北京大学中国经济研究中心
China Center for Economic Research
讨论稿系列
Working Paper Series

No.E2017007 2017-07-08

Schistosomiasis Control and Later Life Outcomes: Evidence from China

By CHAO LIU AND GORDON G. LIU*

Schistosomiasis is endemic in over 70 countries, bringing about large economic and public health loss in tropical and subtropical areas. This study uses an exogenous disease control campaign in China to identify causal impact of early life exposure to schistosomiasis on lifetime outcomes. We find that the program led to increases in educational attainment, labor participation and the number of births. The results are robust to a set of specification checks and using IV estimates. Using the CFPS data set, we confirm the above results and find evidence of increased height for men, better cognitive functions and economic status.

Schistosomiasis Control and Later Life Outcomes: Evidence from China

By CHAO LIU AND GORDON G. LIU[*]

Schistosomiasis is endemic in over 70 countries, bringing about large economic and public health loss in tropical and subtropical areas. This study uses an exogenous disease control campaign in China to identify causal impact of early life exposure to schistosomiasis on lifetime outcomes. We find that the program led to increases in educational attainment, labor participation and the number of births. The results are robust to a set of specification checks and using IV estimates. Using the CFPS data set, we confirm the above results and find evidence of increased height for men, better cognitive functions and economic status.

Schistosomiasis, one of the most prevalent parasitic infectious diseases in the world, is endemic in more than 70 countries and territories. Since schistosomiasis is a chronic insidious disease, it is hard to identify and treat the disease during early stages. In 2014, at least 258 million people are estimated to need treatment for schistosomiasis, but the actual reported number of people treated in 2014 was only 61.6 million (WHO 2016). Schistosomiasis negatively affects men and women's productivity during their adulthood, becoming a huge threat to socioeconomic development in tropical and subtropical areas. Although schistosomiasis is the second most severe parasitic infectious disease after malaria in many tropical and subtropical areas, especially in Africa, economists have paid little attention to it.

A growing literature tries to answer this question: does a heavy health burden hold back socioeconomic development? A key challenge here is the reverse causality, making a naïve regression biased, as the failure to eradicate a certain disease might be the result of underdevelopment. This paper exploits an exogenous intervention in public health, the schistosomiasis control campaign in China in late 1950s as a natural

---

[*] Liu: National School of Development, Peking University, Beijing, 100871, China (e-mail: ccerliuchao@pku.edu.cn). Liu: National School of Development, Peking University, Beijing, 100891, China (gordonliu@nsd.pku.edu.cn). We have benefited from comments and help of Anna Aizer, Richard Freeman, Lixing Li and participants at 2017 CCER summer institute.

experiment, to identify the long-run effect of childhood exposure to schistosomiasis on adult outcomes. The strategy of this paper is similar to that employed by Bleakley (2010), Cutler et al. (2010), and Lucas (2010).

In this paper, we focus on four kinds of later life outcomes: educational attainment, labor productivity, female fertility and adult health. Schistosomiasis might affect educational attainment through at least three channels. First, maternal schistosomiasis results in low birth weight and preterm delivery, which may impair subsequent cognitive and physical development of infants. Second, early-life infections may cause poor growth and learning difficulty. Third, if children live in areas with schistosomiasis prevalence, they are at high risk of malnutrition and poor schooling due to the underdevelopment caused by disease infections. Chronic schistosomiasis could impair people's ability to work, so we estimate the impact on the probability of having a job in adulthood. Similar to Cutler et al. (2010), we also use per capita household consumption expenditure to study the impact on labor productivity. Female genital schistosomiasis may cause infertility, preterm labor, anemia, menstrual disorders, and dyspareunia (Nour 2010), and hence we examine the effect on female fertility. Lastly, it is straightforward to estimate the effect of a health intervention in early life on adult health, measured by height and cognitive functions.

We combine individual data from 1990 China Population Census and the China Family Panel Studies (CFPS) with schistosomiasis prevalence of infection at county level. The difference-in-differences model compares gains of adult outcomes (e.g. years of education) for cohorts born before and after the disease control program in areas with varying pre-control incidence. Our baseline estimates of census sample show that schistosomiasis prevalence reduction significantly increased years of schooling, possibility of having a job and the number of births. The preferred estimates suggest that reducing schistosomiasis incidence by 30 percentages led to an increase in about 0.34 (0.97) schooling years for rural males (females). Exploiting the more information we can get in the CFPS dataset, we find the disease control program also generated better health and economic status in adulthood.

To the best of our knowledge, this is one of the first papers to assess the effect of early life exposure to schistosomiasis on human capital accumulation and adult socio-economic status. The most relevant study to our paper is Li and Wei (2017), which finds that the prevention and cure of schistosomiasis since 1950s in China had a positive impact on population growth. Besides our specific estimates, this paper also contributes to three strands of literature in economics.

Firstly, our paper relates to a set of studies about assessing the effects of disease eradication programs. Most of these studies exploit the utilization of DDT as an exogenous event to identify the impact of childhood exposure to malaria on education, but they have contradictory results. On the one hand, Barreca (2010), Lucas (2010) and Barofsky et al. (2015) show malaria eradication significantly raised educational attainment in the United States, Sri Lanka and Paraguay, and southwestern Uganda, respectively. On the other hand, Cutler et al. (2010) find that malaria eradication did not increase educational level for men, and the effects for women was mixed in India. Bleakley (2010) and Venkataramani (2012) show that malaria eradication has limited impact on schooling in Mexico. This paper examines another important but neglected tropical disease, and finds robust positive effect of disease control program on education.

Our paper fits into a large existing literature that examines the relationships between early life health status and adult outcomes. British physician David Barker (1990) first proposed the fetal origins hypothesis that fetal development may predetermine certain chronic health in later life, then economics researches emerge to empirically test this hypothesis (Almond 2006; Case and Paxson 2009; Lee 2014). In a China context, previous studies mainly focus on the effect of the Great Famine from 1959 to 1961 (Chen and Zhou 2007; Meng and Qian 2009). Our findings show that people exposed to a higher schistosomiasis risk around their birth tend to have worse health outcomes (measured by height and cognitive functions) in their old ages, supporting the fetal-origins hypothesis.

This paper also joins the debate about the sources of underdevelopment. Previous studies argue that economic growth could benefit from the improvements in health (e.g., Gallup and Sachs 2001; Bloom and Canning 2005). Weil (2007) finds that the variance of log GDP per worker becomes smaller after taking health gaps among countries into consideration. However, Acemoglu and Johnson (2007) find large increase in life expectancy failed to raise income per capita. Our estimates show that the reduction of schistosomiasis infection could release part of the economic growth potential by improving human capital, in both education and health levels.

This paper proceeds as follows. Section I introduces the disease and the efforts against it in China. Section II describes the data and empirical strategy used in this paper. Section III presents our main empirical results. Section IV shows the results of various robustness tests and Section V concludes.

I. Background

A. Schistosomiasis

Schistosomiasis is an acute and chronic parasitic disease caused by blood flukes (trematode worms) of the genus Schistosoma. People get infected when larval forms of the parasite penetrate the skin through infested water. Patients' excreta contain parasite eggs, so disease transmission occurs if they contaminate freshwater. Typical symptoms of this disease include abdominal pain, diarrhea, bloody stool, or blood in the urine. Long-time adult patients may develop kidney failure, liver damage, and even bladder cancer. Children may suffer from poor growth and learning difficulty as a result (WHO 2017).

Several archeological evidence has demonstrated the long history of schistosomiasis in China. Experts in medical history found a description recording similar clinical symptoms of acute schistosomiasis in a traditional Chinese medicine book dating back to about 400 B.C. Archeologists detected schistosome eggs in a female corpse that can be traced back to the Western Han dynasty (Mao and Shao 1982).

China was the largest endemic area of S. japonicum infection.[1] Schistosomiasis spread through 10 provinces which in the south of China or along the Yangtze River: Anhui, Fujian, Guangdong, Guangxi, Hunan, Hubei, Jiangsu, Jiangxi, Shanghai, Sichuan, Yunnan, and Zhejiang. In the 1950s, the estimated number of patients was 10 million (Mao and Shao 1982). Endemic areas have three types: marshland and lake regions, hilly and mountainous regions, and plain regions with waterway networks.

B. Schistosomiasis Control in China

Chinese scientists started to carry out epidemiological surveys to determine the prevalence of S. japonicum infections after 1949. Faced with the huge disease burden, central government set up a leading group to be in charge of the nationwide schistosomiasis control in 1955. After Chair Mao's public resolution to eradicate schistosomiasis in 1956, the central government issued an instruction on the eradication of schistosomiasis in the next year. A disease control campaign was carried out in twelve endemic provincial administrative units shortly after the instruction. This campaign relied heavily upon central government's willing from outside the affected areas, which offers us an exogenous change in the schistosomiasis infection rate for our research design.

Special health stations for anti-schistosomiasis was set up at three levels in endemic areas (provincial, county and community). These health stations took various actions to fight against the disease, such as chemotherapy and snail control. Hygienic movement and health education were also used as preventive measures. Although health workers in medical schools and cities were dispatched to help with control in rural endemic areas, the majority work of health care was extended through 'barefoot doctor' system.[2] In addition, government in the endemic areas usually bore the cost of treatment, so that farmers could receive treatment freely.

---

[1] Schistosomiasis is caused by five main species of blood fluke: S. mansoni, S. japonicum, S. mekongi, S. guineensis, and S. haematobium.

[2] In China, barefoot doctors are special medical workers in rural areas where urban doctors would not settle. They received minimal basic medical and paramedical training. In addition to treating common diseases, they provided basic preventive health care and family planning.

The eradication efforts resulted in a great decline of schistosomiasis infections. According to statistic data in 1989, 268 among 373 counties that formerly were endemic for schistosomiasis succeeded to interrupt and control the disease transmission, which proves the effectiveness of the schistosomiasis control campaign in China (Yuan, 1989).

II. Data and Empirical Strategy

A. Data

For measures of schistosomiasis prevalence, we collect data from China Schistosomiasis Atlas. This book contains detailed information on number of patients, prevalence of infection, and snail habitat areas at county level for both before and after the disease control campaign. We made some adjustments accordingly to correct for the administrative unit changes to match our individual data. We use prevalence of infection as the indicator for the pre-control schistosomiasis prevalence. Observations with missing information on this indicator are excluded, leaving a final sample of 329 endemic counties in 12 provincial administrative units (858 counties). Panel A and B of Figure 1 show the pre-and post-control prevalence respectively.[3] Uninfected counties in these 12 endemic provinces are served as control groups. We match these disease data to individual level dataset from two sources.

[Insert Figure 1 here]

The main individual level data used in this paper comes from the 1990 Population Census in China. This dataset contains basic demographical variables: gender, year of birth, ethnicity, type of hukou (urban/rural), county of hukou, marital status and educational level. The data have information about labor market participation for respondents older than 16 years old. We can also know how many births a woman over age 16 has given at the survey time. The original education variable in the dataset has six levels: illiterate, primary school, junior school, senior middle school, college, and

[3] Pre-control data were collected in 1950s, and post-control data reflect infection rates in early 1980s.

graduate or above. We generate a continuous variable measuring schooling years baesd on this categorical variable. Using the information of individual's relation to the head of household, we also calculate educational years of one's parents. We limit the sample to those with rural hukou, born after 1930 and aged 18 or older at the time of the census (1990). Table 1 provides summary statistics for census data in 12 endemic provinces. 50.83% of the sample are male. Males on average have higher education levels than females. Most people have Han ethnicity and are at work. On average, each woman gave 2.37 births.

[Insert Table 1 here]

Another dataset used in this paper is the China Family Panel Studies (CFPS). The CFPS is a large-scale panel survey which collects data on household members in 25 provinces every two years. This study uses the first national wave data collected in 2010. In consistency with the census data, we keep respondents born between 1930 and 1972. As the CFPS data were collected in 2010, we have to address the migration problem, so we exclude mover individuals.[4] We use the type of hukou when the respondent was 3 years old as the proxy of childhood living areas. Due to the much smaller sample size of the CFPS, we keep individuals in all available provinces.[5] This sample includes 19 endemic counties in 25 provinces (117 counties). Summary statistics of key variables are shown in Table 2. Compared with the census sample, the CFPS sample has lower average years of schooling but with larger variations, so more people in the CFPS sample completed junior school.

[Insert Table 2 here]

B. Empirical Strategy

We use a differences-in-differences model to identify causal impacts of early life

[4] We keep individuals whose residential county is the same as the county of birth.

[5] Besides 12 endemic provinces, we also include Beijing, Tianjin, Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Shandong, Henan, Chongqing, Guizhou, Shaanxi, and Gansu in this sample.

exposure to schistosomiasis on later life outcomes. This specification relies on two sources of variation. First, different areas have different schistosomiasis incidence prior to the control program. Second, different cohorts benefit from the disease control campaign differently. Unlike the traditional DID setup, the treatment variable in our econometric model is continuous.[6] We run regressions of the following form:

$$(1)\ Y_{ijc} = \alpha + \beta\left(sch_j \times post_c\right) + X_{ijc}{}'\Theta + \delta_j + \delta_c + \epsilon_{ijc},$$

where Y_{ijc} is an adult outcome of individual i in county j, who belongs to cohort c. sch_j is the pre-control schistosomiasis intensity at county level; $post_c$ is a dummy variable indicating the individual was born after the program; X_{ijc} is a set of control variables, which includes ethnicity and educational years of parents if Y_{ijc} is an educational outcome; δ_j and δ_c are regional fixed effects and cohort fixed effect. We define $post_c$ equals to 1 for individuals born after 1957 and 0 otherwise. Our coefficient of interest is β, representing the difference-in-differences estimate of the effect of schistosomiasis control. We estimate equation (1) for men and women separately.

Equation (1) actually tests fetal-origins hypothesis since it assumes only the year of birth matters. However, school-aged children are most vulnerable to infection of schistosomiasis because they are more likely to play in infested water. Therefore, following Lucas (2010), we define $exposure_c$ of cohort c as the percentage of years prior to age 18 that are spent in the post-control period. The corresponding specification is

$$(2)\ Y_{ijc} = \alpha + \beta\left(sch_j \times exposure_c\right) + X_{ijc}{}'\Theta + \delta_j + \delta_c + \epsilon_{ijc},$$

with other notations the same as in equation (1).

---

[6] Figure A1 illustrates the validity of our identification strategy in a traditional DID setup. It shows the educational years of different cohorts in endemic counties and those free of this disease. The treatment group and the control group show similar trends before 1957, but they diverge significantly when the schistosomiasis control program started in 1957.

A few additional issues are worth noting. First, the census data only report the county of current residence and hukou type. Ideally, we would know each individual's county of birth, and whether one lives in rural areas during childhood. Thus, an assumption here is that people live in their birthplace and hukou type is a good proxy for living areas. Fortunately, before 1990 this is the case in China due to the strict household registration system. In addition, we know exactly the birth location in the CFPS sample, alleviating the migration problem. Second, a necessary assumption of our research design is that areas where schistosomiasis was highly endemic benefited more from the control campaign. Data from the China Schistosomiasis Atlas allow us to examine this. Figure 2 illustrates that areas with higher infection rates indeed saw a greater drop in disease prevalence. If the county eradicated schistosomiasis, the dot would locate on the diagonal. Third, an underlying assumption of difference-in-differences identification strategy is that there are no differential changes in dependent variables of our estimations between infected and uninfected areas in the absence of the schistosomiasis control program. A pre-existing regional convergence would violate this assumption and bias our estimations. To address this problem, we estimate additional specifications that add province × post controls or control regional trends, and we also check mean-reversion problems in Section IV.

[Insert Figure 2 here]

III. Empirical Results

A. The Effect on Educational Attainment

Table 3 presents descriptive evidence of the effects of on educational attainment. We use pre-control schistosomiasis infection rates to classify endemic counties into highly endemic or less endemic groups. sch_j between 0 and 15 are classified as low epidemic, and sch_j over 15 are classified as highly epidemic. Cohorts born during 1930-1956 are the pre-control group, whereas those born from 1957-1972 are the post-control group. Cohorts born after 1957 had higher educational attainment than those born before the schistosomiasis control program. As shown in the seventh column, highly endemic counties usually had larger increases in educational attainment than less endemic counties, and both two types of endemic areas benefited more than uninfected areas.[7] In addition, females had a larger gain in educational attainment after the disease control campaign.

[Insert Table 3 here]

Then we investigate the effects of early life exposure to schistosomiasis on educational attainment by estimating equation (1). Table 4 reports the results of our baseline specification for schooling years, whether one is literate, completes primary and junior school education. Standard errors are clustered at county level.

Panel A of Table 4 presents the results for men. In even columns, we add educational level of parents, ethnicity (=1 if belongs to Han ethnicity) as individual-level controls. The results are robust if we exclude these covariates. Our estimates demonstrate that schistosomiasis control had a significantly positive effect on educational attainment. Based on the coefficients in column 2, 6, and 8, a 30-percentage point decrease in incidence is associated with an increase of 0.34 years in the expected years of schooling for rural males, an increase of 4.23 percent in the probability of completing primary school education, and an increase of 6.39 percent in the probability of junior school completion.[8] Thus, the increase in junior school completion contributes most to the increase of schooling years. However, it seems that schistosomiasis control had little impact on literacy rates for men. We attribute this phenomenon to a high pre-control literacy rate (89%) for rural males, leaving less room for improvement.[9] Another possible explanation is selective mortality. Since the weakest members in one cohort may survive after the disease control, the education attainments of these individuals stay at a low level regardless of the reduction of incidence. We will discuss concerns

7 The increase of educational years in uninfected counties was 2.09 years, while highly endemic and less endemic areas had a gain of 2.37 and 2.34 years.

8 The ninetieth percentile of the schistosomiasis prevalence is 31.4%.

9 According to Cutler et al. (2010), only 53% males (born in 1912-1952 and 1962-1972) in their India sample are literate.

about selective mortality in more detail below. Statistically significant findings also exist for rural women, as is shown in Panel B of Table 4. In all specifications, coefficients are larger than those for men and are significantly positive in column 3-4 (literate), which implies rural females benefited more from the reduction of schistosomiasis prevalence.

[Insert Table 4 here]

To capture the potential nonlinear effects of schistosomiasis exposure, we use a categorical classification of pre-control endemicity. Pre-control prevalence of infection that equals to zero is classified as non-epidemic, sch_j between 0 and 15 are classified as low epidemic (low_j), and sch_j over 15 are classified as high epidemic ($high_j$). The coefficients on $low_j \times post_c$ and $high_j \times post_c$ capture the effect of being born after control versus before control in a county which was mildly or highly epidemic, relative to the effect of being born in a county free of schistosomiasis.[10] If schistosomiasis reduction improved educational attainment in rural areas, we would see positive coefficients on these two interaction terms. If there was a linear effect, the first coefficient ($low_j \times post_c$) should be smaller than the second one. Table A1 reports the results that are similar to those if we use a continuous measure of pre-control endemicity. We can observe a significant gain of years of schooling, primary school and junior school completion for people in infectious counties, except for literacy rates. In addition, we do find a nonlinear effect of schistosomiasis control, as individuals in formerly high-endemic areas benefited more with respect to primary school completion, while mildly areas gained more with respect to junior school completion. This may reflect that areas with less schistosomiasis prevalence could boost educational development at a relatively higher level.

We re-estimate the baseline specification in a restricted sample including control counties (uninfected) and the eradicated counties (i.e. have very low post-control prevalence of infection, say below 0.5%). By doing so, we can better interpret our result

[10] $non_j \times post_c$ is omitted as a reference group.

as a treatment effect of the disease eradication, circumventing the endogeneity problems. If we directly use the difference between pre-control and post-control infection rates to measure the treatment effect of schistosomiasis control, one may concern that counties experiencing a larger reduction of infection were also good at improving local education. However, schistosomiasis levels in these eradicated countries were reduced to approximately zero, so the pre-control infection rate is also the regional intensity of the control campaign. Table A2 presents results for this eradicated sample. The schistosomiasis control program greatly improved educational level except for male literacy rate, just like our baseline results. These results do not change much if we choose 1% as the threshold of eradication (see Table A3).

One more potential concern is whether the indicator of disease incidence used here truly reflect the geographic variations in schistosomiasis prevalence in the pre-control period. For instance, some counties may have incentives to report a higher infection rate to get more fiscal support from upper level governments. If classic measurement error exists, our estimates would be biased, attenuating to zero. The transmission of schistosomiasis relies on infected water in most endemic counties in China, so we use the county's proportion of water area as an instrument for pre-control infection rate. We replicate the baseline estimates in Table 4 and display the results in Table 5. The pattern of the results is very similar, and the point estimates are larger in magnitude than the OLS estimates in all specifications. Our IV estimates reinforce our confidence in the validity of the baseline results.

[Insert Table 5 here]

We next estimate equation (2) to allow for a differential incidence of disease control benefits across cohorts. Due to the property of this disease that school-aged children are more likely to be infected, this specification is pertinent in this paper. The results in Table 6 are very similar with what we have found in Table 4, except the magnitude of coefficients because of the different econometric models. Cohorts born after schistosomiasis control campaign are expected to have higher level of education than those born prior to control. We again find no evidence that the program had impact on literacy rate for males. From column 1, if the schistosomiasis infection rate drops 30 percentages, rural males born after 1957 ($exposure_c$=1) would gain 0.63 more years of education than cohorts born before 1939 ($exposure_c$=0). For rural females, this effect extends to 2.2 years.

[Insert Table 6 here]

B. Cohort Analysis

We next conduct a cohort analysis so that we can present the results visually. Formally, we estimates the following specification

$$(3)\quad Y_{ijc} = \alpha + \sum \beta_c \left(sch_j \times \delta_c \right) + X_{ijc}{}' \Theta + \delta_j + \delta_c + \epsilon_{ijc},$$

where we replace $\beta\left(sch_j \times post_c\right)$ in equation (1) with a separate interaction term for each cohort, allowing the effect to vary by each cohort. If the schistosomiasis control campaign had a positive treatment effect, we will see the coefficients for post-control cohorts are larger than those for pre-control cohorts.

Inspired by our estimates in Table 4, Figure 3 plots the cohort-specific coefficient of interaction terms in equation (3) in female sample born during 1948-1968, with a dummy for junior school completion as the dependent variable. The coefficients fluctuate around 0 for cohorts born before 1957. After 1957, all coefficients are significantly positive. This supports that we can treat people born prior to 1957 as pre-control cohorts as well. Results for other educational dependent variables are shown in Figure A2 and Figure A3 for females and males, respectively.

[Insert Figure 3 here]

C. The Effect on Other Adult Outcomes

In addition to educational attainment, we examine the effect of schistosomiasis on other life cycle outcomes. As is discussed in Section I.A, chronic schistosomiasis greatly impairs people's ability to work. This disease may cause infertility in the long term. For instance, urogenital schistosomiasis can induce pathology of the seminal vesicles in men, and genital lesions, vaginal bleeding in women. Therefore, we assess the impact of early life exposure to schistosomiasis on labor participation and female fertility. Column 1-2 of Table 7 contain the estimates of equation (1) for working status as the dependent variable, and column 3-4 report results for the number of births for females. Individual controls in column 1-2 are ethnicity and marital status, and we add individual educational years as a covariate in column 3-4. Column 1 presents a reduction of schistosomiasis infection can significantly increase the possibility of having a job in adulthood for both males and females.[11] IV estimates are reported in column 2. The effect on female labor participation is robust, but coefficient of males turns to insignificant because of the large standard error.

[Insert Table 7 here]

From column 3 and 4 of Table 7, the schistosomiasis control program increased the number of births for females. Our results are inconsistent with Li and Wei (2017). They find the prevention and cure of schistosomiasis since 1950s had a significant positive impact on population growth, but the faster population growth only depended on lower mortality rather than fertility. They run regressions at the city level, while we use more micro-level data to show a positive relationship between the schistosomiasis control program and fertility.

D. Evidences from the CFPS

In this part, we use the CFPS data set to confirm what we have found in the census data. All regressions in Table 8 add the county fixed effects, cohort fixed effects, and individual controls that are identical to the counterparts using the census data. As is

[11] The results do not change too much if we control individual educational level in column 1-2, which reflects that besides the increase of education, other factors like health capital also help to find a job.

shown in column 1-4, the schistosomiasis control program increased educational attainment for both males and females in rural areas. The effects are more robust for females. Because early cohorts have retired in 2010, here we fail to replicate the results about working status. Although the coefficient in column 5 (numbers of birth) are not significant, it has the right sign and is not far apart from what we get in Table 7. This may be due to the small sample size.

[Insert Table 8 here]

Thanks to the more detailed information available in the CFPS, we can evaluate the effect of early life exposure to schistosomiasis on more adult outcomes. People make decisions to invest in education by comparing opportunity cost and payoff which disease control program could impact both, so previous studies find ambiguous evidence of the effect on education. From this perspective, health outcomes are better for observing the causal impact of disease eradication (Chang et al. 2011). We use two kinds of health outcomes in later life: adult height and cognitive functions. Adult height are considered to be a good proxy of past health conditions (Fogel 1986; Steckel 2009). From column 1 in Table 9, we see a significant increase of adult height for males, though the effect vanishes for females. Recent studies suggest that early childhood is a key period for brain development (Cunha and Heckman 2008). Column 2 and 3 show that schistosomiasis control in early life led to better cognitive functions. From Panel B, a 30 percentage point decrease in county-level schistosomiasis incidence is associated with about 2.9 and 3.4 higher scores in math and word tests for rural women, and the effect is smaller for men. Our results show that improvements in the early childhood disease environment may increase cognitive test scores in adulthood, providing an explanation for Flynn Effect (Flynn 1984).

In column 5, we examine the effect on economic status which here is measured by per capita household consumption expenditure. Since we are more interested in the effect of schistosomiasis control on productivity, an intuitive measure should be individual income. However, in agricultural households it is difficult to define individual income and separate individual contributions to household income (Liu et al. 2008). Given the unsuitability of individual income data, we use log per capita household expenditure, like Cutler et al. (2010). Based on our estimates, a reduction of infection rate from 30% to 0 would lead to an approximately 20 percent increase in per capita household expenditure for treated rural males, but we do not observe the same effect for treated females.[12] In addition, the effects are larger in magnitude for men than for women. The huge household consumption increase reveals a great gain in labor productivity. We expect this increase in economic status is due to a more prestigious job or higher productivity because of the human capital accumulation and better physical health induced by schistosomiasis reduction.

[Insert Table 9 here]

IV. Robustness Tests

In this section, we conduct several robustness tests for the results about educational attainment, using variants of equation (1). We first add interaction terms of control variables with the post-control dummy to allow their influences to vary across cohorts.[13] Column 1 of Table 10 implies that the baseline results survive this specification test. Our second exercise limits the sample to people born 10 years before (1947) and 8 years after the schistosomiasis control program, right before the start of Cultural Revolution (1965), making pre- and post-control cohorts more comparable. This sample has very similar results, as is shown in column 2 of Table 10. In column 3, we add five adjacent provinces in our original sample, and the results are robust to the inclusion of the areas without schistosomiasis.

If counties had different pre-trends of educational attainment, the increase in education attainment could have happened even in the absence of the disease control. To address this concern, we first add province×post fixed effects to equation (1). Next,

---

[12] This result pattern is also found in Cutler et al. (2010). Since women have lower participation rate in labor market, this result suggests that the effects for males may be attributed to improvements in productivity. See Cutler et al. (2010) for a more detailed discussion.

[13] Formally, we estimate $Y_{ijc} = \alpha + \beta(sch_j \times post_c) + X'_{ijc}\Theta + post_c \times X'_{ijc}\Gamma + \delta_j + \delta_c + \epsilon_{ijc}$.

we include county-specific time trends, allowing different countries have different growth paths. As a final specification check, we include interaction terms constructing as the average level of the dependent variable in 1920-1930 times post dummy ($\bar{Y}_{j,20-30} \times post_c$) as an additional covariate to address mean reversion concerns. The results are in column 4-6 of Table 10. Although some coefficients are insignificant, they all have the expected signs. Compared with rural male sample, the results for females are less sensitive to these specification checks.

[Insert Table 10 here]

We undertake two kinds of placebo tests as our final robustness checks. First, we re-estimate equation (1) in the urban sample. We expect that schistosomiasis control had little impact on urban residents because schistosomiasis is prevalent in rural areas where people lack safe water sources, while urban areas have much better sanitary conditions. Second, we create a fake intervention 10 years earlier in 1947 as an additional falsification check. Thus, we would expect to see a limited educational improvement. Results in column 7-8 of Table 10 confirm our hypothesis, as most coefficients are insignificant. A marginally significant effect appears for literacy rate and primary school education completion in rural females.

Admittedly, China experienced many significant events in our study period (e.g. the great famine, Cultural Revolution), so one may concern that the observed results can be explained by other programs, rather than schistosomiasis control campaign. However, in our DID setup only if other programs' intensity were correlated with pre-control prevalence could generate our results, which seems implausible. We examine one confounder that has salient impact on both education and health: the great famine.[14] The correlation between pre-control disease incidence and famine shock in county level is only 0.0007 and is statistically insignificant (p-value=0.98). One limitation of this study is that we fail to test the correlation of Cultural Revolution or Great Leap Forward

---

[14] We use the 1990 census data to construct the measurement of the great famine as follows: $famine_j = \frac{(cohort1_j - cohort2_j)}{cohort1_j} \times 100\%$, where $cohort1_j$ is the cohort size of people born during 1956-1958, $cohort2_j$ is the cohort size of people born during 1959-1961. A larger value means a severer famine impact.

with pre-control schistosomiasis prevalence, but it seems the natural determined disease environment was not associated with political movement intensity.

Selective mortality is another common threat to studies like this paper, but we argue that even with this kind of problem, our results at lease provide lower bounds for the effect of schistosomiasis. Typically, infectious-disease kills the least healthy people. Given that health is positively associated with educational attainment and productivity, our results would be downward biased when the schistosomiasis control program reduced mortality. This could also explain why we find there is no effect or even significantly negative effect on male literacy in some specifications.

V. Conclusions

This study exploits the schistosomiasis control campaign in China (circa 1957) to examine the impact of early life exposure to schistosomiasis on human capital accumulation and economic outcomes in adulthood. We find that the exogenous disease control campaign led to increases in years of schooling, and most of this increase comes from the junior school completion. This positive effect on educational attainment is more pronounced for rural females. However, schistosomiasis control had very limited effect on literacy rate for men in rural areas. The program also increased the possibility of having a job in adulthood, the number of births for females, men height, cognitive functions and later life economic status. Compared to the existing literature on malaria, our estimates imply a larger impact of disease control in China on educational attainment.[15]

Our paper also present clear policy implications for African countries that are still suffer from this disease. Compared with malaria, schistosomiasis is neglected and underestimated. However, migration and urbanization are spreading schistosomiasis to new areas. Our results provide support that schistosomiasis control can indeed lead to positive long-run human capital accumulation and socioeconomic improvements.

[15] Adrienne Lucas (2010) finds that reducing malaria incidence by 10 percentage points leads to an increase in completed schooling of about 0.1 years and an increase in the probability of being literate by 1 to 2 percentage point for females in Paraguay and Sri Lanka.

Although anti-schistosomiasis may not be a major engine of economic development, it is worthwhile to fight against this disease for the sake of future economic well-being.

REFERENCES

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6): 925-85.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy*, 114(4): 672–712.

Barker, D. J. P. 1990. "The Fetal And Infant Origins Of Adult Disease: The Womb May Be More Important Than The Home." *BMJ: British Medical Journal*, 301(6761): 1111.

Barofsky, Jeremy, T. D. Anekwe, and C. Chase. 2015. "Malaria eradication and economic outcomes in sub-Saharan Africa: Evidence from Uganda." *Journal of Health Economics,* 44(3): 118–136.

Barreca, Alan I. 2010. "The long-term economic impact of in utero and postnatal exposure to malaria." *Journal of Human Resources*, 45(4): 865–892.

Bleakley, Hoyt. 2010. "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure." *American Economic Journal: Applied Economics*, 2(2): 1-45.

Bloom, David E., and David Canning 2005. "Health and Economic Growth: Reconciling the Micro and Macro Evidence." Center on Democracy, Development, and the Rule of Law Working Paper 42.

Case, Anne, and Christina Paxson. 2009. "Early life health and cognitive function in old age." *The American Economic Review: Papers and Proceedings*, 99 (2): 104-109.

Chang, Simon, Belton Fleisher, Seonghoon Kim and Shi-yung Liu. 2011. "Long-term Effects of Early Childhood Malaria Exposure on Education and Health: Evidence from Colonial Taiwan." *Economic Development and Cultural Change*, 62(3): 519–536.

Chen, Yuyu, and Li'an Zhou. 2007. "The long-term health and economic consequences of the 1959–1961 famine in China." *Journal of Health Economics*, 26 (4): 659-681.

Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl. 2010. "Early-life malaria exposure and adult outcomes: evidence from malaria eradication in

India." *American Economic Journal: Applied Economics,* 2(2): 72–94.

Cunha, Flavio, Heckman James J. 2008. "Formulating, identifying and estimating the technology of cognitive and non cognitive skill formation." *Journal of Human Resources,* 43(4): 738–782.

Flynn, James R. 1984. "The mean IQ of Americans: massive gains 1932 to 1978." *Psychological Bulletin*, 95(1): 29–51.

Fogel, Robert. 1986. "Physical Growth as a Measure of the Economic Weil-Being of Populations: The Eighteenth and Nineteenth Centuries." In *Human Growth: A Comprehensive Treatise*, Vol. 3, edited by F. Falkner and J. M. Tanner, 263-81. New York: Plenum Press.

Gallup, John L., and Jeffrey D. Sachs. 2001. "The Economic Burden of Malaria." *American Journal of Tropical Medicine and Hygiene*, 64(1-2): 85-96.

Lee, Chulhee. 2014. "In Utero Exposure to the Korean War and Its Long-Term Effects on Socioeconomic and Health Outcomes." *Journal of Health Economics*, 33(1): 76–93.

Liu, Gordon. G., W. H. Dow, Alex Z. Fu, J. Akin, and P. Lance. 2008. "Income Productivity in China: On the Role of Health." *Journal of Health Economics*, 27(1): 27-44.

Li, Nan, and Xin Wei. 2017. "The Effect of Prevention and Cure of Schistosomiasis on Population Growth in China (1950-1990)." *Researches in Chinese Economic History,* 30(1): 84-95 (in Chinese).

Lucas, Adrienne M. 2010. "Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics,* 2(2): 46-71.

Mao, Shoupai, and Baoruo Shao. 1982. "Schistosomiasis control in the People's Republic of China." *The American Journal of Tropical Medicine and Hygiene,* 31(1): 92-99.

Meng, Xin, and Nancy Qian. 2009. "The long term consequences of famine on survivors: evidence from a unique natural experiment using China's great famine." NBER Working paper 14917.

Nour, N. M. 2010. "Schistosomiasis: health effects on women." *Reviews in Obstetrics & Gynecology*, 3(1): 28-32.

Qian, Xinzhong. 1988. *China Schistosomiasis Atlas*. Shanghai: Chinese Map Publishing House.

Steckel, Richard H. 2009. "Heights and Human Welfare: Recent Developments and New Directions." *Explorations in Economic History*, 46(1): 1-23.

Venkataramani, A.S. 2012. "Early life exposure to malaria and cognition in adulthood: evidence from Mexico." *Journal of Health Economics*, 31(5): 767–780.

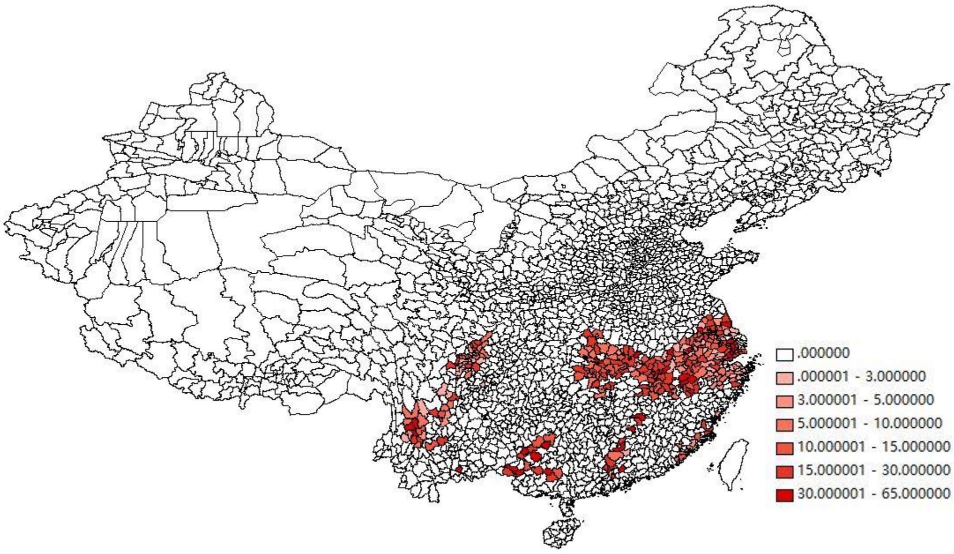
Weil, David N, 2007. "Accounts for the Effect of Health on Economic Growth." *Quarterly Journal of Economics*, 122(3): 1265-1306.

World Health Organization, 2016. 10 Facts about Schistosomiasis. http://www.who.int/features/factfiles/schistosomiasis/en/.

World Health Organization, 2017. Schistosomiasis Fact Sheet. http://www.who.int/mediacentre/factsheets/fs115/en/.

Yuan, Hongchang. 1989. "Epidemic situation and strategies for schistosomiasis control in marshland and lake regions." *Chinese Journal of Schistosomiasis Control,* 1(1): 2-6 (in Chinese).

Panel A. Pre-control



Panel B. Post-control

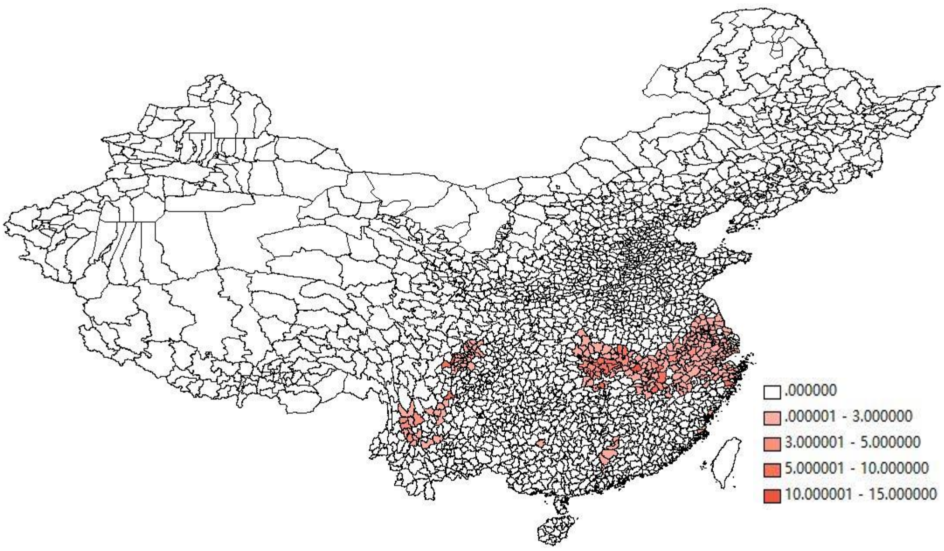


FIGURE 1. SCHISTOSOMIASIS ENDEMICITY MAP

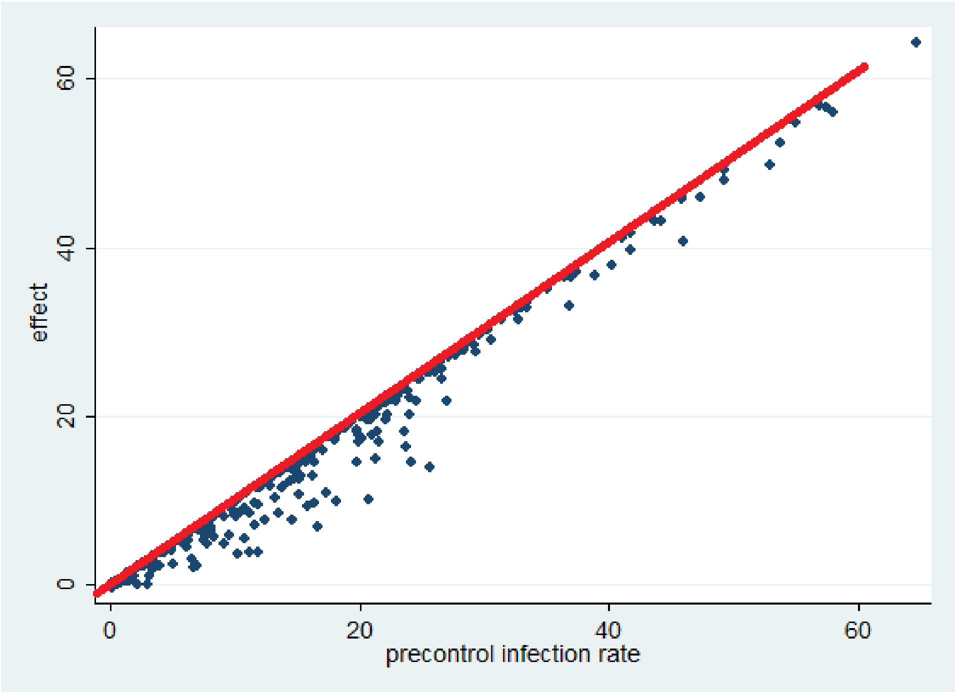


FIGURE 2. PRE-CONTROL INFECTION RATES AND CONTROL PROGRAM EFFECT

Notes: The y axis displays the effect of schistosomiasis control campaign measured by pre-control infection rates minus post-control infection rates. The x axis is the pre-control schistosomiasis prevalence. The red line is the diagonal.

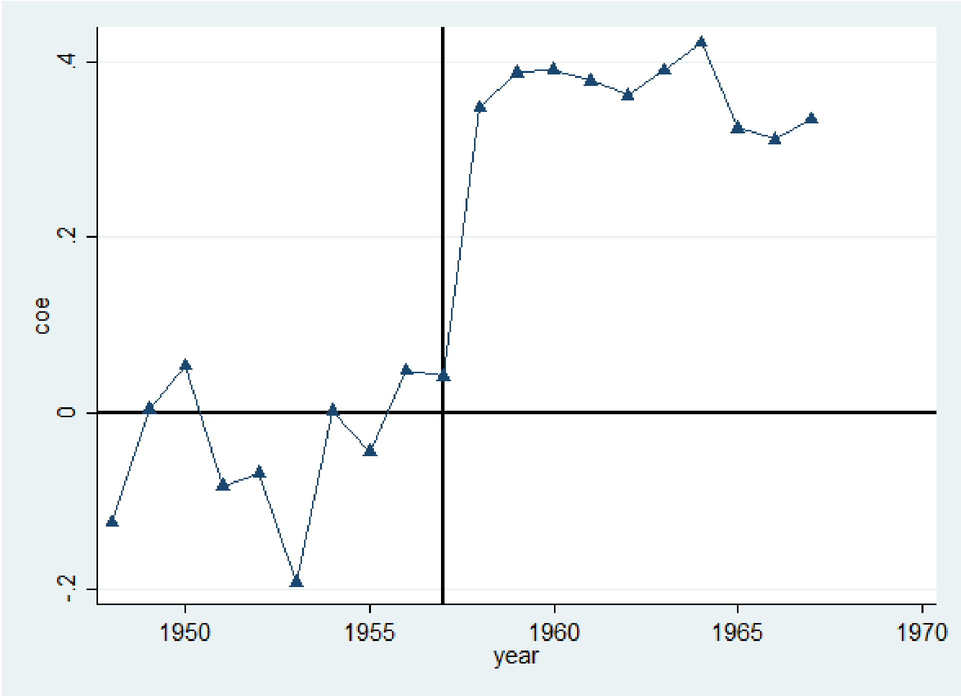


FIGURE 3. COHORT ANALYSIS

Notes: Plots of coefficients on $sch_j \times \delta_c$ in equation (3) for rural females. Dependent variable is a dummy of junior school (=1 if completed).

TABLE 1. SUMMARY STATISTICS OF THE CENSUS DATA

	Male		Female	
	Mean	SD	Mean	SD
Age	34.02	11.92	33.90	11.95
Married	0.70	0.46	0.80	0.40
Ethnicity (Han)	0.92	0.27	0.92	0.28
Schooling Years	6.34	3.26	4.40	3.59
Literate	0.89	0.31	0.69	0.46
Primary School	0.75	0.43	0.53	0.50
Junior School	0.36	0.48	0.20	0.40
At Work	0.97	0.18	0.90	0.31
Number of Births			2.37	2.07
Number of Observations	1,253,044		1,212,189	
Number of Counties	858			
Number of Provinces	12			

Notes: The sample includes people with rural hukou and born during 1930-1972 in 12 endemic provinces (Jiangsu, Zhejiang, Anhui, Jiangxi, Hunan, Hubei, Yunnan, Sichuan, Fujian, Guangdong, Shanghai and Guangxi).

TABLE 2. SUMMARY STATISTICS OF THE CFPS DATA

	Male		Female	
	Mean	SD	Mean	SD
Age	53.66	10.66	52.95	10.50
Married	0.91	0.29	0.89	0.31
Ethnicity (Han)	0.92	0.27	0.92	0.27
Schooling Years	6.31	4.12	3.93	4.20
Literate	0.77	0.42	0.52	0.50
Primary School	0.71	0.45	0.46	0.50
Junior School	0.43	0.50	0.24	0.42
Number of Births			2.45	1.25
Height	167.47	6.42	157.97	6.19
Math Test Score	8.89	5.87	5.53	5.61
Word Test Score	15.53	9.83	9.59	9.89
Log per capital Household Expenditure	8.52	0.84	8.54	0.85
Number of Observations	7,191		6,649	
Number of Counties	117			
Number of Provinces	25			

Notes: The sample includes people with rural hukou at 3 years old and born during 1930-1972 in 12 endemic provinces (Jiangsu, Zhejiang, Anhui, Jiangxi, Hunan, Hubei, Yunnan, Sichuan, Fujian, Guangdong, Shanghai and Guangxi) plus 13 uninfected provinces (Beijing, Tianjin, Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Shandong, Henan, Chongqing, Guizhou, Shaanxi, and Gansu).

TABLE 3. DIFFERENCE IN MEANS

	High Endemic			Low Endemic			
	Pre	Post	Increase	Pre	Post	Increase	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Rural Male							
Years of Schooling	5.82	8.19	2.37	5.77	8.11	2.34	0.03
Literate	0.85	0.98	0.13	0.85	0.97	0.12	0.01
Primary School	0.65	0.89	0.24	0.65	0.89	0.24	0.00
Junior School	0.29	0.60	0.31	0.28	0.58	0.30	0.01
Panel B. Rural Female							
Years of Schooling	3.17	6.67	3.50	3.47	6.86	3.39	0.11
Literate	0.54	0.90	0.36	0.58	0.90	0.32	0.04
Primary School	0.35	0.74	0.39	0.39	0.77	0.38	0.01
Junior School	0.12	0.43	0.31	0.14	0.45	0.31	0.00

Notes: We use pre-control schistosomiasis infection rates to classify endemic counties into highly endemic or less endemic groups. sch_j between 0 and 15 are classified as low epidemic, and sch_j over 15 are classified as highly epidemic. Cohorts born during 1930-1956 are the pre-control group, whereas those born from 1957-1972 are the post-control group. Column 3 is the difference between column 1 and 2. Column 6 is the difference between column 4 and 5. Column 7 is the difference between column 3 and 6.

TABLE 4. EFFECT ON EDUCATIONAL ATTAINMENT

Dependent Variable	Schooling Years		Literate		Primary School		Junior School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rural Male								
$sch_j \times post_c$	1.350***	1.132***	-0.016	-0.051	0.186***	0.141***	0.191***	0.213***
	(0.275)	(0.367)	(0.036)	(0.032)	(0.042)	(0.050)	(0.038)	(0.057)
Observations	1,253,044	332,668	1,253,044	332,668	1,253,044	332,668	1,253,044	332,668
Panel B. Rural Female								
$sch_j \times post_c$	2.188***	3.229***	0.173***	0.139*	0.247***	0.356***	0.233***	0.471***
	(0.420)	(0.568)	(0.048)	(0.072)	(0.051)	(0.085)	(0.065)	(0.081)
Observations	1,212,189	249,755	1,212,189	249,755	1,212,189	249,755	1,212,189	249,755
County Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Individual Controls	NO	YES	NO	YES	NO	YES	NO	YES

Notes: Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 5. INSTRUMENTAL VARIABLES ESTIMATE

Dependent Variable	Schooling Years	Literate	Primary School	Junior School
	(1)	(2)	(3)	(4)
Panel A. Rural Male				
$water_j \times post_c$	4.490***	0.133	0.570***	0.718***
	(1.469)	(0.155)	(0.203)	(0.182)
K-P F Statistic	37.682	37.682	37.682	37.682
Observations	331,784	331,784	331,784	331,784
Panel B. Rural Female				
$water_j \times post_c$	7.704***	0.518**	1.060***	0.852***
	(1.920)	(0.227)	(0.307)	(0.225)
K-P F Statistic	37.792	37.792	37.792	37.792
Observations	249,197	249,197	249,197	249,197
County Fixed Effects	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES

Notes: $water_j$ is the proportion of water area in county *j*, instrumenting for pre-control schistosomiasis infection rates. All regressions are estimated by 2SLS. Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 6. CHILDHOOD EXPOSURE

Dependent Variable	Schooling Years	Literate	Primary School	Junior School
	(1)	(2)	(3)	(4)
Panel A. Rural Male				
$sch_j \times exposure_c$	2.100***	-0.116	0.308***	0.337***
	(0.709)	(0.073)	(0.101)	(0.099)
Observations	332,668	332,668	332,668	332,668
Panel B. Rural Female				
$sch_j \times exposure_c$	7.347***	0.397**	0.875***	1.000***
	(1.345)	(0.190)	(0.195)	(0.177)
Observations	249,755	249,755	249,755	249,755
County Fixed Effects	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES

Notes: $exposure_c$ is the percentage of years prior to age 18 that are spent in the post-control period. Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 7. EFFECT ON OTHER LATER LIFE OUTCOMES

Dependent Variable	At Work		Number of Births	
	(1)OLS	(2)IV	(3)OLS	(4)IV
Panel A. Rural Male				
$sch_j \times post_c$	0.015**	0.007		
	(0.007)	(0.032)		
K-P F Statistic		23.010		
Observations	1,253,044	1,249,598		
Panel B. Rural Female				
$sch_j \times post_c$	0.074**	0.198*	0.678***	2.629***
	(0.037)	(0.111)	(0.202)	(0.868)
K-P F Statistic		20.321		20.321
Observations	1,212,189	1,209,023	1,212,189	1,209,023
County Fixed Effects	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES

Notes: Dependent variable in column 1-2 is a dummy for working status (=1 if have a job). Individual controls in column 1-2 are ethnicity and marital status, and we also add individual educational years in column 3-4. Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 8. RESULTS FROM CFPS

Dependent Variable	Schooling Years	Literate	Primary School	Junior School	Number of Births
	(1)	(2)	(3)	(4)	(5)
Panel A. Rural Male					
$sch_j \times post_c$	3.748*	0.011	0.375	0.849***	
	(1.963)	(0.254)	(0.233)	(0.155)	
Observations	4,153	4,153	4,153	4,153	
Panel B. Rural Female					
$sch_j \times post_c$	10.250***	0.721***	1.020***	1.246***	0.327
	(1.884)	(0.235)	(0.188)	(0.297)	(0.371)
Observations	3,826	3,826	3,826	3,826	3,826
County Fixed Effects	YES	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES	YES

Notes: Individual controls in column 1-4 are educational level of parents and ethnicity, in column 5 are years of education, ethnicity, and marital status. Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 9. OTHER LATER LIFE OUTCOMES FROM CFPS

Dependent Variable	Height	Math Test	Word Test	Consumption
	(1)	(2)	(3)	(4)
Panel A. Rural Male				
$sch_j \times post_c$	4.769*	4.007*	7.212*	0.683*
	(2.519)	(2.392)	(3.748)	(0.371)
Observations	4,095	4153	4153	3726
Panel B. Rural Female				
$sch_j \times post_c$	2.836	9.550***	11.27***	-0.005
	(2.511)	(2.641)	(2.924)	(0.383)
Observations	3,624	3826	3826	3368
County Fixed Effects	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES

Notes: Dependent variable in column 4 is log per capita household consumption expenditure. Individual controls are educational level of parents and ethnicity. Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE 10. ROBUSTNESS CHECKS

	Add More Controls	Cohorts 1947-1965	Add Adjacent Provinces	Province-Post Fixed Effects	Regional Trends	Mean Reversion	Urban Sample	Fake Interventio n:1947 (1930-1957 Cohorts)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rural Male								
Schooling Years	1.102***	1.258***	1.026***	0.519	0.972**	1.080***	0.477	0.366
	(0.343)	(0.368)	(0.348)	(0.338)	(0.376)	(0.351)	(0.840)	(0.655)
Literate	-0.048	-0.048	-0.117***	-0.087***	-0.025	-0.048*	0.022	0.043
	(0.031)	(0.031)	(0.032)	(0.032)	(0.027)	(0.028)	(0.044)	(0.077)
Primary School	0.141***	0.104**	0.149***	0.061	0.037	0.096**	0.025	0.115
	(0.047)	(0.052)	(0.047)	(0.045)	(0.048)	(0.046)	(0.055)	(0.094)
Junior School	0.204***	0.223***	0.198***	0.136***	0.079	0.213***	-0.057	-0.090
	(0.051)	(0.054)	(0.055)	(0.049)	(0.061)	(0.057)	(0.108)	(0.082)
Observations	332,668	92,875	461,247	332,668	332,668	332,668	27,254	29,040
Panel B. Rural Female								
Schooling Years	3.172***	2.629***	2.723***	1.445***	1.708***	3.152***	0.121	1.593
	(0.556)	(0.551)	(0.544)	(0.536)	(0.575)	(0.543)	(1.453)	(1.220)
Literate	0.123*	0.098	0.062	0.048	0.169***	0.140**	0.020	0.329*
	(0.070)	(0.067)	(0.070)	(0.071)	(0.062)	(0.063)	(0.091)	(0.193)
Primary School	0.346***	0.284***	0.348***	0.127*	0.213***	0.300***	0.106	0.288*
	(0.083)	(0.083)	(0.083)	(0.075)	(0.080)	(0.081)	(0.136)	(0.174)
Junior School	0.476***	0.388***	0.367***	0.238***	0.074	0.471***	-0.028	-0.043
	(0.078)	(0.080)	(0.078)	(0.068)	(0.093)	(0.081)	(0.187)	(0.107)
Observations	249,755	48,636	355,826	249,755	249,755	249,755	22,503	10,316

Notes: In column 1, we interact controls with the post-control dummy. Adjacent provinces are Henan, Shandong, Shanxi, Guizhou and Gansu. Column 3 controls province×post fixed effects. Column 4 includes county-specific trends as additional controls. In column 6, we add an interaction term with county average of the dependent variables during 1920-1930 as an additional control. All regressions add county fixed effects, cohort fixed effects and individual controls. Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

APPENDIX

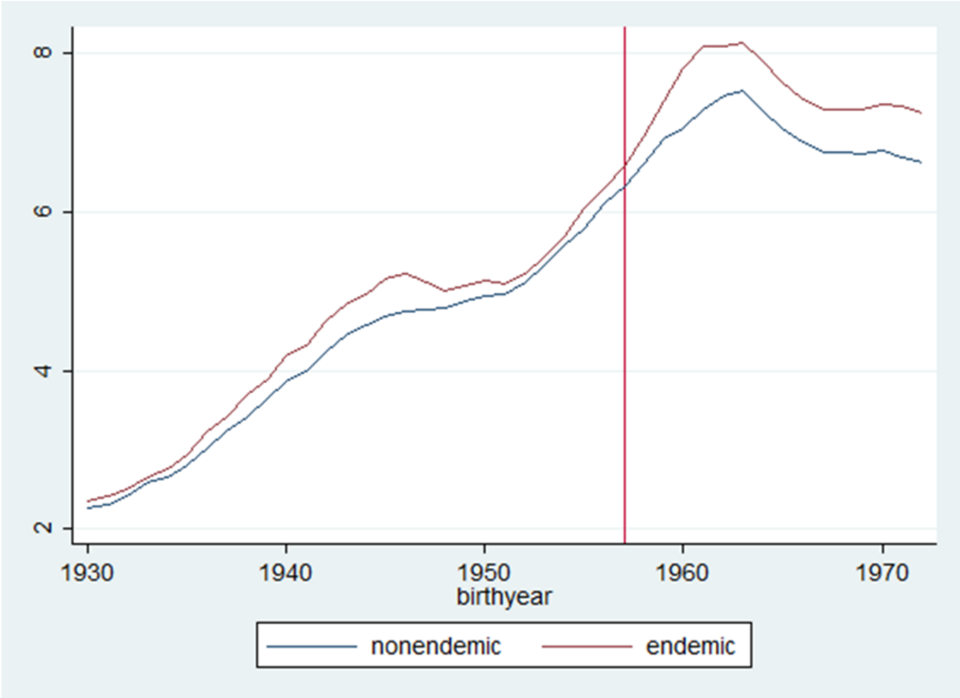


FIGURE A1. Years of Schooling in each cohort

Notes: This figure shows the educational years of different cohorts in endemic counties and those free of schistosomiasis.

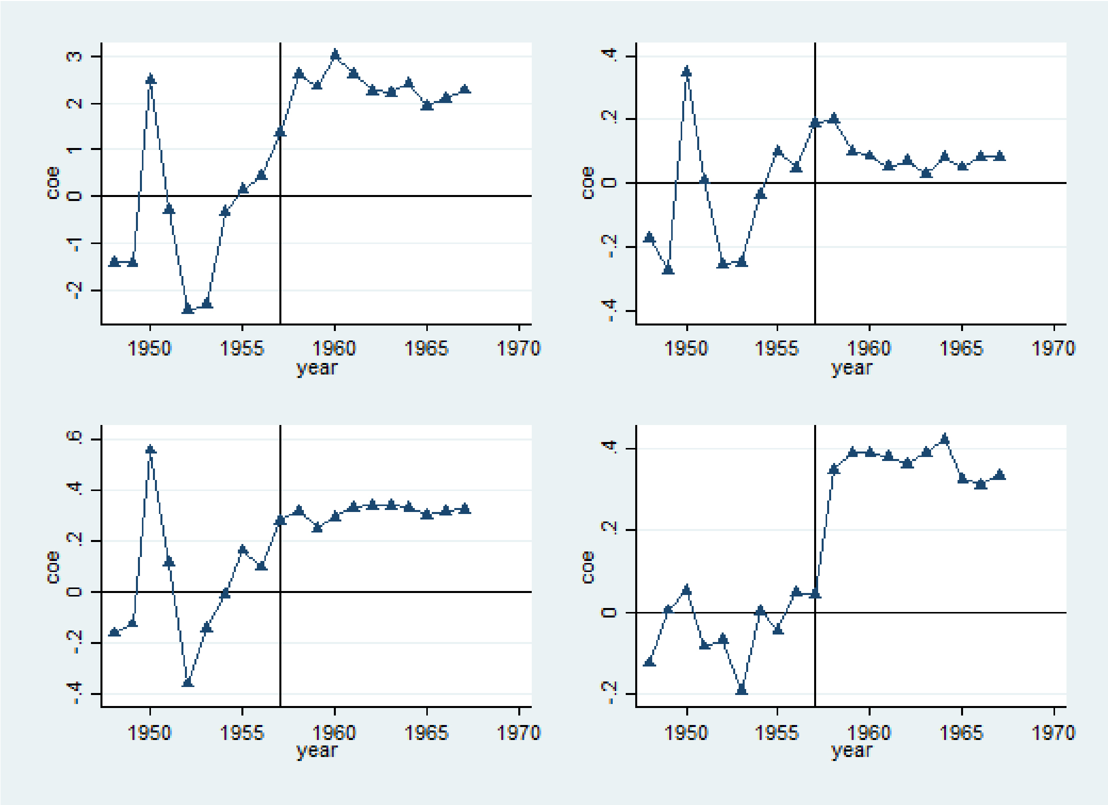


FIGURE A2. COHORT ANALYSIS OF FEMALES

Notes: Plots of coefficients on $sch_j \times \delta_c$ in equation (3) for rural females. Dependent variable in the upper-left is years of education, in the upper-right a dummy of literate, in the lower-left a dummy of primary school (=1 if completed), in the lower-right a dummy of junior school (=1 if completed).

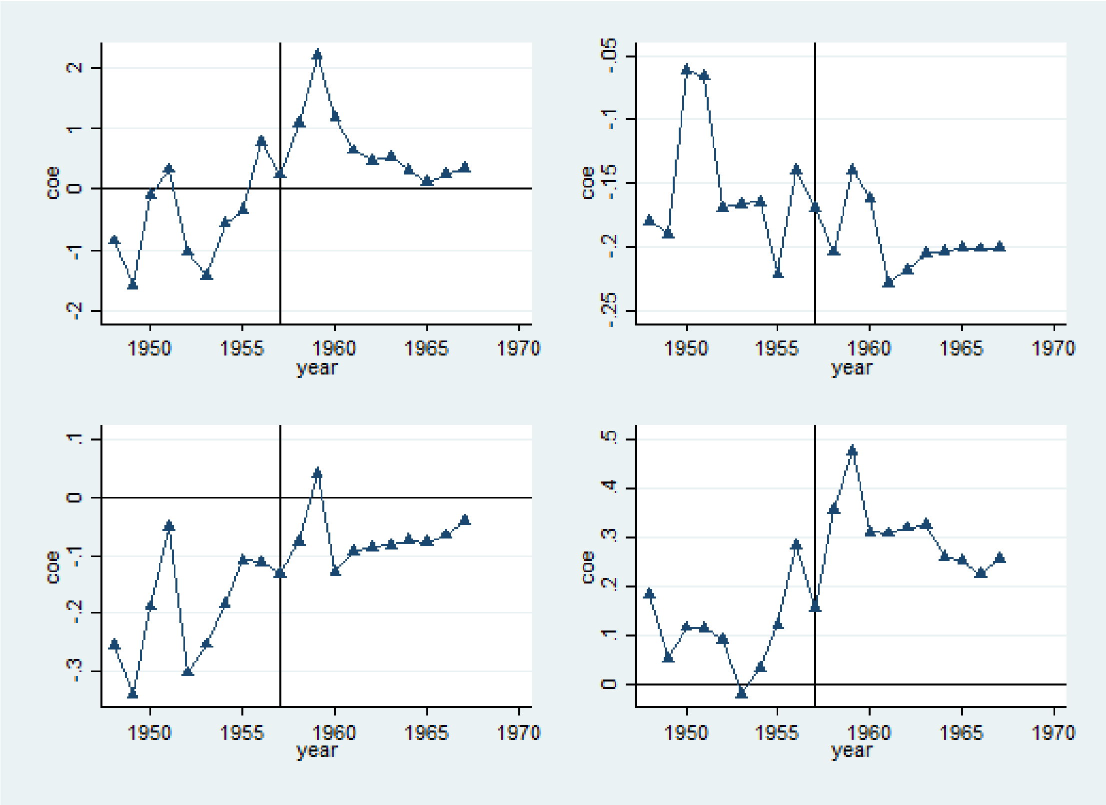


FIGURE A3. COHORT ANALYSIS OF MALES

Notes: Plots of coefficients on $sch_j \times \delta_c$ in equation (3) for rural males. Dependent variable in the upper-left is years of education, in the upper-right a dummy of literate, in the lower-left a dummy of primary school (=1 if completed), in the lower-right a dummy of junior school (=1 if completed).

TABLE A1. BASELINE RESULTS USING A CATEGORICAL CLASSIFICATION

Dependent Variable	Schooling Years		Literate		Primary School		Junior School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rural Male								
$low_j \times post_c$	0.459***	0.501***	-0.004	-0.005	0.057***	0.050***	0.070***	0.081***
	(0.065)	(0.095)	(0.010)	(0.010)	(0.009)	(0.013)	(0.010)	(0.014)
$high_j \times post_c$	0.422***	0.316***	-0.007	-0.020**	0.0579***	0.0413***	0.0611***	0.0611***
	(0.080)	(0.101)	(0.009)	(0.008)	(0.012)	(0.014)	(0.012)	(0.016)
Observations	1,253,044	332,668	1,253,044	332,668	1,253,044	332,668	1,253,044	332,668
Panel B. Rural Female								
$low_j \times post_c$	0.605***	0.812***	0.011	-0.004	0.064***	0.096***	0.096***	0.148***
	(0.099)	(0.169)	(0.012)	(0.02)	(0.012)	(0.027)	(0.016)	(0.019)
$high_j \times post_c$	0.639***	0.900***	0.046***	0.023	0.0694***	0.104***	0.0726***	0.137***
	(0.138)	(0.161)	(0.015)	(0.020)	(0.017)	(0.025)	(0.020)	(0.021)
Observations	1,212,189	249,755	1,212,189	249,755	1,212,189	249,755	1,212,189	249,755
County Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Individual Controls	NO	YES	NO	YES	NO	YES	NO	YES

Notes: Pre-control infection rates that equal to zero are classified as non-epidemic, sch_j between 0 and 15 are classified as low epidemic (low_j), and sch_j over 15 are classified as high epidemic ($high_j$). $non_j \times post_c$ is omitted as a reference group. Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE A2. BASELINE RESULTS IN AN ERADICATED SAMPLE (BELOW 0.5%)

Dependent Variable	Schooling Years		Literate		Primary School		Junior School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rural Male								
$sch_j \times post_c$	1.005**	1.094**	-0.098**	-0.102***	0.114*	0.127*	0.221***	0.250***
	(0.464)	(0.556)	(0.045)	(0.038)	(0.068)	(0.068)	(0.065)	(0.091)
Observations	1,076,417	287,039	1,076,417	287,039	1,076,417	287,039	1,076,417	287,039
Panel B. Rural Female								
$sch_j \times post_c$	2.531***	3.531***	0.055	0.010	0.282***	0.376***	0.396***	0.635***
	(0.593)	(0.765)	(0.063)	(0.083)	(0.078)	(0.116)	(0.086)	(0.106)
Observations	1,039,072	216,547	1,039,072	216,547	1,039,072	216,547	1,039,072	216,547
County Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Individual Controls	NO	YES	NO	YES	NO	YES	NO	YES

Notes: The eradicated sample includes control counties and the eradicated counties that have very low post-control prevalence (below 0.5%). Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE A3. BASELINE RESULTS IN AN ERADICATED SAMPLE (BELOW 1%)

Dependent Variable	Schooling Years		Literate		Primary School		Junior School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rural Male								
$sch_j \times post_c$	1.216***	1.255***	-0.051	-0.083**	0.152***	0.150**	0.227***	0.268***
	(0.351)	(0.456)	(0.046)	(0.036)	(0.054)	(0.061)	(0.048)	(0.070)
Observations	1,125,297	300,428	1,125,297	300,428	1,125,297	300,428	1,125,297	300,428
Panel B. Rural Female								
$sch_j \times post_c$	2.435***	3.510***	0.122**	0.054	0.286***	0.378***	0.314***	0.599***
	(0.467)	(0.680)	(0.054)	(0.074)	(0.058)	(0.105)	(0.082)	(0.092)
Observations	1,087,221	226,774	1,087,221	226,774	1,087,221	226,774	1,087,221	226,774
County Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Cohort Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Individual Controls	NO	YES	NO	YES	NO	YES	NO	YES

Notes: The eradicated sample includes control counties and the eradicated counties that have very low post-control prevalence (below 1%). Individual controls are educational level of parents and ethnicity (=1 if belongs to Han ethnicity). Standard errors clustered at county level appear in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.