examine cholesterol measures, which reflect not only health endowments but also health-related behaviors such as diet, exercise, and health care access. I find no impacts of de-worming on cholesterol measures.

Finally, I find no effects of de-worming on FEV1/FVC, a common pulmonary function test that is sensitive for asthma.

3 Discussion

I examine the impact of de-worming before the age of 5 on lifespan at older age, as well as a series of morbidity outcomes measured in midlife. In doing so, I provide direct evidence on the long-term impacts of de-worming, as well as novel tests of 1) the linkage between early-life inflammatory exposure with later life morbidity and mortality, and 2) the role of hookworm in the "hygiene hypothesis" in the 20th century. I find substantial gains in life expectancy at older age, as well as long-term reductions in a biomarker of inflammation and the number of skin-tested allergies, as well as increases in BMI and hemoglobin. Similar changes are not observed in health outcomes that should be unaffected by de-worming.

The estimated impact of de-worming on life expectancy at older age is large. These findings are consistent with, and provide new rigorous evidence for, influential theories that infectious disease burdens early in life shape mortality across the life course (Finch, 2010;

Finch and Crimmins, 2004). It is notable that both men and women benefit similarly from de-worming, though point estimates are larger for women. Prior work has suggested that males and females may have different returns to changes in the infectious disease environment (Goldin and Lleras-Muney, 2019; Noymer and Garenne, 2000). The morbidity findings provide mechanistic support for many of the theorized links between early life infectious disease and later life mortality. Changes in ESR are consistent with long-term changes in chronic, low-grade, inflammation that may raise cardiovascular disease and cancer risk (Fulop et al., 2018; Gurven et al., 2008). Further, disease-changes in early-life health such as nutritional status and hemoglobin may be persistent, leading to long-term benefit (Blackwell et al., 2011; Crimmins and Finch, 2006). In this sample of older adults with low BMI in the 1970s, I interpret increases in BMI as likely positive protections against frailty (Kıskaç et al., 2022; Kopinska et al., 2021).

While the improvements in mortality can be attributed to meaningful observed improvements in underlying immunologic and nutrition-related changes from early childhood, it is possible that de-worming may lead to economic benefits that may also play a role (Bleakley and Lange, 2009; Miguel and Kremer, 2004). While I do not have educational attainment and mortality data in the same individuals, it is unlikely that changes in education or income fully explain the observed impacts. In a study of the economic ramifications of under-18 exposure to de-worming, Bleakley finds no long-term impact on educational attainment, but small increases in income. In my analyses of morbidity, I directly control for a variety of potentially relevant socio-economic pathways including education, poverty, and rurality. These controls minimally impact observed results.

The joint reductions in background inflammation as well as allergies further an argument that helminths may in fact raise allergies through chronic inflammatory processes (Briggs et al., 2016). Notably, these findings contrast sharply with a potential role of hookworm in the "hygiene hypothesis" (Briggs et al., 2016; Loukas et al., 2016; Ramanan et al., 2016; Strachan, 1989; Wammes et al., 2014; Yazdanbakhsh et al., 2002). This study provides rigorous evidence of long-term impacts between hookworm and allergies, in a space where mouse models and short-term human trials are conflicting, and active debate ensues. However, the interpretation of these findings is limited by the context of the RSC de-worming campaign in the early 20th century. The role of hookworm in auto-immune processes may vary in contexts with differing levels of background pathogens.

This study also has substantial implications for mass de-worming programs around the globe. While some data from randomized trials with short and medium-term follow-up into early adulthood and adolescence finds persistent health and economic gains, the consequences of de-worming programs continue to be debated (Croke and Atun, 2019; Croke et al.,

2024; Hamory et al., 2021; Miguel and Kremer, 2004; Taylor-Robinson et al., 2019). Other work from the Rockefeller Sanitary Commission's de-worming program find contemporaneous benefits to school attendance in childhood, with small benefits in adulthood (Bleakley, 2007). Work by Noghanibehambari and Fletcher takes a similar approach to this study, but focuses only on men, examines broader education impacts across childhood and adolescence rather than focusing on a critical period, and does not investigate morbidity changes (Noghanibehambari and Fletcher, 2024). This paper includes women in the primary sample, using data with a wider coverage of deaths and a sample twice as large for men, as well as testing the physiology of critical-period exposure before age 5 in the long term. I provide evidence for life-long benefits to de-worming, with persistent impacts on health and life expectancy, making a strong case for expanded de-worming programs in endemic areas.

This study has several limitations. One limitation is that the primary source of mortality data primarily draws from the second half of the 20th century, potentially missing earlier-life mortality. Hookworm should have minimal impact on childhood mortality. Analyses at the state level using census data on survival rates until 1960, which should reflect mid-life mortality, find similar large impacts of hookworm exposure to the older life samples. Additionally, individuals with "marginal" health dying before observation at older age means estimates at older age likely under-state results. For both empirical and theoretical reasons, the primary estimates are thus likely to be underestimates of true effects.

Another potential limitation is unobserved changes in the health status across cohorts before and after de-worming begins, in a way that is related to hookworm prevalence. I address this in several ways. 1) Using a model that includes state-by-birth year fixed effects in the primary mortality specification, I am able to account for potential state-specific trends and changes over time, and localize the variation used to across-SEAs within states. 2) Event-study analyses suggest that in the years before the campaign, and the years after full effects are realized, mortality dynamics are not evolving in a way related to baseline hookworm prevalence. 3) I provide evidence from placebo health outcomes that there were not major changes in health status unrelated to hookworm-specific morbidity. Finally, interpretation is limited by external validity - the de-worming campaign studied occurred in an environment that ultimately faces very different ecological forces than present day.

I conclude that de-worming in early childhood has large life-long benefits for morbidity and mortality. Mechanistic evidence finds support for long-term immunologic changes due to de-worming, and other health benefits. In this context, hookworm reduces allergies in a way that is not consistent with a role in the "hygiene hypothesis". Potential long-term benefits should be considered when evaluating mass de-worming programs.

4 Methods

The intuition for these analyses is analogous to a difference-in-differences approach, similar to Bleakley, 2007. However, while conventional difference-in-differences uses a treatment variable that takes the value of 1 for groups eventually treated, and 0 untreated, I use continuous hookworm prevalence from prior to the intervention that takes values between 0 and 1. Individuals from areas with higher hookworm prevalence should disproportionately benefit from the de-worming campaign. Comparing birth cohorts before and after the de-worming campaign from areas of varying hookworm prevalence enables an implied estimate of the long-term consequences of childhood de-worming. I first describe the data sources for hookworm prevalence, and then the specific analyses I conduct for lifespan and morbidity.

4.1 Data sources for hookworm prevalence

Prior to the initiation of the de-worming campaign, the RSC sampled at least 200, but often many more, children in each county. They examined stool microscopically for hookworm ova. Using data digitized by Roodman from the Rockefeller foundation archives, I aggregate counties into State Economic Areas with stable borders over time (Commission, 1914; Roodman, 2018). As shown in figure 1, prevalence varied wildly pre-intervention within and across states. Hookworm suitability was primarily a function of land characteristics and temperature: sandy soils were preferable for hookworm reproduction. Note that many of the largest urban centers were excluded from the pre-intervention prevalence data and activity. People born in these areas are also excluded from analysis. Prior analyses of pre-intervention hookworm prevalence found more rural and less Black counties had higher prevalence, but the first-order variation was predicted by soil type and number of frost-free days (Elman et al., 2014). I use this pre-intervention hookworm prevalence data in my primary mortality specification.

In other analyses, I take a complementary approach comparing hookworm prevalence across states. In order to expand my sample to the contiguous United States, I reference complementary hookworm data from army recruits (Kofoid and Tucker, 1921). Analyses of the 11 overlapping states suggests that the supplementary dataset approximates well systematic county-level data collection from the RSC (supplement figure A1). The primary difference is that army data is lower, on average, than the county data for a given prevalence level. This is likely due to differences in the age of the samples. Kofoid notes that the burdens identified were the remaining light-intensity residual hookworm burden as hookworms in young soldiers died and they were not re-infected.

4.2 Defining exposure to the RSC campaign

In order to account for 1) partial treatment of children born right before the initiation of the RSC campaign, and 2) the fact that it took several years to achieve high penetrance of de-worming, I construct a continuous measure of "effective years" of exposure to the campaign before age 5. In order to model the increasing efficacy of the campaign over time, I use aggregated data from a subset of counties where repeated follow-ups occurred. In these samples, in 1910 59.7% of children had hookworm, falling to 39.7 in 1915, to 21.7 by 1920. This suggests yearly reduction in hookworm prevalence by a factor of 0.9. The functional form is as follows: $\sum_{y=Y1}^{Y5} \mathbb{1}[y \geq 1912](1 - 0.90^{(y-1912)})$ where $Y1, Y5$ are the years of birth through age 5 respectively. Thus, each individual gets the sum of "effective years" of exposure to the campaign before they reach the age of 5. The functional form is plotted in supplement figure A2. Analogous to more simpler approaches that use "years of exposure", in the pre-period individuals have 0 effective years, and by 1920 the functional form is essentially flat (though increasing slightly). Results are robust to using simpler functional forms.

4.3 Primary analyses of mortality

The empirical strategy follows closely from intuition developed in Bleakley (2007). The strategy rests on two core ideas. First, it leverages the varying hookworm prevalence across different areas at baseline. While all areas in the RSC treatment states were able to be exposed to the benefits of de-worming, given higher prevalence in some areas, some areas stood much more to benefit from the RSC treatment than others. Second, the study design treats the initiation of the RSC campaign as an exogenous natural shock - given poor baseline understanding of hookworm and the amount of capital investment needed to intervene, the timing and initiation of the RSC campaign is considered exogenous to later health outcomes.

The core specification for the lifespan outcome is as follows:

$$Y_{icst} = \beta(W_c * EffectiveYrsExposed_t) + \theta_{st} + \gamma_c + \epsilon_{ict} \tag{1}$$

Where Y_{icst} is the outcome of interest for an individual i born in state economic area c, and state s, in year t. W_c refers to the hookworm prevalence at baseline in SEAs indexed by c, and $EffectiveYearsExposed$ is based on birth year and is the effective number of years exposed to the RSC campaign. θ_{st} represents the state of birth-by-year fixed effects, which capture state-specific shocks and time-trends. X_i represents covariates of interest - while a general label is used, in lifespan analyses the only covariate used is sex. γ_c are fixed effects for SEA c (functionally an aggregation of several counties). These terms are allowed to vary by sex as potentially important sex-specific trends may have been different for males and

females during this time (Bhalotra et al., 2023; Miller, 2008). In regressions separated by sex, these effects collapse to just one term for birth county, state, or year.

Note, γ_c absorbs the baseline worm prevalence, along with observed and unobserved SEA-level characteristics that may be relevant for lifespan and are not changing over time. While there are some concerns for this project about the full accounting of all deaths, which may be disproportionately skewed in these data towards higher-coverage death years of the BUNMD, the birth state by birth year fixed-effect capture cohort-level endogeneity in the probability a given individual is reflected in the data, to the extent it is fixed across a given cohort. θ_{st} also accounts for observed and unobserved state and cohort effects in lifespan. The analytical sample is restricted to RSC-treated SEAs.

Taken together, the coefficient of interest β can be interpreted as the marginal increase of experiencing a full effective year (maximum 5 total) of de-worming on longevity, estimated by comparing across SEAs, but within states, that have differing baseline prevalence of hookworm. Standard errors are clustered at the SEA level.

4.4 State-level survival estimates

In order to address concerns that the primary mortality data are at older ages and my miss earlier-life mortality, I conduct a second series of analyses using survival rates I construct from census data. I detail the approach in the supplement. In brief, I construct survival rates until 1960 from the census for birth cohorts from 1900-25, where the numerator is population in 1960 and the denominator the population alive at ages 5-14, whenever I first observe the cohort after infant mortality has stabilized. Using these state level survival rates, I estimate a similar approach with state-level hookworm prevalence, and a state fixed effect instead of SEA.

4.5 Morbidity analyses

I use a parallel strategy to the primary analyses of lifespan, in order to evaluate the impacts of de-worming on morbidity. I use a series of biomarkers from the NHANES. Since the NHANES does not have county or SEA of birth, I construct treatments at the state of birth, using state level hookworm prevalence from Kofoid and Tucker, 1921.

The core specification for each morbidity outcome is as follows:

$$Y_{ist} = \beta(W_s * EffectiveYrsExposed_t) + \lambda_t + \gamma_s + \phi X_i + \epsilon_{ist} \tag{2}$$

Where Y_{ist} is the outcome of interest for an individual i born in state s in year t. W_c refers to the hookworm prevalence at baseline in states indexed by s, and $EffectiveYearsExposed$ is based on birth year and is the effective number of years exposed to the RSC campaign.

λ_t, γ_s are birth state and birth-year fixed effects, which capture time-invariant state characteristics, and cohort-specific effects. These terms are allowed to vary by sex. One strength of the NHANES is richer socio-economic data. I include X_i as a vector of covariates of interest, including sex, NHANES wave, rural/urban status, Black race, educational attainment, and low income level. It is worth noting that γ_s will absorb baseline state hookworm prevalence, along with observed and unobserved state-level characteristics. λ_t will absorb year-specific effects, including age. β retains a similar interpretation to the primary specification, as the marginal increase of experiencing a full effective year (maximum 5 total) of de-worming on longevity, estimated by comparing across states that have differing baseline prevalence of hookworm, before and after the de-worming campaign. Standard errors are clustered at the state level.

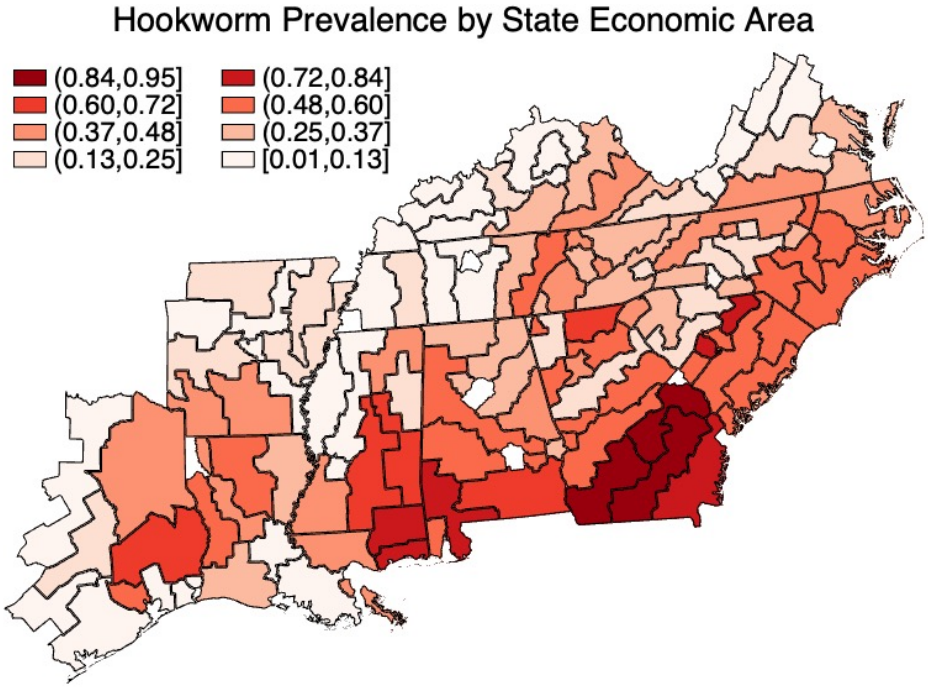

Figure 1: Distribution of pre-intervention hookworm by state economic area

Notes: This figure shows estimates of childhood hookworm prevalence from systematic county surveys of children pre-intervention. County data are aggregated and displayed at the State Economic Area (SEA).

Table 1: Summary statistics by sample

	BUNMD	**NHANES**
Lifespan (years)	77.6	-
	(7.2)	-
Age (years)	-	57.18
	-	(10.69)
Proportion Male	0.47	0.46
	(0.50)	(0.50)
Proportion Black	0.27	0.15
	(0.45)	(0.35)
Effective Years	2.92	2.34
	(1.34)	(1.85)
Proportion RSC State	1	0.34
	-	(0.47)
Proportion in Major Urban Areas	-	0.46
	-	(0.50)
Proportion in Rural areas	-	0.39
	-	(0.49)
Proportion HS Grad+	-	0.47
	-	(0.50)
Proportion College Grad+	-	0.08
	-	(0.28)
Proportion low income	-	0.24
	-	(0.43)
N	3,980,291	18,377

Notes: This table presents summary statistics for individuals in both of the analytic samples used for these analyses. Mortality data are drawn from the Berkeley Unified Numident Mortality Dataset. Morbidity data are drawn from the NHANES I and NHANES II.

Table 2: Impacts of childhood de-worming on lifespan

	Longevity (years)		
	All	**Males**	**Females**
A. Primary specification			
SEA Prevalence X Effective Exposure Years	0.042	0.034	0.046
	[0.012]	[0.015]	[0.016]
Constant	77.765	76.100	79.191
	[0.012]	[0.016]	[0.015]
N	3,980,291	1,851,939	2,128,352
B. Allow for SEA-specific mean reversion			
SEA Prevalence X Effective Exposure Years	0.042	0.035	0.045
	[0.012]	[0.015]	[0.017]

Notes: This table presents estimates of the effect of exposure to the de-worming campaign on the age of death, first pooled then stratified by sex. Data are drawn from the CenSoc BUNMD. Hookworm prevalence is assigned at the SEA of birth. Estimates are conducted using fixed effects by sex that reflect birth SEA, and birth state-by-birth year. Standard errors presented in brackets, clustered at the SEA level. The coefficients presented can be interpreted as the implied effects of one full year of exposure to de-worming, and each individual received up to five years of exposure.

Table 3: Impacts of exposure to de-worming on later-life allergy and inflammation

	ESR			# of Allergies		
Immunologic changes	All	M	F	All	M	F
Prev. X Eff. Exp. Yrs	-1.502	-2.314	-1.039	-0.144	-0.089	-0.184
	[0.871]	[1.187]	[1.017]	[0.068]	[0.106]	[0.100]
N	6,218	2,660	3,558	8,639	4,219	4,420

Notes: This table presents estimates of the effect of exposure to the de-worming campaign, first pooled then stratified by sex. Prevalence is assigned by state of birth. Data are drawn from the NHANES I and NHANES II. Models include fixed effects for birth year and birth state, controls for sex, age, rural/urban, status, Black race, educational attainment, and low income level. Standard errors presented in brackets, clustered at the State level.

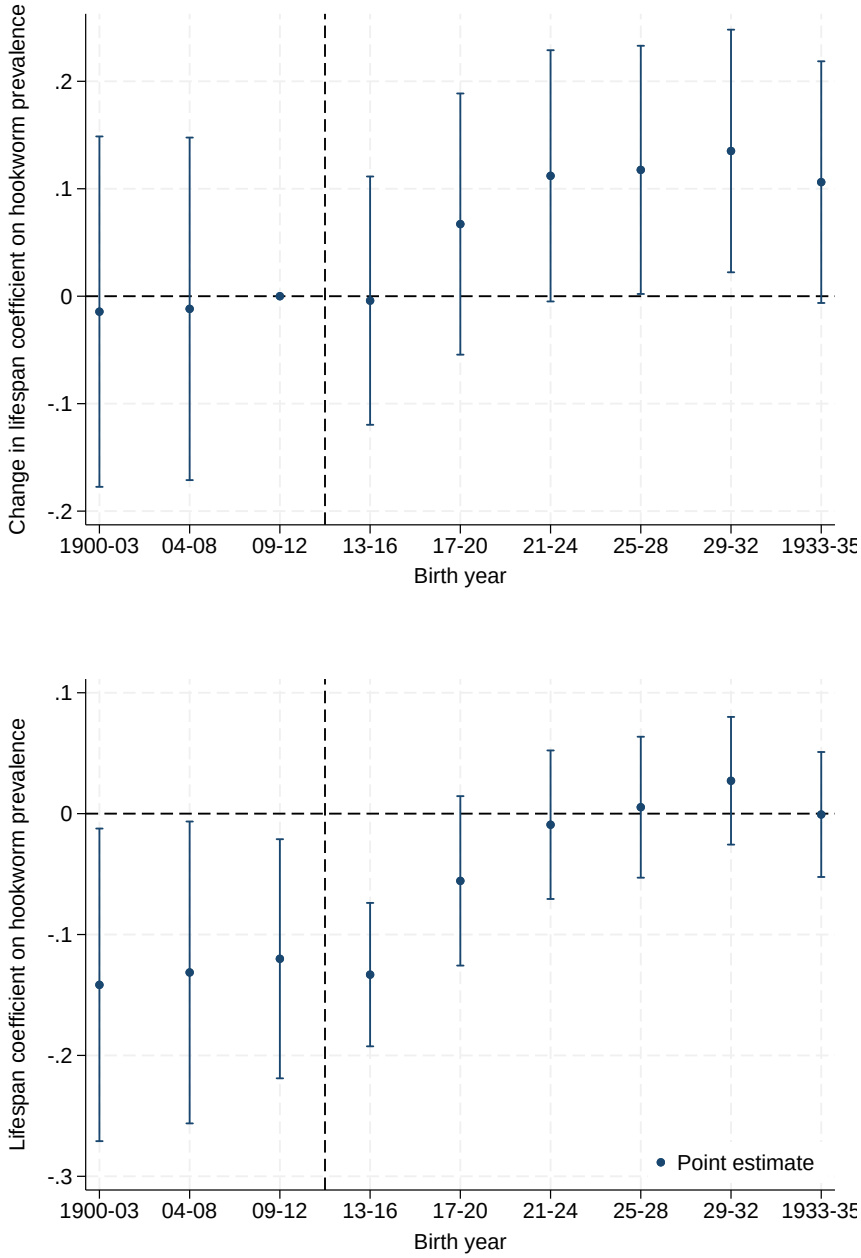

Figure 2: Relationship between baseline hookworm prevalence and lifespan over time

Notes: This figure shows estimates of the relationship between *baseline* state economic area (SEA) hookworm prevalence and longevity for people born in a given SEA, estimated within each birth year cohort and with state fixed-effects. Only SEAs where the RSC was active are included. These coefficient estimates and 95% confidence intervals are plotted within each cohort over time, to visualize the trajectory of the changing relationship between baseline hookworm prevalence and longevity around the time of the initiation of the RSC campaign. The Rockefeller Campaign's first full year of effect was 1913.

Table 4: Impacts of exposure to de-worming on later life BMI and hemoglobin

	BMI			Hemoglobin			Anemia		
Variables	All	M	F	All	M	F	All	M	F
Prev. X Eff. Exp. Yrs	0.923	1.050	0.670	0.143	0.151	0.128	-0.012	-0.019	-0.008
	[0.273]	[0.291]	[0.327]	[0.058]	[0.097]	[0.046]	[0.010]	[0.017]	[0.012]
N	17,450	8,162	9,288	17,068	7,856	9,212	17,068	7,856	9,212

Notes: This table presents estimates of the effect of exposure to the de-worming campaign, first pooled then stratified by sex. Prevalence is assigned by state of birth. Data are drawn from the NHANES I and NHANES II. Models include fixed effects for birth year and birth state, controls for sex, age, rural/urban, status, Black race, educational attainment, and low income level. Standard errors presented in brackets, clustered at the State level.

Table 5: Placebo tests

Variables	Height (m)	Total Cholesterol	HDL	FEV1/FVC
Prev. X Eff. Exp. Yrs	0.002	-1.395	-1.895	-0.009
	[0.006]	[1.431]	[1.301]	[0.007]
N	17,714	15,526	5,695	3,338

Notes: This table presents estimates of the effect of exposure to the de-worming campaign. Prevalence is assigned by state of birth. Data are drawn from the NHANES I and NHANES II. Models include fixed effects for birth year and birth state, controls for sex, age, rural/urban, status, Black race, educational attainment, and low income level. Standard errors presented in brackets, clustered at the State level.

References

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

Arias, E., Heron, M., & Tejada-Vera, B. (2013). United States life tables eliminating certain causes of death, 1999-2001. *National Vital Statistics Reports*, *61*(9). Retrieved April 22, 2024, from https://stacks.cdc.gov/view/cdc/13714

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). https://doi.org/10.1086/719757

Bengtsson, T., & Lindström, M. (2003). Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766–1894. *International Journal of Epidemiology*, *32*(2), 286–294. https://doi.org/10.1093/ije/dyg061

Bhalotra, S., Clarke, D., Gomes, J. F., & Venkataramani, A. (2023). Maternal Mortality and Women's Political Power. *Journal of the European Economic Association*, *21*(5), 2172–2208. https://doi.org/10.1093/jeea/jvad012

Blackwell, A. D., Gurven, M. D., Sugiyama, L. S., Madimenos, F. C., Liebert, M. A., Martin, M. A., Kaplan, H. S., & Snodgrass, J. J. (2011). Evidence for a Peak Shift in a Humoral Response to Helminths: Age Profiles of IgE in the Shuar of Ecuador, the Tsimane of Bolivia, and the U.S. NHANES [Publisher: Public Library of Science]. *PLOS Neglected Tropical Diseases*, *5*(6), e1218. https://doi.org/10.1371/journal.pntd.0001218

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

Breen, C., & Goldstein, J. R. (2022). Berkeley Unified Numident Mortality Database: Public administrative records for individual-level mortality research. *Demographic Research*, *47*, 111–142. Retrieved April 18, 2024, from https://www.demographic-research.org/articles/volume/47/5/

Briggs, N., Weatherhead, J., Sastry, K. J., & Hotez, P. J. (2016). The Hygiene Hypothesis and Its Inconvenient Truths about Helminth Infections [Publisher: Public Library of Science]. *PLOS Neglected Tropical Diseases*, *10*(9), e0004944. https://doi.org/10.1371/journal.pntd.0004944

Caruso, C., Candore, G., Colonna-Romano, G., Lio, D., & Franceschi, C. (2005). Inflammation and Life-Span [Publisher: American Association for the Advancement of Science]. *Science*, *307*(5707), 208–209. https://doi.org/10.1126/science.307.5707.208

Cohen, A. A., Tillinghast, J., & Canudas-Romo, V. (2010). No consistent effects of prenatal or neonatal exposure to Spanish flu on late-life mortality in 24 developed countries. *Demographic research*, *22*(20), 579–634. https://doi.org/10.4054/DemRes.2010.22.20

Commission, R. (1914). The Rockefeller Sanitary Commission for the Eradication of Hookworm Disease, Fourth Annual Report for the year 1913. *HathiTrust*. Retrieved April 18, 2024, from https://hdl.handle.net/2027/mdp.39015073478136?urlappend=%3Bseq=5

Crimmins, E. M., & Finch, C. E. (2006). Infection, inflammation, height, and longevity [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *103*(2), 498–503. https://doi.org/10.1073/pnas.0501470103

Croke, K., & Atun, R. (2019). The long run impact of early childhood deworming on numeracy and literacy: Evidence from Uganda. *PLoS Neglected Tropical Diseases*, *13*(1), e0007085. https://doi.org/10.1371/journal.pntd.0007085

Croke, K., Hamory, J., Hsu, E., Kremer, M., Maertens, R., Miguel, E., & Więcek, W. (2024). Meta-analysis and public policy: Reconciling the evidence on deworming [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *121*(25), e2308733121. https://doi.org/10.1073/pnas.2308733121

Cutler, D., & Meara, E. (2001, October). Changes in the Age Distribution of Mortality Over the 20th Century. https://doi.org/10.3386/w8556

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

Elman, C., McGuire, R. A., & Wittman, B. (2014). Extending Public Health: The Rockefeller Sanitary Commission and Hookworm in the American South [Publisher: American Public Health Association]. *American Journal of Public Health*, *104*(1), 47–58. https://doi.org/10.2105/AJPH.2013.301472

Ettling, J. (2013, October). *The Germ of Laziness: Rockefeller Philanthropy and Public Health in the New South*. Harvard University Press. https://doi.org/10.4159/harvard.9780674333345

Finch, C. E. (2010, July). *The Biology of Human Longevity: Inflammation, Nutrition, and Aging in the Evolution of Lifespans* [Google-Books-ID: 4rGOLEOC4mEC]. Elsevier.

Finch, C. E., & Crimmins, E. M. (2004). Inflammatory Exposure and Historical Changes in Human Life-Spans [Publisher: American Association for the Advancement of Science]. *Science*, *305*(5691), 1736–1739. https://doi.org/10.1126/science.1092556

Fletcher, J., & Noghanibehambari, H. (2023, December). Toxified to the Bone: Early-Life and Childhood Exposure to Lead and Men's Old-Age Mortality. https://doi.org/10.3386/w31957

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

Goldin, C., & Lleras-Muney, A. (2019). XX>XY?: The changing female advantage in life expectancy. *Journal of Health Economics*, *67*, 102224. https://doi.org/10.1016/j.jhealeco.2019.102224

Goldstein, J., Alexander, M., Breen, C., Miranda Gonzalez, A., Menares, F., Osborne, M., Mallika, S., & Ugur, Y. (2023). CenSoc Mortality File.

Gurven, M., Kaplan, H., Winking, J., Finch, C., & Crimmins, E. M. (2008). Aging and Inflammation in Two Epidemiological Worlds. *The Journals of Gerontology: Series A*, *63*(2), 196–199. https://doi.org/10.1093/gerona/63.2.196

Hamory, J., Miguel, E., Walker, M., Kremer, M., & Baird, S. (2021). Twenty-year economic impacts of deworming [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *118*(14), e2023185118. https://doi.org/10.1073/pnas.2023185118

Hotez, P. J., Alvarado, M., Basáñez, M.-G., Bolliger, I., Bourne, R., Boussinesq, M., Brooker, S. J., Brown, A. S., Buckle, G., Budke, C. M., Carabin, H., Coffeng, L. E., Fèvre, E. M., Fürst, T., Halasa, Y. A., Jasrasaria, R., Johns, N. E., Keiser, J., King, C. H., … Naghavi, M. (2014). The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases [Publisher: Public Library of Science]. *PLOS Neglected Tropical Diseases*, *8*(7), e2865. https://doi.org/10.1371/journal.pntd.0002865

Kıskaç, M., Soysal, P., Smith, L., Capar, E., & Zorlu, M. (2022). What is the Optimal Body Mass Index Range for Older Adults? [Publisher: The Korean Geriatrics Society]. *Annals of Geriatric Medicine and Research*, *26*(1), 49. https://doi.org/10.4235/agmr.22.0012

Kofoid, C. A., & Tucker, J. P. (1921). ON THE RELATIONSHIP OF INFECTION BY HOOKWORM TO THE INCIDENCE OF MORBIDITY AND MORTALITY IN 22,842 MEN OF THE UNITED STATES ARMY AT CAMP BOWIE, TEXAS,

FROM OCTOBER 1917, TO APRIL 1918. *American Journal of Epidemiology*, *1*(1), 79–117. https://doi.org/10.1093/oxfordjournals.aje.a118027

Kopinska, J. A., Atella, V., Bhattacharya, J., & Miller, G. (2021, May). The Changing Relationship between Bodyweight and Longevity in High- and Low-Income Countries. https://doi.org/10.3386/w28813

Loukas, A., Hotez, P. J., Diemert, D., Yazdanbakhsh, M., McCarthy, J. S., Correa-Oliveira, R., Croese, J., & Bethony, J. M. (2016). Hookworm infection [Publisher: Nature Publishing Group]. *Nature Reviews Disease Primers*, *2*(1), 1–18. https://doi.org/10.1038/nrdp.2016.88

Martorell, R. (1995). Results and Implications of the INCAP Follow-up Study. *The Journal of Nutrition*, *125*, 1127S–1138S. https://doi.org/10.1093/jn/125.suppl_4.1127S

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

Miguel, E., & Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.146 0262.2004.00481.x]. *Econometrica*, *72*(1), 159–217. https://doi.org/10.1111/j.1468-0262.2004.00481.x

Miller, G. (2008). Women's Suffrage, Political Responsiveness, and Child Survival in American History*. *The Quarterly Journal of Economics*, *123*(3), 1287–1327. https://doi.org/10.1162/qjec.2008.123.3.1287

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 [Publisher: American Public Health Association]. *American Journal of Public Health*, *103*(7), e83–e90. https://doi.org/10.2105/AJPH.2012.301060

Noghanibehambari, H., & Fletcher, J. (2024, December). Unhooking the Past: Early-life Exposure to Hookworm Eradication and Later-life Longevity.

Noymer, A., & Garenne, M. (2000). The 1918 Influenza Epidemic's Effects on Sex Differentials in Mortality in the United States [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.17 4457.2000.00565.x]. *Population and Development Review*, *26*(3), 565–581. https://doi.org/10.1111/j.1728-4457.2000.00565.x

Ramanan, D., Bowcutt, R., Lee, S. C., Tang, M. S., Kurtz, Z. D., Ding, Y., Honda, K., Gause, W. C., Blaser, M. J., Bonneau, R. A., Lim, Y. A., Loke, P., & Cadwell, K. (2016). Helminth infection promotes colonization resistance via type 2 immunity [Publisher:

American Association for the Advancement of Science]. *Science*, *352*(6285), 608–612. https://doi.org/10.1126/science.aaf3229

Roodman, D. (2018). The Impacts of Hookworm Eradication in the American South. A replication study of Bleakley (The Quarterly Journal of Economics, 2007) [Publisher: Kiel, Hamburg: ZBW – Leibniz Information Centre for Economics]. *International Journal for Re-Views in Empirical Economics (IREE)*, *2*(2018-3), 1–45. https://doi.org/10.18718/81781.7

Strachan, D. P. (1989). Hay fever, hygiene, and household size. *BMJ : British Medical Journal*, *299*(6710), 1259–1260. Retrieved April 1, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1838109/

Strauss, J., & Thomas, D. (1998). Health, Nutrition, and Economic Development [Publisher: American Economic Association]. *Journal of Economic Literature*, *36*(2), 766–817. Retrieved December 8, 2024, from https://www.jstor.org/stable/2565122

Taylor-Robinson, D. C., Maayan[a], N., Donegan, S., Chaplin, M., & Garner, P. (2019). Public health deworming programmes for soil-transmitted helminths in children living in endemic areas [Publisher: John Wiley & Sons, Ltd]. *Cochrane Database of Systematic Reviews*, (9). https://doi.org/10.1002/14651858.CD000371.pub7

Wammes, L. J., Mpairwe, H., Elliott, A. M., & Yazdanbakhsh, M. (2014). Helminth therapy or elimination: Epidemiological, immunological, and clinical considerations [Publisher: Elsevier]. *The Lancet Infectious Diseases*, *14*(11), 1150–1162. https://doi.org/10.1016/S1473-3099(14)70771-6

Yazdanbakhsh, M., Kremsner, P. G., & van Ree, R. (2002). Allergy, Parasites, and the Hygiene Hypothesis [Publisher: American Association for the Advancement of Science]. *Science*, *296*(5567), 490–494. https://doi.org/10.1126/science.296.5567.490

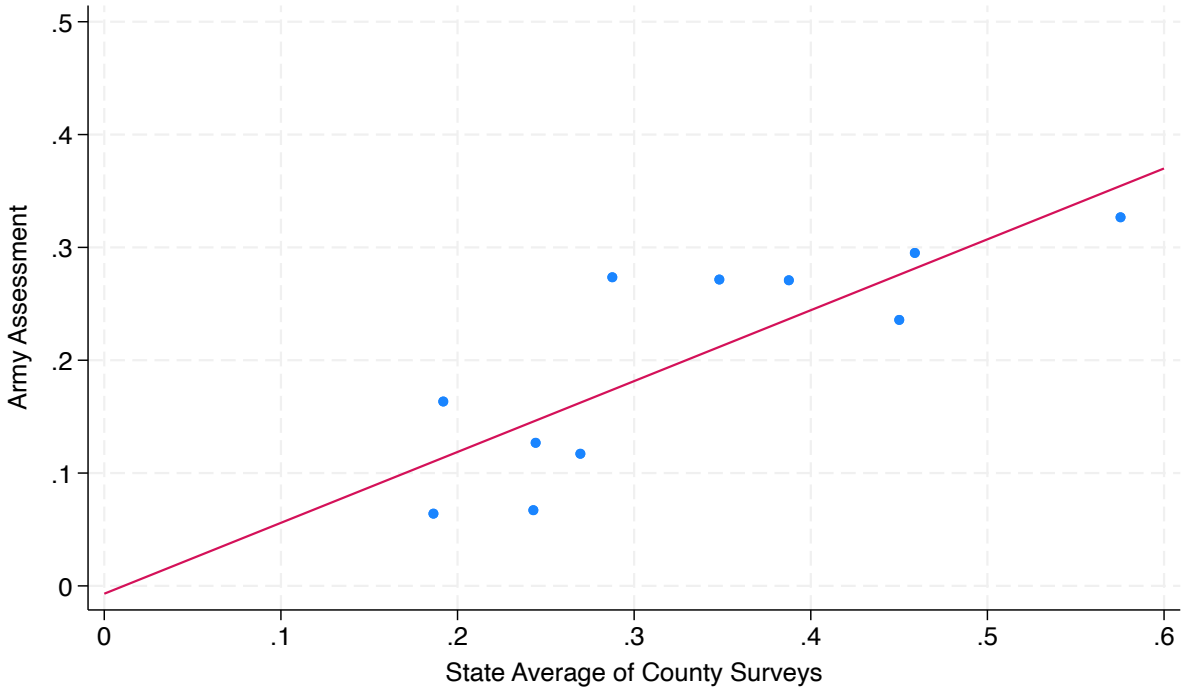

Figure A1: Relationship between army hookworm prevalence and county-level RSC assessed prevalence in overlapping states

Notes: This figure shows the relationship between average hookworm prevalence from systematic RSC surveys of children in each county, and average hookworm prevalence from surveys of army recruits. The line of best fit is included.

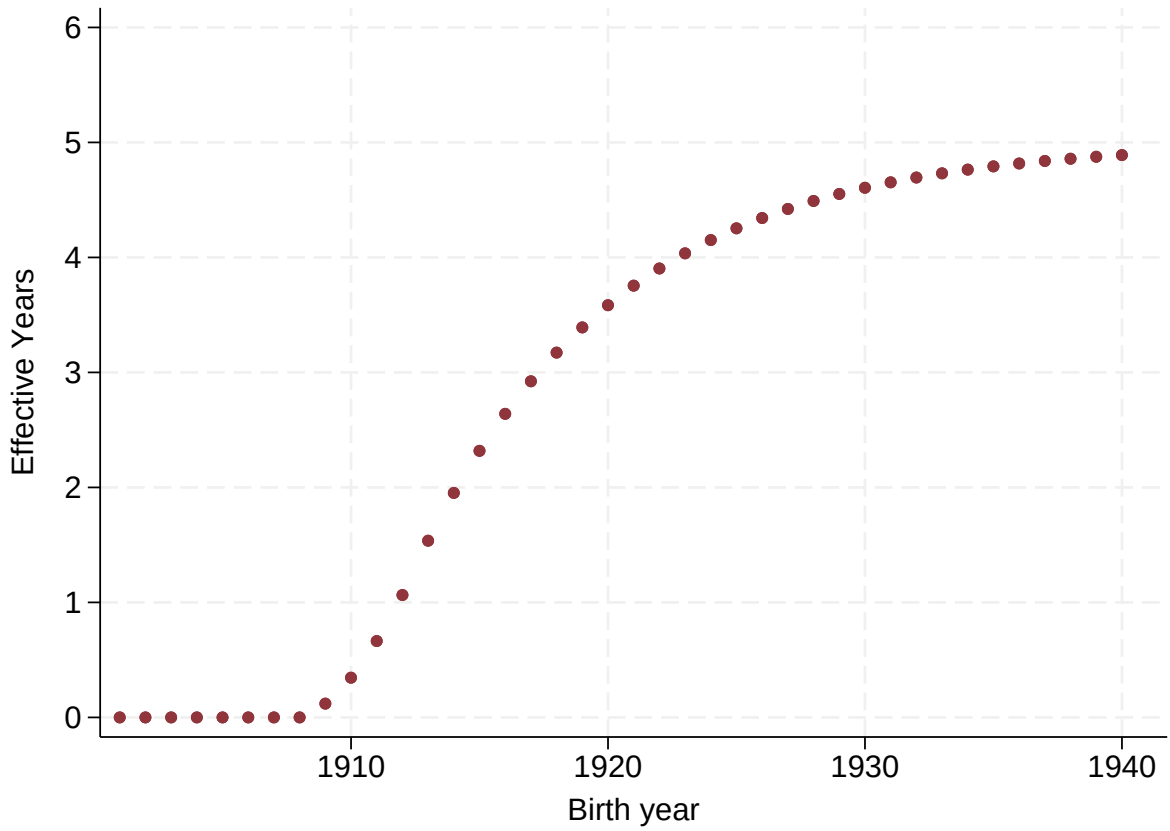

Figure A2: Effective years of exposure assigned to each birth cohort

Notes: This figure shows the number of effective years of exposure assigned to each birth cohort over time.

Table A1: Results from alternative specifications

	Longevity (years)		
	All	**Males**	**Females**
A. Semi-parametric specification			
SEA tertile 1 X Effective Exposure Years (omitted)	-	-	-
SEA tertile 2 X Effective Exposure Years	0.016	0.014	0.017
	[0.006]	[0.008]	[0.008]
SEA tertile 3 X Effective Exposure Years	0.019	0.017	0.020
	[0.007]	[0.008]	[0.009]
Constant	77.772	76.104	79.199
	[0.011]	[0.014]	[0.013]
N	3,980,291	1,851,939	2,128,352
B. Pre-post specification			
Baseline SEA prevalence X [1] >2 effective exposure years	0.112	0.049	0.142
	[0.048]	[0.049]	[0.061]
Constant	77.777	76.122	79.197
	[0.013]	[0.014]	[0.015]
N	3,980,291	1,851,939	2,128,352
C. Linear years of exposure before age 5			
Baseline SEA prevalence X years since 1912	0.028	0.023	0.030
	[0.011]	[0.011]	[0.014]
Constant	77.769	76.104	79.195
	[0.014]	[0.015]	[0.017]
N	3,980,291	1,851,939	2,128,352

Notes: This table presents estimates of the effect of exposure to the de-worming campaign, first pooled then stratified by sex. Panel A presents coefficients from an alternative semi-parametric specification using the interaction between tertile of baseline hookworm prevalence and years of exposure, with coefficients relative to the lowest tertile. Panel B presents coefficients from a simplified pre-post framework, using those with less than 2 effective years of exposure as the reference group. Panel C most closely resembles the paper's primary specification, but uses a linear years of exposure before age 5 term, where the years of exposure are calculated as the birth year-1912, truncated at 0 on the low end and 5 on the high end. All empirical specifications reflect the primary specification using SEA fixed-effects that are allowed to vary by sex, and state-by-birth year fixed-effects. Data are drawn from the CenSoc BUNMD, weighted such that the distribution matches that from the Human Mortality Database. Hookworm prevalence is assigned at the SEA. Standard errors presented in brackets, clustered at the SEA level.

Table A2: Impacts of de-worming on survival until 1960

	Survival rates until 1960		
	All	**Males**	**Females**
A. Primary specification			
State Prevalence X Effective Exposure Years f	0.051	0.038	0.064
	[0.028]	[0.029]	[0.026]
Constant	0.770	0.738	0.804
	[0.008]	[0.008]	[0.007]
N	341	341	341

Notes: This table presents estimates of the effect of exposure to the de-worming campaign, first pooled then stratified by sex. Observations only include RSC-treated states, and calculate the survival from early life (after infant mortality) until 1960 in the census. The empirical specification includes birth year and birth state fixed-effects. Standard errors presented in brackets, clustered at the state level.