Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure[*]

Hoyt Bleakley[†]

Hoyt Bleakley: chicagogsb.edu

[†]Associate Professor of Economics, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL, 60637. Telephone: (773) 834-2192

Abstract

This study uses the malaria-eradication campaigns in the United States (circa 1920), and in Brazil, Colombia and Mexico (circa 1955) to measure how much childhood exposure to malaria depresses labor productivity. The campaigns began because of advances in health technology, which mitigates concerns about reverse causality. Malarious areas saw large drops in the disease thereafter. Relative to non-malarious areas, cohorts born after eradication had higher income as adults than the preceding generation. These cross-cohort changes coincided with childhood exposure to the campaigns rather than to pre-existing trends. Estimates suggest a substantial, though not predominant, role for malaria in explaining cross-region differences in income.

Keywords

1 Introduction

The disease known as malaria, an affliction of mankind through recorded history, persists in tropical regions up to the present day. These same tropical areas have, generally speaking, a much lower level of economic development than that enjoyed in the temperate climates. These facts lead us to a natural question: does malaria hold back economic progress? Unfortunately, simple correlations between tropical disease and productivity cannot answer this question. Malaria might depress productivity, but the failure to eradicate malaria might equally well be a symptom of underdevelopment, itself caused by poor institutions or bad luck.[1] How can we disentangle this circular causality? The standard econometric answer is

to consider plausibly exogenous variation in malaria. A possible source of such variation comes from targeted interventions in public health.

The present study considers two major attempts to eradicate malaria in the Americas during the Twentieth century. The first episode analyzed took place in the southern US, largely in the 1920s. In the decades before, the cause and transmission mechanism of the disease were first understood by European physicians, and this knowledge allowed the US Army to attempt scientifically based campaigns against malaria in areas it had recently occupied: Havana and the Panama Canal Zone. Subsequently, this wealth of new knowledge and experience related to disease was applied to the malaria problem in the South. The second episode is the worldwide malaria eradication campaign, and in particular as it was implemented in Brazil, Colombia, and Mexico (starting in the 1950s). The efforts to eradicate malaria worldwide were spurred on by the discovery of DDT, a powerful pesticide. After World War II, the World Health Organization (WHO) helped many afflicted countries put together programs of spraying to combat malaria transmission. The campaigns in these regions partially interrupted the malaria transmission cycle and brought about marked drops in infection in a relatively short period of time. (Further background on the disease and the eradication efforts is found in Section 2.) Sufficient time has passed that we can evaluate the long-term consequences of these eradication campaigns.

The relatively rapid impact of the treatment campaigns combine with cross-area heterogeneity to form the research design of the present study. These four countries are geographically variegated, such that, within each country, some regions have climates that support malaria transmissions, while other regions do not. Areas with high malaria infection rates had more to gain from eradication, but the non-malarious areas serve as a comparison group, filtering out common trends in national policy, for example. The reductions in disease burden occur in the space of a few years, allowing me to identifying likely childhood exposure with sufficient precision. Moreover, these episodes of rapid progress against malaria were made possible by critical innovations to knowledge and spending, which are shocks that came from outside the studied areas. This latter fact mitigates the usual concern about policy endogeneity and reverse causality.

The particular goal of the present study is to identify the effect that childhood exposure to malaria eradication has on subsequent labor productivity as an adult. While direct effects of malaria on adults can be partially measured with lost wages from work absences, little is known about effects that persist from infection in early life. Why would one expect childhood malaria to have an impact that is potentially larger than the direct effect of lost work time during adulthood? To begin with, children are more susceptible to malaria than adults, most likely because prolonged exposure to the disease brings some degree of resistance. Although partial immunity is conferred by age, the damage from childhood exposure to malaria may be hard to undo: Most of a person's human-capital and physiological development happens in childhood. On the physiological side, a malaria-free childhood might mean that the individual is more robust as an adult, with concomitant increases in labor supply. On the human-capital side, fewer school absences and less anemia translates into more learning. This would be manifested in the data as greater literacy, higher

---

[1]For instance, see McCarthy, Wolf, and Wu (2000), Gallup and Sachs (2001), and Gallup, Sachs, and Mellinger (2001) for cross-country evidence on malaria and income, or Bloom, Canning, and Sevilla (2004), Lorentzen, McMillan, and Wacziarg (2005), Acemoglu and Johnson (2006), and Weil (2007) for evidence on the broader health/income link across countries. To say the least, a consensus on the magnitude of health impacts on income per capita has not emerged from this literature. Estimates range from the modestly negative to the large and positive. A common assumption in the latter set of studies is to treat health as a single index that is well proxied by life expectancy for the purposes of productivity impacts. This assumption is hard to sustain on theoretical grounds (see Meltzer, 1992, for example) and is therefore likely to be a source of misspecification. The former set of studies, on the other hand, treat only malaria, but use cross-section or time-series variations, which are unlikely to identify the effect of malaria in the face of numerous omitted correlates of the disease.

adult earnings, and, for a fixed time in school, higher returns to schooling. This also affects the schooling decision, but, because malaria also affects the childhood wage (the opportunity cost of schooling), this latter effect is ambiguously signed by theory. On the other hand, malaria eradication would have also reduced adult mortality, effectively extending the time during which educational investments could be utilized. Additionally, early adulthood (the steep part of the age-earnings profile) is a time of human-capital development that might be hindered by malaria. Finally, malaria's possible effect on contemporaneous wages implies that an additional channel is via the impact of parental income on a child's development.

To identify the effects of early-life malaria infection, I begin by noting that the timing of the eradication campaign induces variation in childhood malaria that has a marked pattern across year-of-birth cohorts. Cohorts that were already adults before the campaigns were too old to have any early-life exposure to the eradication efforts. In contrast, later cohorts experienced reduced malaria infection during their childhood. I therefore compare cohorts based on (i) the pre-eradication malaria burden in their place of birth and (ii) their year of birth relative to the malaria-eradication campaigns. Being born later *and* in an area with high pre-campaign malaria burden implies more exposure to the benefits of the eradication programs. To test this hypothesis, census microdata samples of native-born males from all four countries are used to construct panel data of cohorts by birth year and birthplace. I construct a year-of-birth-varying index of childhood (temporal) exposure to the eradication efforts, which I then interact with pre-campaign malaria intensity by place of birth. This interaction forms the central explanatory variable of the present study. (The identification strategy is further described in Section 2.4, while the data construction, including the various estimates of malaria intensity, is detailed in Section 3 and in the appendices.)

The main result of the study is that cohorts more exposed to the eradication efforts as children went on to earn higher incomes as adults. In Section 4, I examine long differences across cohorts by comparing two groups—those born well before and those born just after the campaign. This is therefore a simple pre/post comparison, with less-malarial areas serving as a comparison group. I show that the cross-cohort change in income was higher in areas that had higher pre-campaign burdens of malaria. I interpret this as being due to childhood exposure to malaria (or lack thereof). I also find that cohorts with less childhood exposure to malaria have higher literacy rates, but results are mixed for years of schooling. These results are not sensitive to accounting for a variety of alternative hypotheses. First, scatterplots show that this shift was systematic across areas of these countries, and not due to a few outliers. Second, I obtain essentially similar estimates of malaria coefficients even when controlling for different indicators of health and economic development, or when using corrections for migration or measurement error. Third, I present evidence (in Section 6.2) that these results are not due to mortality selection, the failure to account for general-equilibrium spillovers across cohorts, or the decline of other vector-borne diseases.

The shift in the relationship between income and pre-campaign malaria coincides with childhood exposure to the eradication efforts. I present evidence on this point in in Section 5, where I examine cohort data disaggregated by year of birth. Regression analysis of these data generally support the childhood-exposure hypothesis, even when controlling for trends at the cohort or time level. Furthermore, graphs of the cohort-specific coefficients are also consistent with a model in which childhood exposure to malaria depresses income later in life. First, for cohorts that were already adults by the time of the eradication efforts, the malaria-income coefficients are generally negative and do not exhibit a pronounced trend. Second, these coefficients begin to shift upwards for those cohorts born late enough to be partially exposed to eradication campaign during childhood. Third, for those cohorts born after the onset of eradication efforts, these malaria-income estimates are roughly stable. This

pattern of estimates—which is visible to varying degrees in all four countries studied—suggests that the rise in income across cohorts following eradication was indeed due to reduced childhood exposure to malaria rather than because of pre-existing trends or convergence.

The estimates below indicate that childhood malaria has a large, depressing effect on adult productivity. Reduced-form effects on income, when comparing the least malarious to the most malarious areas within a country, are on the range of 12 (in the US) to 40 percent (in Latin America). To get a sense of the magnitude of the effect per probability of childhood infection, I normalize the reduced-form numbers with estimates of the pre-campaign infection rates in Section 6.1. Although it is impossible to completely rule out that the intervention had effects through channels besides estimated malaria infection, the results that suggest that *persistent* childhood malaria infection reduces adult income around fifty percent. This estimate is similar across the four countries considered, which is noteworthy given the heterogeneity of income levels and institutions across the sample. In Section 6.4, I show that these estimates are factor of four smaller than some cross-country estimates, but nevertheless explain a modest fraction of the gap between the areas under study and more developed economies. Section 7 concludes the study.

2 Malaria and the Eradication Campaigns

2.1 The Disease

Malaria is a parasitic disease that afflicts humans. Acute symptoms of infection include fever, headache, and nausea. An important chronic symptom is anemia. Malaria results in death in some cases, but the strains prevalent in the Americas (*vivax* and to a lesser extent *malariae*) have quite low case-fatality rates, especially compared with the predominantly African variety (*falciparum*).

The parasite has a complicated life-cycle that is partly spent in a mosquito vector and partly in the human host. The disease is transmitted when a mosquito takes a blood meal from an infected person and, some time later, bites another person. Because of the crucial role played by mosquitoes in the transmission cycle, warm and wetter climates are more likely to sustain endemic malaria.

2.2 Efforts against Malaria in the Southern US, circa 1920

The turn of the 20th century saw considerable advances in the scientific understanding of the disease. Dr. Charles Louis Alphonse Laveran, of the French army, showed in the early 1890s through microscopic studies that malaria is caused by a single-celled organism. Dr. Ronald Ross, of the British Indian Medical Service, discovered in the late 1890s that malaria is transmitted via mosquitoes. These discoveries proved invaluable to addressing the malaria problem in a scientific and systematic way, and both men later won Nobel Prizes for Medicine. But a workable package of techniques for malaria control was still a ways off.

The US government's interest in vector-borne diseases arose in the 20th century, not because of a new-found interest in the Southern region, but because of its involvement in Cuba and in the Panama Canal Zone. Early in the occupation of Cuba, the US Army dispatched a team of physicians, among them Dr. Walter Reed, to Havana to combat yellow fever and malaria. Armed with the new knowledge about these diseases (from Laveran, Ross, and others), the Army was able to bring these diseases under control in that city. Another team of American physicians, this time led by Dr. William Gorgas, was able to bring these diseases under control in the Canal Zone, which was a considerable challenge given that much of the area was a humid, tropical jungle.

The knowledge generated by US Army doctors on malaria control in Cuba and Panama inspired work back home in the South in the latter half of the 1910s. Several physicians in the United States Public Health Service (PHS) began collecting information on the distribution of malaria throughout the South and the prevalence of the various species of parasites and mosquitoes.[2] The PHS began actual treatment campaigns in a limited way, first by controlling malaria in a handful of mill villages. The Rockefeller Foundation, having mounted a successful campaign against hookworm in the early 1910s, also funded anti-malarial work later in the decade through its International Health Board (IHB). These two groups sponsored demonstration projects in a number of small, rural towns across the South. They employed a variety of new methods (spraying of larvacides, water management, window screening, and mass administration of quinine) and most of these demonstrations were highly successful, resulting in 70% declines in morbidity (i.e., sickness that does not result in mortality).

The federal government's large-scale efforts against malaria in the South began with World War I (WWI). In previous wars, a significant portion of the troops were made unfit for service because of disease contracted in or around encampments. The PHS, working now with both a strong knowledge base on malaria control and greatly increased funding, undertook drainage and larviciding operations in Southern military camps as well as in surrounding areas. After the War, the IHB and PHS expanded the demonstration work further. By the mid-1920s, the boards of health of each state, following the IHB/PHS model, had taken up the mantle of the malaria control in all but the most peripheral areas of the region (Williams, 1951). The South experienced a drop in malaria mortality of more than 60 percent in the decade of the 1920s.

2.3 The Worldwide Campaign to Eradicate Malaria, circa 1950

While some of the innovations in malaria control diffused to less-developed regions, the tropical countries of the Americas would wait for further technological advance before launching serious campaigns against malaria.[3] These campaigns had a peculiar starting point: In 1941, a Swiss chemist seeking to build a better mothball re-discovered a chemical known today as DDT (short for dichloro-dipenyl-trichloro-ethane). Early tests showed this new chemical to be of extraordinary value as a pesticide: it rapidly killed a variety of insects and had no immediately apparent effects on mammals. DDT proved enormously valuable to the Allied war and occupation efforts in combating typhus and later malaria.[4] The United Nations Reconstruction and Relief Agency used DDT in the late 1940s to essentially eradicate malaria from Sardinia in the lapse of a few years.

The World Health Organization (WHO) proposed in the early 1950s a worldwide campaign to eradicate malaria. While the WHO mostly provided technical assistance and moral suasion, substantial funding came from USAID and UNICEF. The nations of Latin America took up this task in the 1950s. While individual nations had formal control of the design and implementation of the programs, their activities were comparatively homogeneous as per the dictates of their international funders. The central component of these programs was the spraying of DDT, principally in the walls of houses. Its purpose was not to kill every mosquito in the land, but rather to interrupt the transmission of malaria for long enough that

---

[2]Williams (1951) presents a thorough history of the US Public Health Service. Humphreys (2001) summarizes the history of malaria-control efforts in the United States. The annual reports of the Rockefeller Foundation's International Health Board (1919) provide information about its anti-malaria demonstration projects. Much of the historical detail for the US is drawn from these sources.
[3]The historical narrative on the worldwide campaign is drawn from Harrison (1978). Country-specific details are drawn primarily from Pesquiera (1957) and the Colombian Servicio Nacional de Erradicación de la Malaria (SEM, 1957).
[4]DDT was also instrumental is wiping out remaining traces of malaria in the US, but this was a small change relative to earlier declines in the disease. Humphreys (2001) colorfully characterizes the use of DDT in US malaria eradication as "kicking a dying dog."

the existing stock of parasites would die out. After that, the campaigns would go into a maintenance phase in which imported cases of malaria were to be managed medically.

The Latin American countries analyzed in the present study (Mexico, Colombia, and Brazil)[5] all mounted malaria eradication campaigns, and all saw large declines in malaria prevalence. Panel A of Figure 1 shows malaria cases per capita in Colombia. (Comparable time-series data were not available for the other four countries. Nevertheless, there is little doubt that malaria declined in all four countries immediately following these campaigns.) A decline of approximately 80 percent is seen in the graph. Throughout Latin America, the campaign ultimately proved inadequate to the task, and, in many areas, malaria partially resurged two decades later. But in almost all parts of the hemisphere, malaria never returned to its levels from before the application of DDT.

2.4 Research Design

The first factor in the research design is that the commencement of eradication was substantially due to factors external to the affected regions. The eradication campaign relied heavily upon critical innovations to knowledge from outside the affected areas. Such innovations were not related to or somehow in anticipation of the future growth prospects of the affected areas, and therefore should not be thought of as endogenous in this context. This contrasts with explanations that might have potentially troublesome endogeneity problems, such as, for example, positive income shocks in the endemic regions.

Second, the anti-malaria campaigns achieved considerable progress against the disease in less than a decade. This is a sudden change on historical time scales, especially when compared with trend changes in mortality throughout recent history, or relative to the gradual recession of malaria in the midwestern US or Northern Europe. Moreover, I examine outcomes over a time span of 60 to 150 years of birth, which is unquestionably long relative to the malaria eradication campaigns.

An additional element in the identification strategy is that different areas within each country had distinct incidences of malaria. In general terms, this meant that the residents of the US South, southern Mexico, northern Brazil, and the *tierra caliente* of Colombia were relatively vulnerable to infection.[6] Populations in areas with high (pre-existing) infection rates were in a position to benefit from the new treatments, whereas areas with low endemicity were not. This cross-regional difference permits a treatment/control estimation strategy.

The advent of the eradication effort combines with the cross-area differences in pre-treatment malaria rates to form the research design. The variable of interest is the pre-eradication malaria intensity. By comparing the cross-cohort evolution of outcomes (e.g., adult income) across areas with distinct infection rates, one can assess the contribution of the eradication campaigns to the observed changes. (Specific estimating equations are presented below.)

How realistic is the assumption that areas with high infection rates benefited more from the eradication campaign? Mortality and morbidity data indicate drops of 50 to 80 percent in the decade after the advent of the eradication efforts. Such a dramatic drop in the region's average infection rate, barring a drastic reversal in the pattern of malaria incidence across the region, would have had the hypothesized effect of reducing infection rates *more* in

[5]The choice of these three countries was dictated mainly by the availability of census data in the early stages of this study, but also by these four being countries with both malarious and non-malarious regions.
[6]Humid areas with slow-moving water were the preferred habitat for mosquitoes, the vector that transmits malaria.

highly infected areas than in areas with moderate infection rates. Data on malaria cases by Colombian departaments allow us to examine this directly: the decline in malaria incidence as a function of intensity prior to the eradication campaign is found in Panel B of Figure 1. The basic assumption of the present study — that areas where malaria was highly endemic saw a greater drop in infection than areas with low infection rates — is borne out. (Similar results are seen for US and Mexican states; data for Brazilian states were not available.)

Finally, the timing of the eradication campaign should induce variation in childhood malaria infection that has a marked pattern across year-of-birth cohorts. The present study considers the effects of childhood malaria infection on later-life outcomes, so it is useful to characterize childhood exposure to an eradication campaign. This is shown in Figure 2. Consider a campaign that starts in year zero and takes effect instantaneously. Cohorts born after this date will be exposed to the campaign for their entire childhood. On the other hand, those cohorts who were already adults in year zero will have no childhood exposure to the campaign, while the 'in between' cohorts will be partially exposed during childhood, as shown in the figure. I exploit this timing in two ways. First, in Section 4, I compare the 'born after' cohorts to the 'already adult' cohorts by taking differences across these cohort groups. Second, in Section 5, I use the functional form of childhood exposure in estimation using data for all cohorts. (I discuss some alternative functional forms in Section 6.2.)

These four factors (the external origin of the campaigns, the quick reduction of malaria that followed, the use of nonmalarious areas for comparison, and the differential incidence of eradication benefits across cohorts) combine to form the research design of the present study.

2.5 Related Literature

An important first step in quantifying the output costs of malaria has been to survey the contemporaneous effects of malaria fever. Numerous studies address the question: Following a measured case, how much time is lost at work and/or in leisure/home production? For example, Conly (1975) presents such an analysis in Paraguay, while Bonilla Castro, Kuratomi, Rodríguez, and Rodríguez (1991) perform this exercise in Colombia. Furthermore, Conly links the time-allocation data to contemporaneous measures of farm output.

There are important reasons to believe that these estimates provide a limited picture of malaria's economic impact, however. First, the fever is simply the most acute realization of morbidity from malaria. For a variety of reasons, malaria infection can cause anemia that persists for a considerable time after infection, and Thomas *et al.* (2003) show that anemia has depressing effects on contemporaneous adult productivity. Second, while this methodology considers costs that are evaluated at current prices and constraints, a much larger cost might be that investment in physical capital and land improvement are suppressed by the threat of malaria. Third, and most relevant for the present study, an accounting of how malaria in childhood might affect adult outcomes has been absent from the literature.

The present study is related to several recent papers on human capital and exposure to malaria during childhood. Using a database of Union Army veterans, Hong (2007) correlates early-life exposure to malaria with numerous later-life health outcomes in 19th century America. Barreca (2007) studies the effect on later-life outcomes of specifically *in utero* exposure to short-term fluctuations in malaria caused by within-year variation in rainfall and temperature in the United States. Note that cross sectional and interannual comparisons are both likely to differ from the casual effect of eradication.

Several other papers analyze eradication campaigns, and compare the evolution of outcomes across cohorts in malarious and nonmalarious areas, which is similar to the design on the present study. Lucas (2009) examines women born before and after malaria eradication in Sri Lanka and Paraguay. Cutler, Fung, Kremer, Singhal, and Vogl (2009) compare outcomes across pre- and post-eradication-campaign cohorts in India. The methodologies and results from these two papers are broadly similar to those of the present study, but there are certain important differences as well. I make precise comparisons in

Also relatedly, Acemoglu and Johnson (2006) show that the global decline in mortality from several diseases (including malaria) in the 1950s increased life expectancy in developing countries. They argue, however, that this increase in life expectancy was not matched with an increase in income, relative to high-income countries. Nevertheless, their study does not analyze the impact of malaria separately from other mortality risks. Furthermore, it relies on an instrument that is mechanically weighted towards diseases with greater mortality (in large measure among infants, as it turns out). In contrast, prior to the studied campaigns, malaria in the Americas had a relatively low case-fatality rate. While Acemoglu and Johnson and the present study both measure impacts of health in some broadly defined sense, there is little theoretical basis for assuming that different components of health can be compressed into a single index that summarizes their impact on productivity. Meltzer (1992), for example, shows that health at different points in the life cycle will impact investment and fertility decisions differentially. Furthermore, Malthusian effects (Acemoglu and Johnson's main interpretation for their results) will be attenuated for a disease with comparatively high childhood morbidity, such as is the case for the malaria in the Americas.

3 Data Sources and Definitions

The micro-level data employed in the present study come from the *Integrated Public Use Micro Sample* (IPUMS), a project to harmonize the coding of census microdata from the US and several other countries (Ruggles and Sobek (1997); Sobek *et al.* (2002)). I analyze the census data from the US, Brazil, Colombia, and Mexico.

The geographic units employed in this analysis are place of birth rather than current residence. Matching individuals with malaria rates of the area where they end up as adults would be difficult to interpret because of selective migration. Instead, I use the information on malaria intensity in an individual's place of birth to conduct the analysis, which is therefore an intention-to-treat design. For the US, Mexico, and Brazil, this means the state of birth. The Colombian census also contains information on birthplace by *municipio*, a second-order administrative unit similar to US counties.

For the United States, the base sample consists of native-born white[7] males in the *Integrated Public Use Micro Sample* or IPUMS (Ruggles and Sobek, 1997) and North Atlantic Population Project (NAPP, 2004) datasets between the ages of 35 and 55, inclusive, for the census years 1880–2000, which includes cohorts with years of birth ranging from 1825 to

---

[7]I focus on US whites for several reasons. First, a only small proportion of blacks lived outside of the most malarious states among the earlier cohorts, which means that they make for an imprecisely measured point of comparison. Second and more importantly, that same population of blacks was less likely to have been enslaved, which means that they make for an inappropriate control group for those blacks born into slavery in the malarious South. The estimates reported below (for whites) are similar to those obtained if I include native blacks in the base sample. Estimates using blacks only, however, are imprecise and sensitive to control sets employed. Race was not measured consistently in the Latin America sample, so I work with native males of any race/ethnicity.

1965.[8] I consider males rather than females because their labor-force participation is higher and more consistent across the wide swath of years.

I use two proxies for labor productivity that are available for a large number of censuses. The occupational income score and Duncan socioeconomic index are both average indicators by disaggregated occupational categories that were calibrated using data from the 1950 Census. The former variable is the average by occupation of all reported labor earnings. The measure is instead a weighted average of earnings and education among males within each occupation. Both variables can therefore measure shifts in income that take place between occupations. The Duncan measure has the added benefit of picking up between-occupation shifts in skill requirements for jobs. Occupation has been measured by the Census for more than a century, and so these income proxies are available for a substantial stretch of cohorts. (These and the other income variables below are annual measures; data on hours worked were not available consistently over a long enough span of cohorts.) The income proxies here and throughout are transformed into natural logarithm. There are over 300 categories of occupation on which these indices are based.

The data on native-born males from the Brazilian and Mexican IPUMS-coded censuses from 1960 to 2000 are similarly pooled, and males aged 25–55 are included, which results in birth cohorts from 1905 to 1975. These censuses contain questions on literacy, years of education, and income (both total and earned).

For Colombia, I use the IPUMS microdata on native males from the censuses of 1973 and 1993 (those for which municipio of birth was available). This yields birth cohorts from 1918 to 1968. I use the census-defined variables for literacy and years of schooling. I also construct an income score based on industry and class of worker and benchmarked from the Mexican and Brazilian data. There are over 35 distinct industry/class categories on which this index is based.

I combine microdata from various censuses to construct panels of average outcomes by cohort. Cohorts are defined by both when they were born and where they were born. To construct these panels, I pool the micro-level census data. The individual-level outcomes in the microdata averaged up to the level of year-of-birth × census year × place of birth. (Cohorts can appear in multiple censuses in this pooling strategy, so the resulting dataset is a three-dimensional panel.) To account for the changing precision with which the cell means are estimated, in the regressions below, observations are weighted by square root of the cell size. In Section 4, I compare two groups— cohorts born well before or just after the campaign—so the averages by period of birth are computed accordingly. In Section 5, however, I consider how cross-area outcomes change by year of birth, so the panels are constructed with year of birth × place of birth × census year as the units of observation.

Malaria data are drawn from a variety of sources. US mortality data are reported by the Census (1894), Maxcy (1923), and later in the *Vital Statistics* (Census, 1933.) Malaria mortality for Mexico is reported by Pesqueira (1957) and by the Mexican *Anuario Estadístico* (Dirección General de Estadística, 1960). SEM (1957) and the Colombian Anuario de Salubridad (DANE, 1970) are the sources for the Colombian data. Data on malaria ecology are derived from three sources. Andrew Mellinger (personal communication) provided a raster file of malaria ecology around the globe based on the methodology of Kiszewski, Spielman, Malaney, Mellinger, Sachs, and Sachs (2004).

---

[8]The choice of the lower age bound of 35 for the US excludes those ages on the steeper part of the age-earnings profile. This age was chosen heuristically by inspecting the age-income profiles for several census years. A lower age (25) was chosen for the Latin American countries, which reflected an earlier flattening out of the age-earnings slope in those data. Further, because literacy tends to be realized earlier in life, the lower age bound was set younger (15 years).

Poveda, Graham, and Epstein *et al.* (2000) develop and map an index of malaria ecology for Colombia, which I digitized. I also make use of an index of malaria ecology developed by Hong (2007) for the 19th-century US. I refer to these below as, respectively, the Mellinger, Poveda, and Hong indices. The ecology data were matched with states and municipios using a geographic information system. To facilitate interpretation of the results using these various indices, I re-normalize each malaria measure by the gap between areas at the 95th and 5th percentiles in the malaria distribution of each country. Appendix B contains further details.

A number of additional variables are also employed below as controls. These proxy for cross-area differences in income, health, and other factors (generally measured prior to the campaigns) that might affect or correlate with the evolution of outcomes across cohorts. A description of these control variables is found below and in Appendix C. Inspection of the correlation among these variables reveals that malarious areas in these countries were (i) less developed and (ii) closer to the equator (except in Colombia).

Finally, I remark on the precise definition of childhood exposure. Widespread adoption of new malaria control techniques began in the South around 1920, so this year is chosen as the approximate start of the campaign. The start of large-scale spraying of DDT in Brazil, Colombia, and Mexico was in the mid-to-late 1950s, so 1957 is chosen as an approximate start date. Cohorts born on or after the start year in their respective country are therefore assigned a potential childhood exposure of one. (The exposure variable here is defined only by year of birth; it will be interacted with pre-campaign malaria below to measure a cohorts effective exposure to the eradication campaigns.) I treat childhood as being the first 21 years of life, so that cohorts born more than 21 years before the campaign are assigned a childhood exposure of zero. I postulate, somewhat agnostically, uniform effects of malaria per years of childhood exposure, which implies that the formula for exposure is $\max(\min(21, k-(y-21)), 0)/21$ for year of birth k and a starting year of y. (This functional form provides a reasonable fit to the data, although, in Section 5, I discuss some of its deficiencies.) Section 4 considers the cohorts with childhood exposure equaling either zero or one, Section 5 analyzes all of the cohorts in the data, including those with partial childhood exposure.

4 Pre/Post Comparisons

I compare outcomes across cohorts while separating along two lines: (i) by the timing of birth relative to the campaign and (ii) by the degree of pre-campaign malaria intensity in the place of birth. In this section, to simplify the analysis, I aggregate the different year-of-birth cohorts into two groups: those born well before the campaigns and those who were already adults when the campaign began. (The partially exposed cohorts are therefore not treated until Section 5.) Therefore, for each place of birth, the outcome variables employed in this section are cross-cohort differences (i.e., 'born after' minus 'born well before') in the socioeconomic measures. The basic equation to be estimated is

$$Y_{j,post} - Y_{j,pre} = \beta M_{j,pre} + X_{j,pre}\Gamma + \alpha + \varepsilon_{j,post} \quad (1)$$

in which Y is some socioeconomic outcome for state or municipio j. The time subscript 'post' refers to a year of birth following the malaria-eradication campaign, while 'pre' indicates being born (and having become an adult) prior to the advent of the campaign. The pre-program malaria incidence is $M_{j,pre}$, the X variables are a series of controls, and   is a constant term. The parameter of interest is   . This equation is estimated with weighted least squares, where the square root of the cell sizes are used to construct weights to account with the different precisions with which cohort means are estimated. (Note that this is

numerically equivalent to a 2-period fixed-effect estimator with the malaria×post on the right-hand side instead.) An advantage of this long differencing is to reduce the bias in inference stemming from higher-frequency serial correlation. A disadvantage of this methodology is that it does not account for pre-existing trends, which are considered in the next section.

4.1 United States

Areas in the US with higher malaria burdens prior to the eradication efforts saw larger cross-cohort growth rates in income, as measured by the occupational proxies. These results are found in Table 1. The first row of Panel A contain estimates for the basic specification of equation 1, which includes a dummy for being born in the South plus the natural logarithm of state unskilled wages in 1899. The former variable allows for differential income shifts across regions, while the latter variable, drawn from Lebergott (1964), serves as a correction for possible mean reversion in income. If the oldest cohorts had high malaria infection and low productivity because of some mean-reverting shock, we might expect income gains for the subsequent cohorts even in the absence of a direct effect of malaria eradication on productivity. I present results for both income proxies available for the US: the occupational income score and the Duncan socioeconomic indicator. The first two columns of the table report results using the measure of 1890 malaria mortality, while the remaining two columns use an alternative measure of malaria intensity: a malaria-ecology variable due to Hong (2007).

The estimates for malaria are not substantially affected by the inclusion of a number of additional control variables. The balance of Table 1, Panel A contains these results. The second row controls for additional state-of-birth-level measures of health, including fertility and infant mortality in 1890; late-1910s hookworm infection; state public-health spending and the number of doctors per capita in 1898; and the fraction of recruits rejected from service for health reasons by WWI-era Army physicians. The third row of Panel A shows the estimated effect of malaria when controlling for several education-related controls: the 1910 adult literacy rate, and the logarithmic change (*circa* 1902–32) of teacher salaries, pupil/teacher ratios. The fourth row, marked "Other", includes a mixed basket of controls: male unemployment rates from 1930, the 1910 fraction black, and the 1910 fraction living in urban areas. The specification employed in the final row includes all of the above control variables simultaneously in the regression. Relatedly, the upper left graph in Figure 3 displays a scatter plot of the orthogonal component of cross-cohort income growth versus malaria (the 1890 measure), after having projected each variable onto the broad set of state-level controls.

If these noisy proxies of malaria are measured with independent errors, then the measurement-error bias in any one can be corrected by using the other malaria variable as an instrument. Indeed, as seen in first row of Panel B, the instrumented (2SLS) estimate is higher than the OLS estimate in almost every case. Furthermore, similar results (shown in the second row of the same panel) are obtained using state-average temperature and altitude (plus the interaction of the two) as instruments. The assumption of independence of errors might seem inappropriate for the climate-related instruments, but similar results are obtained by using various subsets of the instruments, and accordingly, a Hausman/NR^2 test fails to reject the null of identical parameter estimates in the second stage.

These results are most likely not because of migration. First, recall that the malaria variable is, throughout all of the main analysis, assigned by area of birth, never by area of residence. This rules out the mechanism that high-income workers migrate to areas where malaria was controlled. Second, I obtain similar results with two additional checks for migration confounds. These results, estimated using the "full controls" specification from Panel A, are

seen in Panel C of Table 1. The first two rows decompose the results by residence in one's state of birth. To compute these estimates, I construct two cohort-level datasets, one for movers and another for nonmovers. Significant and positive effects on income are seen for those who reside in their state of birth ("nonmovers") and those who reside in a different state ("movers"). I then re-do the analysis using the state of birth of the individual's father (results for mother's birthplace are similar). Results suggest an effect of childhood exposure to malaria on income, but data availability places certain limitations on this approach. Namely, parental state of birth is not available for people in the sample born after 1915, so this requires me to redefine the treatment group to those with only partial childhood exposure to the campaign (born 1905–1915 in this case). Standard errors are therefore larger, in accordance with the narrower span of cohorts employed. In the third row of Panel C, I re-estimate the model above, but restrict the sample to those with data on parents' birthplace. The coefficients for this model are broadly similar to those above, although larger in some cases and smaller in others. Next, I consider results for just those individuals with native-born fathers. I use the redefined treatment group, but otherwise replicate the specification from above in that the malaria variable is based on one's own state of birth. Finally, I assign malaria based on father's state of birth, and repeat the analysis. These estimates, found in the fifth row of Panel C, are similar to their baseline estimates in the third and fourth rows.

Finally, results are similar across broadly defined regions, but more precisely determined when considering the Southern and border regions. These estimates are found in Panel D of Table 1. The first row replicates the "full controls" specification from above for the Southern states and states that border the South. Estimates are similar to those above, and statistically significant in 3 of the 4 columns. The second row of Panel D uses the remaining states to estimate the same equation. The coefficients are similar to those above, although imprecisely determined. This latter result is perhaps not surprising given that the malaria problem was largely concentrated in the Southern region of the country.

These point estimates imply substantial, but not unreasonable, reduced-form magnitudes for the effect of childhood exposure to malaria. The income variables are in natural logarithms, so the exposure coefficients can be interpreted approximately as percentage changes in income per unit increase in the independent variable. Recall that the malaria measure is renormalized by the difference across the 95th and 5th percentile states.[9] Therefore, these point estimates suggest a reduced-form effect on income of ten to fifteen percent when comparing the non-malarious to the highly malarious states. That is, in the states with high levels of malaria, cohorts born after the anti-malaria campaign earned 10–20% more than the previous generation, relative to the benchmark of cohorts in malaria-free states.

4.2 Brazil, Colombia, and Mexico

In Brazil and Mexico, malarious areas saw faster cross-cohort growth in income and literacy, but there is mixed evidence regarding differences in years of schooling. Table 2 reports the estimates for these two countries for a variety of outcome and control variables.

Panel A contains estimates from ordinary least squares. The first row presents results from the basic specification, which contains just malaria, region-of-birth dummies, and the log of electrical capacity and of population density in 1950 are used as controls for the pre-program level of development.[10] The next row uses a specification with infant mortality in 1950 to control for possible catch-up in general health. The third row includes controls for

---

[9]The 5th-percentile state (Wyoming) was essentially malaria free, while the 95th-percentile state (Mississippi) had almost 9% of its deaths attributed to malaria.

the 1950 sectoral composition of the labor force: the fraction of the population that is economically active, as well as employment shares in agriculture, extractive industries, manufacturing, transportation, and services. These variables control for sectoral shocks, which is important given the region's adoption of import substitution policies during this time. For the final row of Panel A, all the controls listed above are included as regressors. (These control variables are further described in Appendix C.) The components of malaria and income that are orthogonal to these controls are plotted in Figure 3. The estimates for income are similar across control sets, albeit somewhat larger when I control for sector. Childhood malaria's effect on literacy is consistently positive, although variable in its magnitude and statistical significance. Results for years of education are mixed, as per the theoretical ambiguity.

Table 2, Panel B replicates specifications from Panel A, but corrects for measurement error using state-of-birth data on temperature, altitude, and the interaction of the two as instruments. Estimates using 2SLS are generally larger in magnitude, although less precisely determined. Reduced-form estimates for earned income are comparable between these two countries: around 0.25 for OLS and 0.45 for 2SLS. Estimates for total income in Brazil are higher: around 0.37 for OLS and 0.7 for 2SLS.

Results from Colombia suggest that childhood exposure to the anti-malaria campaign raised income. Cross-cohort growth in income, literacy, and education was higher in the areas with more perverse malaria ecology, as shown in Table 3. I consider three measures of malaria: the Poveda and Mellinger indices of malaria ecology as well as malaria case rates registered by the Colombian eradication campaign in 1955 (SEM, 1957). Specifications with alternative sets of controls are shown in Panel A, The basic specification (seen in the first row) also includes birth-region dummies, log population density and a measure of the general level of economic development, *circa* 1960, from the Colombian Banco de la República (1964). The second row adds controls for conflict, including variables describing the intensity of the 1950s civil war in Colombia known as "La Violencia" and a variable indicating the concentration of land holdings. The third row reports estimates controlling for various types of economic activity: manufacturing employment per capita, proximity to major markets, and dummies for being in the coffee, coal-mining, and cattle-ranching areas. For the fourth row, I control for health differences across areas by including measures of the fraction of each area in which the following diseases occur: leishmaniasis, yellow fever, hookworm, and non-hookworm helminth diseases. (The first two categories are vector-borne diseases and would themselves have been affected by the campaign. I return to this issue below.) The final row of the panel reports estimates controlling for all of the above controls simultaneously. Additionally, the residualized components (after projecting variables onto the full set of controls) of the cross-cohort income changes and malaria ecology are shown in Figure 3.

When correcting for measurement error, estimates of the malaria coefficient rise substantially, which suggests attenuation bias in the OLS estimates. Panel B contains 2SLS estimates produced using several sets of instruments, and with the "full controls" specification. In the first row, the municipio's average temperature, altitude, and the interaction of the two are used as instruments. The other malaria measures are used as instruments in the second row, while dummies for the municipio's Holdridge climate classification are employed in the third row. Finally, the fourth row contains 2SLS estimates

---

[10]Some measure of income (as used for the United States) would be preferable as a mean-reversion control, but no suitable pre-campaign data were found for the full set of states. It is likely that electrical consumption is highly correlated with the level of development, however. According to 2005 data from the CIA World Factbook (2006), gross domestic product per capita and electricity production per capita (in logs) have a correlation coefficient of 0.83 across countries.

produced with all of the above-mentioned instrumental variables. At conventional confidence levels, Hausman/NR^2 tests do not reject null hypothesis of identical parameters estimates derived from each instrument.

For Colombia, I estimate reduced-form magnitudes that are roughly similar across malaria measures. This is especially the case when instruments are used to correct for measurement error, which is likely large, in particular for cases notified. The 2SLS estimates indicate that malarious areas saw faster cross-cohort increases to the tune of five percentage points of literacy, 0.1–0.6 years of schooling, and 0.08–0.16 log points of income.

These estimates suggest similar reduced-form impacts on income of childhood exposure to the anti-malaria campaigns across the three Latin American countries. The income variables are all measured in natural logarithms, and the malaria variables are rescaled so that the gap in malaria between 95th and 5th percentile areas equal one. In these countries, the 5th percentile areas had little to no malaria transmission, so the exposure coefficient measures the evolution of incomes across cohorts in highly malarious regions, relative to the malaria-free benchmark areas. Childhood exposure to the malaria-eradication campaign is associated with a log income gain of around 0.3 and 0.2 in Brazil and Mexico, respectively. The estimate for Colombia (approximately 0.04) is considerably lower, but this is most likely because of the crudeness of the income score based on class of worker and one-digit industry. Because the between- and within-occupation components of the income changes induced by early-life malaria are most likely of the same sign, estimates using this score variable will be underestimates of the total effect on income. To adjust for this, I estimate the same income score for Brazil and obtain numbers 4–6 times smaller than the estimates for total income. On this basis, an adjusted number for Colombia (≈ 0.22) would be quite similar to those found for Brazil and Mexico. These reduced-form magnitudes point to a larger impact in Latin America than in the US South, which is consistent with pre-eradication malaria infection rates being lower in the Southern US than in Latin America. I return to this issue in Section 6 below, where I normalize reduced-form exposure effects by estimated declines in malaria infection.

5 Panel Estimates

Cohorts born late enough to have been exposed to eradication during childhood generally have higher income than earlier cohorts, and this shift correlates with higher potential exposure to the eradication campaign. In this section, I present two types of evidence on this point: (i) panel regressions in which I examine sensitivity of exposure effects to controlling for various other trends, and (ii) graphs of cohort-specific relationships between income and pre-campaign malaria, in which shifts that coincide approximately with childhood exposure are visible.

5.1 Pooled Regression Analysis

Statistical tests indicate that the shift in the income/pre-campaign-malaria relationship coincided with exposure to malaria eradication, rather than with certain trends across cohort or time. These tests are performed using the three-dimensional cohort/panel data structured by year of birth, census year, and area of birth, the construction of which is described in Section 3 and Appendix A. The following equation is estimated:

$$Y_{jkt} = \tilde{\beta}\, M_j \times \mathrm{Exp}_k + \delta_k + \delta_j + \delta_t + \sum_i \left(x_j^i \times \mathrm{Exp}_k \right)\, \gamma_i + \nu_{jkt} \quad (2)$$

in which Exp_k is potential exposure to the malaria-eradication campaign (defined above), M_j is the pre-campaign malaria intensity in area of birth j, the x_j^i are area-of-birth controls, and

the $_k$, $_t$, and $_j$ are fixed effects for year of birth, census year, and area of birth, respectively. (The main effects of the area-of-birth and exposure controls are therefore absorbed by these fixed effects.)

Table 4 reports estimates of equation 2 under a variety of specification assumptions. The dependent variables and countries are indicated in the Panel headings. The Table reports present estimates of , the effect of childhood exposure to the campaign. The controls (the x_j^i) used for first three columns comprise region and mena-reversion controls, which corresponds to the "basic" specification in Section 4. The last three columns uses variables in the "full controls" specifications above.

Panel A of Table 4 contains estimates for the United States. The outcome variable is each cohorts average (log) occupational income score. Estimates most comparable to those in Section 4 are found in columns (1) and (4) of the row labeled "Baseline". The magnitudes here are similar to those estimated using long differences across cohorts, which should be the case. (These two types of estimates will not be exactly the same because partially exposed cohorts are now included in the sample.) The remaining columns of the baseline row present specifications that also include birthplace-specific trends across cohorts, with the trend being modeled as either a first- and second-order polynomial trend in year of birth. Estimates on childhood malaria exposure are not sensitive to controlling for these smooth trend across cohorts.

The rest of Table 4, Panel A considers the issue of childhood exposure versus contemporaneous exposure to malaria. If the effect on income were truly via contemporaneous exposure, this might appear as a cohort effect in that those born earlier spend a larger fraction of their working lives exposed to malaria. (Note that the estimates from the literature suggest relatively small effects on output of episodes of malarial fever, although persistent morbidity would also depress labor productivity contemporaneously.) The first attempt to address this confound is seen in the second row of Panel A. This specification maintains the assumption that areas had similar time trends prior to the eradication campaign, but allows for birthplace-specific trend and level shift after 1920. The estimates on childhood exposure tend to be a bit lower here, although they rise as I also control for differential trends across cohorts as well. The next row allows instead for effects of birthplace×time effects, which relaxes the previous assumption that pre-campaign time trends were the same across areas. In the fourth row of Panel A, I instead drop all of the census years taken prior to 1930, which, by drawing observations from after the campaign, would obviate the compositional problem discussed above. Then, as a final check against area-specific time effects, I control in the fifth row for a full set of region × year × year of birth effects. These estimates are generally similar to the baseline (one cannot reject the null that they are the same in any case), although the estimates incorporating area differences in time effects tend to be a bit lower. Finally, the similarity between estimates for movers and non-movers seen in Section 4 also suggests that the result is not about the local, contemporaneous wage, but instead due to influences that pre-date migration, such as early-life health.

The shift in the income/malaria relationship in Brazil is broadly consistent with a model of childhood exposure to the campaign, rather than some alternative time-series process. These results are found in Panel B of Table 4. The baseline estimates, found in the first row, are comparable to coefficients from the long-difference estimation above, and are robust to including birthplace-specific trends in the regression. (It bears mentioning the span of years in the Latin American data is much shorter than the range for the US, so horse-racing the exposure variable with second-degree trends across cohorts is a more difficult test to pass.)

The rest of Panel B examines the potential confound of contemporaneous exposure to malaria in Brazil and Mexico. Estimates on childhood exposure in the second row are produced from a specification that also includes effects of birthplace × time. On the other hand, note the data from Latin America are all drawn from censuses taken after the campaign had begun. However, the analysis for Brazil and Mexico includes data from 1960, during which time the campaigns were still in full swing. Therefore, in the third row of Panel B, I re-estimate the exposure coefficients excluding the 1960 census data. Finally, the fourth row includes a full set of region × year × year of birth effects. Point estimates that control for time effects that are differential across areas of birth are generally within a standard error of the baseline, but do tend to be lower than earlier estimates. Estimates using the 'full controls' are generally robust to these various checks, but, in several cases, estimates using mean-reversion and region controls only are no longer statistically significant.

Estimates for Colombia are found in Panel C. Baseline estimates, produced without controls for differential trends by area in time or cohort, are seen in columns (1) and (4) of the first row. These are similar to the magnitudes from the long-difference specification above. The second row presents specifications that include birthplace×time effects. The earliest census used for Colombia is from the 1970s, and so the check that involved dropping the 1960 census is not relevant here. Estimates for the final row produced controlling for dummies for region × census year × year of birth. Estimates are broadly similar to the baseline, even when including controls for differential trends across cohorts or across time.

Finally, Panel D reports estimates for Mexico. The baseline estimates are comparable to, albeit a bit lower than, the results for Mexico using the pre/post design above. As above, this Panel includes tests for alternative hypotheses for area-specific trends by cohort or by time. Using the basic specification with mean-reversion and region controls (Column 1)), estimates of childhood exposure remain statistically significant in the face of various checks for time effects that were differential across areas, but none of these results, including the baseline, are robust to including differential trends across cohorts (Columns 2 and 3). In contrast, when the full set of birthplace-level controls is used (Columns 4–6), estimates of the effect of childhood malaria exposure are more robust to these checks for differential trends by cohort or by time.

5.2 Graphical Evidence: Cohort-specific Estimates

The shift in the malaria-income relationship coincides with childhood exposure to the eradication efforts, which be seen graphically in this section. I compare changes in socioeconomic outcomes by cohort across areas with distinct pre-campaign malaria intensities in order to assess the contribution of the eradication campaign to the observed changes. For each year of birth, OLS regression coefficients are estimated on the resulting data by states/municipios of birth. Consider a simple regression model of an average outcome, Y_{jkt}, for a cohort with state of birth j, census year t, and year of birth k:

$$Y_{jkt} = \beta_k\, M_{j,pre} + \delta_k + X_j\, \Gamma_k + \nu_{jkt} \quad (3)$$

in which M_j is the pre-campaign malaria intensity in place of birth j, $_k$ is year-of-birth-specific coefficient on malaria, X_j is a vector of other state-of-birth controls, and $_k$ and $_k$ are cohort-specific intercept and slope coefficients. I estimate this equation using weighted OLS for each year of birth k, thus generating a series of estimates across cohorts. This specification allows one to examine how the relationship between income and pre-eradication malaria ($_k$) differs across cohorts. The coefficients on the control variables are similarly flexible by year of birth. (Note that this is conceptually similar to pooling the data for all years of birth and interacting the independent variables with a full set of year-of-birth

dummies.) The control variables correspond to the "full controls" specifications above. (Appendices available on the author's web site show graphs for the basic specification, as well as graphs of the cohort-specific coefficients on the controls.)

I present simple graphical analysis using the flexible specification (equation 3) for cross-cohort comparison. Figure 4 display plots of the estimated $_k$, for the various countries under study. The x axis is the cohort's year of birth. The y axis for each graphic plots the estimated cohort-specific coefficients on the area-of-birth measure of malaria. Each cohort's point estimate is marked with a dot.

To relate these results to childhood exposure to malaria, I also plot each cohort's potential childhood exposure to the eradication efforts as a dashed line in the figure. For example, observe that US cohorts that were already adults in 1920 were too old to have benefited from the eradication efforts during childhood. On the other hand, later cohorts experienced less malaria infection during their childhood. This benefit increased for those who were exposed to the anti-malaria efforts for a greater fraction of their childhood. The dashed lines therefore measure the number of years of potential childhood exposure to the malaria-eradication campaign. The line is rescaled such that pre-1895 and post-1925 levels match those of the $_k$. The exposure line is not rescaled in the x dimension.

Results for the US are shown in the upper left of Figure 4, which displays the coefficient on state-of-birth 1890 malaria mortality for each year of birth. The pattern of estimates are broadly consistent with the childhood-exposure model. For those born before 1900, more pre-eradication malaria in one's state of birth predicts lower adult income on average. If malaria infection during childhood reduces adult income, we would expect such a negative relationship for the earlier cohorts, who grew up without the benefit of the eradication campaign. However, for those born after 1920, this negative relationship between pre-campaign malaria and income is no longer present. Again, this is to be expected because the anti-malarial efforts interrupted transmission during the early life of these later cohorts, thus breaking the link between malariousness in one's birthplace and adult income.[11] For the intervening cohorts, who had intermediate levels of childhood exposure to the campaign, pre-campaign malaria predicts lower income, but with magnitude that diminishes for later birth years. This is also consistent with the childhood-exposure model, insofar as partial exposure to the campaign confers partial benefits to these middling cohorts. Moreover, estimated coefficients correspond reasonably well to the exposure function: they move when they should move, and they are flat when they should be flat. The main exception is for the cohorts born around 1920: the exposure model predicts a faster transition than is observed. The model, however, is based on a campaign that is instantaneously successful, whereas a slower campaign would slow the transition somewhat in these graphs.

Graphical results for Brazilian states are shown in the upper right of Figure 4. The outcome variable is the natural log of total income and the malaria proxy is the Mellinger index. For Brazil, the malaria-related change in outcomes across cohorts coincides approximately with childhood exposure to the campaign. When comparing cohorts with zero versus full exposure to the campaign, the estimated coefficients tend to be lower for those who were already adults before the campaign began. Furthermore, the shift the estimated series occurs as childhood exposure to the campaign increases.

---

[11]Another feature of the coefficients is that earlier birth years exhibit considerably greater spread about their central tendency than in later years. This is an artifact of the larger sample sizes available in later censuses, which reduces the sampling variance in the estimated cohort-level means.

The results for Mexico are consistent with the childhood-exposure hypothesis, but the evidence is weaker because of imprecise estimates for the earlier cohorts. Graphs of the cohort-specific malaria coefficients for Mexican states are shown in the lower right of Figure 4. The outcome variable is log earned income. Malaria mortality is available by state for *circa* 1950, and this variable is used to measure the malaria prevalence. As was the case for the other countries, the estimated coefficients for the earlier cohorts tend to be below the average for those born after the campaign.

Graphs of the malaria coefficients for Colombia are found in the lower left of Figure 4. The outcome is the log industrial income score, and the malaria proxy is the Poveda index. The pattern of coefficients coincides approximately with potential childhood exposure to spraying, albeit with a delay. Another notable difference is that the coefficients from the earlier cohorts cluster around zero, while the estimates for later cohorts are positive numbers. Nevertheless, the central hypothesis of the present study is that the coefficients should shift in a manner roughly coincident with childhood exposure, which is seen in the graph for Colombia.

Inspecting these graphs is revealing about the importance of the timing of childhood malaria exposure. The definition of childhood exposure above assumed an effect of malaria infection that was uniform across youth. As seen in this section, the pattern of effective exposure to eradication that this generated across cohorts was a reasonable fit for the data. With this result, we can reject some alternative hypotheses about relative importance of exposure at various points in childhood. First, note that trend breaks were evident around birth years that precede the campaigns by 20 years or so. This fact rules out a model with disproportionately large effects of malaria infection *in utero* or during infancy. Second, this early trend break was indeed a trend break rather than a level shift. Because cohorts born then were adolescents at the time of the campaign, this suggests that the result is not driven exclusively by adolescent exposure.

I also have estimate smoothed versions of the $_k$, but this time organizing the data by age at first exposure to the campaign. (These results in an appendix available upon request.) For Brazil and Colombia, the smoothed representation passes through the mean of those exposed at age 30+ at approximately 20 years. Curves for Mexico and the US reach the mean of those with later-life exposure by the late twenties.[12] In any event, the fact that the early trend break occurs twenty years before the campaigns provides further evidence that these results were not due to selective migration in response to the campaign itself. This leaves open the possibility of selective migration in anticipation of the campaign, but the evidence for the US using malaria in paternal birthplace contradicts this idea. In addition, knowledge of DDT was an Allied war secret in WWII, which renders implausible the idea that Latin Americans would change their migration decisions as early as 1940 in expectation of a chemical that they could not possibly have known about.

6 Interpretation

In this section, I characterize the estimates from above in alternate units and consider several mechanisms for the results.

---

[12]These ages are higher than typical school-leaving ages. This is consistent with the hypothesis of the the present study insofar as human-capital accumulation continues apace into young adulthood (the steep part of the age-earning profile), and that malaria might depress this. Relatedly, Hong (2007) shows that farmers aged 20–35 in more malarious counties experienced slower wealth accumulation in the 19th-century United States. On the other hand, if one takes the view that years in school are the only type of human capital that is potentially impeded by malaria, then these estimated ages are anomalously high.

6.1 Normalizing by the Probability of Childhood Infection

Expanding upon the reduced-form estimates above, I renormalize the effects on adult income per probability of malaria infection. Above, data limitations required using measures of malaria that were heterogeneous across countries, but I constructed comparable reduced-form differences by comparing the most malarious to least malarious areas within each country. Representative values of these estimates are reported in the first row of Table 5, which run from 14 to 37 percent. A difficulty in interpreting these numbers, however, is that they are composed of two parts: (i) the effect on adult income of a given childhood malaria burden, and (ii) the magnitude of decline of the malaria burden following the eradication campaigns. The parameter (i) is of interest because it is portable: it is in units of income per infection rate, a number that can be applied to other situations with known infection rates. I therefore estimate the order of magnitude of (ii) and thereby can calculate the approximate effect on adult income of childhood malaria exposure in units of infection rates.

What was the range of pre-eradication malaria infection within each country? Molineaux (1988) reports on the WHO typology of malaria intensity (and associated malaria-infection rates among children): non-endemic (0%), hypoendemic (0–10%), mesoendemic (10–50%), hyperendemic (50–75%), and holoendemic (75–100%). Molineaux also reports estimates of the spatial distribution of different endemic zones throughout the world. (Both the typology and its associated geography are derived from the experience of many experts and do not simply reflect the opinion of that one author, however.) Taking the midpoint of the reported intervals, information about the types of endemicity within each country is used to estimate the cross-area differences in malaria burden prior to the campaigns. The pre-eradication malaria burden in the US ranged from malaria free to mesoendemic, representing a within-country difference in malaria-infection rates of approximately 0.3. Areas within each of the three Latin American countries varied from essentially zero to hyperendemic, for a range of 0.625 in infection probability. (These are reported in the second row of Table 5.) Because infection rates were thought to have dropped precipitously in the decade following the campaign, I take the pre-campaign level to be an adequate measure of the subsequent decline.[13]

I estimate the effect of childhood malaria infection on adult wages to be substantial: being infected with malaria through childhood lead to a reduction in adult income of approximately 50 percent. I calculate this number by normalizing the reduced-form differences with the estimated decline in malaria. (Note that this procedure has the flavor of Indirect Least Squares.) These estimates are shown in the last row of Table 5. For Brazil, the estimated effect is higher for total income (0.59) than for earned income (0.45). In Mexico, the estimate for earned income is 0.41. For Colombia, the raw estimate from Table 3 is small (approximately 0.07), but I adjust this as above using the Brazilian data as a benchmark. This reduced-form number (0.28) is re-normalized to 0.45, based on a maximal malaria infection rate of 0.625 in Colombia. In the United States, the Duncan socioeconomic index shows a larger response to childhood malaria than the occupational income score. The latter variable is calibrated using total labor income, but only incorporates across-occupation changes in income. Accordingly, it is about 25% smaller than the effect on total income for Brazil. It is unclear whether the Duncan socioeconomic index is an under- or overestimate of the full income effect, since the index effectively double counts schooling. I report it nevertheless for completeness.

---

[13]This is partly an assumption of necessity inasmuch as I have not found similar estimates of infection rates for the post-campaign period. Eradication was slightly less than complete in the decade following the campaign, so this will likely result in a downward bias of the estimates in this subsection. If, for example, infection dropped 95% instead of 100%, the effects on income reported in Table 5 will be about 5% too low.

6.2 Mechanisms

Schooling—Formal education had an important, but by no means dominant, role in the Latin-American results. (Lack of data for the earlier cohorts prevents doing this computation for the US) Using a standard return to schooling of 10% per year, I re-calculate the income effect of childhood exposure to malaria, but with years of schooling held constant. Because in Mexico formal education was estimated to have declined in response to the campaign, earned income would have been 10% higher (for a resulting effect size of 0.47 in units of log income per childhood infection rate) had schooling been fixed. In Brazil and Colombia, where schooling rose in response to childhood exposure to the campaign, the increase in education accounts for less than 25% and 10%, respectively, of the income results.[14] (Similar results were found by Bleakley (2007a), who estimated that years of education accounted for approximately 25% of the effect of early-life hookworm exposure on adult income in the Southern United States.)

Note that the economic theory of schooling does not require that years of schooling be the central intermediate variable between childhood health and adult income. Even if years in school do increase with improving childhood health, the increase in schooling will not account for the entirety of the rise in adult income. Let the function $b(e, h)$ be the discounted net benefits of education for a given amount of health endowment h. At the optimal choice of schooling, let $b^* \quad b(e^*, h)$. This will respond to health via two channels, as seen by taking the full derivative of b^* w.r.t. h:

$$\frac{db^*}{dh} = \frac{de^*}{dh} b_e|_{e^*} + b_h|_{e^*}.$$

The first term values the rise in years of schooling ($\frac{de^*}{dh}$) at the marginal return to schooling ($b_e|_{e^*}$). The second term measures the direct effect of health on labor productivity, evaluated at e^*. Which term is larger is an empirical question, although an application of the Envelope Theorem suggests that the first term should be small. (See Bleakley, 2010, for more analysis on this point.) Notice that, if malaria depresses the return to inframarginal schooling, such as by impeding learning or causing absences, this would appear in the $\frac{db^*}{dh}$ term. While absences are likely part of the story, the sheer magnitude of the effect on income is too large to be caused by missed school due to malarial fevers. Cost-of-illness studies such as those cited above find that time lost to fevers sum up to at most a few weeks per year in endemic areas, or much less than 10% of the school year.

At conventional discount rates, these estimates indicate that eradication brought about an increase in the present discounted value of life-time income for the exposed cohorts. For Brazil and Colombia, the ILS calculations above imply that the added schooling consumed at most two additional years at the beginning of their working-age lifetime, but income rose by around 40% in subsequent years. In Mexico, those who benefited from childhood exposure to eradication saw, on average, more income and fewer years of foregone earnings due to schooling investments.

Other vector-borne diseases—The application of DDT most probably reduced the burden of other vector-borne diseases, but these diseases had minuscule prevalences relative to malaria. In 1962, the government of Colombia (DANE, 1963) reported 22 cases of yellow fever, and 167 cases of leishmaniasis. These diseases were dwarfed by the 21,245 cases of

---

[14]This latter result suggests that the simple effect of extending working-age life expectancy—which would include more time in school and thereby raise income—cannot account for the income results of the present study.

malaria reported in the same year. No systematic information is available for dengue, in part because prevalence was so low in those years that it was not a mandatory-notification disease. The last US outbreak of yellow fever was in the 1880s, and mortality from dengue and leishmaniasis were so rare as to escape even being included in mortality statistics. No specific information on these diseases was found for Brazil and Mexico during the period of the anti-malaria campaign, but similar numbers to Colombia should be expected owing to the similar disease ecologies of these three countries. Because the incidence of these other vector-borne diseases was so small, their simultaneous decline with malaria should not induce more than a small bias in the computations above.

Mortality selection—The eradication campaigns brought about a decline in malaria mortality as well as morbidity, but I argue that selection induced by the change in mortality is most likely not responsible for the results of the present study. First, typically infectious-disease mortality is thought to differentially remove from a population the least healthy (and, for a positive health/income correlation, the least productive). Therefore, when the eradication campaigns reduced mortality, the most plausible composition effect would have reduced the income of the cohorts exposed to the eradication campaign as children. Second, even if malaria killed off those who would have been more productive, this mechanism is probably not of the right order of magnitude to explain the cross-cohort rise in income. For example, suppose that malaria had previous killed off ten percent of children, and this non-surviving group would have earned 20% *more* than the rest of the population. Eradication of malaria would generate a 2% increase in income from earlier to later cohorts. Even if malaria killed 30% of children, and these non-survivors would have earned 30% more than the survivors, this would still only generate a cross-cohort difference of 9%.

Spillovers—The present study considers outcomes at the cohort level, so any general-equilibrium effect within the cohort is built into the estimate. Moreover, if the healthier cohorts simply displaced older workers (because of ranking or signaling), we would expect a different time-series pattern to the coefficients in Section 5. A pure ranking/displacement mechanism would generate income differences by childhood exposure for those born within a generation of the intervention. But this mechanism would not affect the endpoints of the time series. Why? Consider the episode from the US. People born in 1820 never worked in the same labor market with cohorts exposed to the anti-malaria campaign as children. Similarly, people born in 1960 always worked alongside those fully exposed to the treatment. As seen in the Figure 4, the end points tend to be a bit closer together, but a gap remains. Similar patterns are seen for the Latin American countries, although the span of years covered in those samples is too short to completely separate these own-cohort versus spillover effects. There is therefore little indication of large, negative spillover effects from the entrance of these cohorts to the labor market. This contrasts with the interpretation of health effects offered in the Acemoglu-Johnson study cited above. Again, it bears noting that their study analyzed large changes in mortality, which in a sense tilt their exercise towards finding Malthusian effects, relative to analyzing a disease with low case fatality, such as the malaria present in the Americas. Bleakley (2007b) looks for across-cohort spillovers effects for malaria and hookworm eradications in the Southern US and finds them to be positive, although considerably smaller than the main effect of childhood exposure to eradication. That same study also examines the effect of hookworm and malaria eradication on *aggregate* income via the childhood-exposure mechanism and finds estimates that are, if anything, a bit higher than those based on cohort comparisons such as the present study.

6.3 Comparison to Other Cohort-Based Studies of Eradication

Two other recent studies (Lucas, 2009, and Cutler *et al.*, 2009) analyze the path of outcomes across cohorts by their effective exposure to eradication campaigns. Their methodologies are

broadly similar to the ones I employ here, and so I now compare and contrast their results with those above. An advantage of these two papers over the present study is that they have access to measurements of rates of spleen inflammation across areas, which are probably better proxies of the rate of malaria infection than the endemicity variable used in Section 6.1. However, the disadvantages of these studies relative to this one are twofold. First, they assign pre-campaign malaria to individuals based on their place of residence rather than their place of birth. Thus, selective migration could be a confound.[15] Second, the two other studies analyze a single cross section per episode, whereas I pool multiple cross sections to study a longer panel of cohorts. The effects of malaria, according to estimates seen above in Figure 4, play out over at least twenty years of cohorts. In a sample of less than 40 years of birth, it might be difficult to distinguish between a slow transition and a smooth trend. This might explain why the more of the emphasis in these two studies is on measuring differential effects of malaria exposure in the first few years, rather than the first few decades, of life.

Lucas (2009) shows that women born after malaria eradication in Sri Lanka and Paraguay completed more years of schooling, suggesting that returns to education rose faster than child wages in that episode. This result is consistent with positive estimates above for education in Colombia, but not with the mixed results for education in Brazil and Mexico. On the other hand, Lucas also reports estimates of the effect of childhood malaria on literacy that are similar in sign and magnitude to the estimates above for all countries (Lucas, 2009). This is consistent with a model in which childhood malaria has, to first order, an ambiguous impact on time in school, but an unambiguously positive effect on outputs such as literacy. It bears mentioning that Lucas considers ever-married females only, while the present study analyzes males only. While the underlying biological process by which malaria impedes human-capital growth might be similar across gender, the response of opportunity costs to malaria could differ between males and females.

Cutler, Fung, Kremer, Singhal, and Vogl (2009) analyze the malaria-eradication campaign in India and find malaria exposure in the first few years of life reduced household expenditures in adulthood, but had mixed effects on educational attainment. Comparing the 5th- and 95th-percentile areas, they report that the cross-cohort growth in expenditures was 3–13% higher in areas that started with higher malaria. This number is considerably smaller than the estimates above for Latin America, and probably smaller than the estimate for the US, which almost surely had less malaria in 1920 than India did in 1950. Nevertheless, there are a few methodological differences that should be taken into account. First, they compare those born after the campaign to those born before, but the earlier cohorts contains people with partial childhood exposure. Comparable years of birth for Brazil had an average childhood exposure of 0.30, and comparably lagged cohorts (the Indian campaign started about five years earlier) had an average childhood exposure of 0.43. The partial exposure of the earlier cohorts would tend to attenuate their estimates. Second, by virtue of the timing of their sample, the treatment group is quite young: the age range was 20–25 years old. It is well known that there is a 'fanning out' of the income distribution with age, which is in part because the return to human capital rises with age. Specifically for malaria, Hong (2007) showed that wealth accumlation was faster among twenty-somethings in areas without malaria in the 19th-century US. Moreover, if I estimate the models above (for the US) with a sample of just those 25–30 years of age, the coefficients come out about 2/3s of the average estimate. If the first and second factors compound, then an effect on log earned income of

---

[15]Cutler *et al.* report that less then 10% of people in India lived outside their their district of birth at the time of their sample, but this is an average number rather than a marginal contribution of malaria. The experience of (lowland) Nepal, for example, suggests that the marginal contribution of malaria control can be large. On the other hand, if I re-run the regressions above using area of residence rather than area of birth, the estimates are not much different from the baseline above.

0.27 (the estimate for Brazil) would be attenuated down to 0.08–0.12, which is within the range estimated by Cutler *et al.*

6.4 Extrapolations

The estimated impact of childhood malaria is large enough that it bears consideration in a macroeconomic context, although it is not so large that it can account for the observed dispersion in income across areas.[16] The log-income gap between the North and the South at 1900 was approximately 0.75. For a 10–20% infection rate in the South and an effect of childhood malaria on log total income of 0.5 from Table 5, we would expect a reduction in Southern incomes of approximately 0.05–0.10 log points. In other words, some 7–13% of this income gap could be attributed to malaria infection in the South. On the other hand, the 1950 difference in log GDP between the United States and the three Latin American countries was between 1.5 and 2. If these countries had 30–40% malaria infection rates among children, we would expect eradication to have reduced this gap by 0.19–0.25 in natural log terms, which would close 10–16% of the gap with the US

The benefits estimated above played out across cohorts rather than time, so some time would have had to pass before the healthier cohorts filled out the working-age population. Indeed, only around 20% of the effect in Latin America would have been realized by 1980, and the full gains will not be seen until almost 2010. For the econometrician working with standard output data from this transition period, there would appear to be a growth effect of malaria, although this is the result of time-aggregation bias.

A number of macroeconomic studies attempt to measure the impact of malaria on economic output across countries, but their estimates are too large to be plausibly generated by the childhood-exposure mechanism identified in the present study. Examples of such work are by Gallup and Sachs (2001) and Sachs (2003). The latter study reports an estimate of −1.3, while the present study quantifies the childhood-exposure channel to be approximately −0.5, about 40% of the cross-country estimate. However, the Sachs estimates are in units of log(GDP) per fraction of population potentially exposed to malaria, in contrast with the results above, whose units are log(income) per probability of childhood infection. Because the fraction exposed is less than the fraction infected, we need to inflate the former number by some amount. If the fully exposed countries have childhood infection rates around 0.6, then the Sachs estimate in my units is −2.16, and I can account for only around 25% of that result. On the other hand, those cross-country studies emphasize the importance of *falciparum*, whereas *vivax* was probably the predominant strain of malaria in the Americas at the time of the studied campaigns. Another point of contrast comes from noting that, on the one hand, the IV estimates of the depressing effect of persistent childhood malaria are approximately 50%, but, on the other hand, the richest and poorest countries are separated by closer to a factor of twenty in income.

7 Conclusions

This study considers the socioeconomic impact of malaria-eradication campaigns in the United States (circa 1920), and in Brazil, Colombia, and Mexico (circa 1955). The specific goal is to measure how much childhood exposure to malaria depresses labor productivity, a channel which is neglected by traditional microeconomic studies of malaria.

Several factors combine to form the research design. The eradication campaigns studied happened because of advances in medical and public-health knowledge, which mitigates concerns about reverse causality of the timing of eradication efforts. Highly malarious areas

---

[16]The approximate infection rates cited in this subsection are estimated from Molineaux (1988).

saw large drops in their malaria incidence following the campaign. Furthermore, these gains against the disease were realized in approximately a decade. Finally, sufficient time has passed that we can evaluate its long-term consequences. Data from regional malaria eradication programs were collected and collated with publicly available census microdata.

In both absolute terms and relative to the comparison group of non-malarious areas, cohorts born after eradication had higher adult income and literacy than the preceding generation. This shift across cohorts is shown to coincide with childhood exposure to the eradication efforts. This suggests that being exposed to malaria in childhood depresses labor productivity as an adult. In the most malarious areas of the Latin American countries studied, cohorts born after the anti-malaria campaign earned approximately 25% more than the previous generation, relative to the cross-cohort change in the malaria-free areas, while the comparable reduced-form change in the US was approximately 12%. Considered in terms of the probability of persistent childhood infection, this effect is substantial: IV estimates from Section 6 indicate that reducing one's point-in-time probability of childhood malaria infection from one to zero results in earning approximately 50 percent less as an adult.

Mixed results are found for years of education, in contrast with consistently positive effects of malaria eradication on income and literacy. Furthermore, in no country can the change in income be accounted for by the change in years of schooling. These facts are in no way discordant with the economic theory of schooling, which compares returns with opportunity costs. Childhood health plausibly raises both, leaving an ambiguous effect on the optimal time to spend in school. This combination of results, interpreted with simple price-theoretic reasoning regarding the education decision, show that we should be cautious in using changes in time in school as a sufficient statistic by which development and health policies are evaluated.

Another notable result of the present study is that estimates of the effect of childhood malaria exposure on adult income, when expressed in comparable units, are similar across the four countries considered. This fact is remarkable particularly given the composition of the sample: one (eventually) developed region of a developed country and three others that are still on the path of development. Put another way, in spite of the differences in culture, institutions, and endowments among the countries studied, the effect of malaria was broadly similar.

Finally, the results indicate potentially large benefits of interventions against malaria (especially the *vivax* strain) in tropical countries where it is still endemic today, although the benefits of reducing childhood exposure to malaria are nevertheless small relative to dispersion of income across countries or to the gains suggested by cross-country regressions.

References

Abbott, Samuel W. The Past and Present Condition of Public Hygiene and State Medicine in the United States. In: Adams, HB., editor. Monographs of American Social Economics, Vol. XIX, Commonwealth of Massachusetts. Prepared for the United States Commission to the Paris Exposition of 1900; 1900.

Acemoglu, Daron; Johnson, Simon. Working Paper 12269. National Bureau of Economic Research; Cambridge, Mass: 2006 Jun. Disease and Development: The Effect of Life Expectancy on Economic Growth.

Banco de la República. Atlas de Economía Colombiana. Bogotá: Departamento de Imprenta del Banco de la República; 1964.

Barlow, Robin. The Economic Effects of Malaria Eradication. American Economic Review May. 1967; 57(2):130–148.

Barreca, Alan. Technical Report. UC-Davis; 2007. The Long-Term Economic Impact of *In Utero* and Postnatal Exposure to Malaria.

Bleakley, Hoyt. Disease and Development: Evidence from Hookworm Eradication in the American South. The Quarterly Journal of Economics. 2007a Feb; 122(1):73–117. [PubMed: 24146438]

Bleakley, Hoyt. Unpublished manuscript. University of Chicago; 2007b. Spillovers and Aggregate Effects of Health Capital: Evidence from Campaigns Against Parasitic Disease in the Americas.

Bleakley, Hoyt. Health, Human Capital, and Development. Annual Reviews in Economics. 2010 forthcoming.

Bloom, David E.; Canning, David; Sevilla, Jaypee. The Effect of Health on Economic Growth: A Production Function Approach. World Development. 2004; 32:1–13.

Castro, Bonilla; Elssy, Luz; Kuratomi, Stella; Rodríguez, Penélope; Rodríguez, Alejandro. Salud y Desarrollo: Aspectos Socioeconómicos de la Malaria en Colombia. Bogotá: Universidad de los Andes, Centro de Estudios sobre Desarrollo Económico (CEDE); 1991 Jan.

Central Intelligence Agency. CIA World Fact Book. Washington: Central Intelligence Agency; 2006 Jun 29. Web page: http://www.cia.gov/cia/publications/factbook/

Conly, Gladys N. Scientific Publication 297. Pan American Health Organization; Washington, D.C: 1975. The Impact of Malaria on Economic Development: A Case Study.

Coordinación General de los Servicios Nacionales de Estadística, Geografía e Informatica. Incidencia de la mortalidad infantil en los Estados Unidos Mexicanos 1940–1975. Mexico, D.F: Secretaria de Programación y Presupuesto; 1981.

Cutler, David; Fung, Winnie; Kremer, Michael; Singhal, Monica; Vogl, Tom. Working Paper 13539. National Bureau of Economic Research; 2007 Oct. Mosquitoes: The Long-term Effects of Malaria Eradication in India. (Revised on August 17, 2009.)

Departamento Administrativo Nacional de Estadística (DANE). Anuario General de Estadística. Bogotá, Colombia: Departamento Administrativo Nacional de Estadística; 1942, 1963, 1965.

Departamento Administrativo Nacional de Estadística (DANE). Censo de Población. Bogotá, Colombia: Departamento Administrativo Nacional de Estadística. International Population Publications, microfilmed by Research Publications, New Haven, Conn; 1951.

Departamento Administrativo Nacional de Estadística (DANE). Anuario de Salubridad. Bogotá, Colombia: Departamento Administrativo Nacional de Estadística; 1968–70.

Departamento Administrativo Nacional de Estadística (DANE). División Político-Administrativa de Colombia. Bogotá, Colombia: DANE; 2000.

Dirección General de Estadística. Censo General de Población 1950. Mexico, D.F: Secretaria de la Economía Nacional, Direccion General de Estadística. International Population Publications, microfilmed by Research Publications, New Haven, Conn; 1952a.

Dirección General de Estadística. Compendio Estadistico. Mexico, D.F: Secretaria de la Economía Nacional; 1952b.

Dirección General de Estadística. Anuario Estadístico 1958–1959. Mexico, D.F: Secretaria de Industria y Comercio; 1960.

Dirección Nacional de Estadística. Primer censo industrial de Colombia, 1945. Vol. 1–10. Bogotá: Imprenta Nacional; 1947.

Gallup, John Luke; Sachs, Jeffrey D. The Economic Burden of Malaria. The American Journal of Tropical Medicine & Hygiene. 2001 Jan-Feb;64(1, 2S):85–96. [PubMed: 11425181]

Gallup, John Luke; Mellinger, Andrew; Sachs, Jeffrey D. Working Paper 1. Center for International Development at Harvard University; Cambridge, Mass: 1999a Mar. Geography and Economic Development.

Gallup, John Luke; Mellinger, Andrew; Sachs, Jeffrey D. "GIS Data." Machine-readable dataset (Arc/info interchange file). Center for International Development (CID) at Harvard University. ; 1999b. http://www2.cid.harvard.edu/ciddata/Geog/gisdata.html

Harrison, Gordon. Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880. New York, N.Y: E. P. Dutton; 1978.

Hong, Sok Chul. Doctoral dissertation. University of Chicago; 2007. Health and Economic Burden of Malaria in Nineteenth-Century America.

Humphreys, Margaret. Malaria: poverty, race, and public health in the United States. Johns Hopkins University Press; 2001.

Instituto Brasileiro de Geografia e Estatistica (IBGE). Rio de Janeiro: Servicio Grafico do IBGE. microfiche by Chadwyck-Healey Ltd; Cambridge, England: 1951. Anuario Estadistico do Brasil. June 1978, part of Latin American & Caribbean official statistical serials on microfiche

Instituto Brasileiro de Geografia e Estatistica (IBGE). VI Recenseamento Geral do Brasil, 1950. Rio de Janeiro: IBGE Conselho Nacional de Estatistico, Servicio Nacional de Recenseamento; International Population Publications, microfilmed by Research Publications; New Haven, Conn: 1956.

Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: the United States, 1790–1970. Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research (ICPSR). Computer file, ; 1984. http://www.icpsr.org [Accessed Sept. 20, 2002]

International Health Board. Annual Report. New York: Rockefeller Foundation; 1919.

Jonnes, Peter G.; Bell, William C. Machine-readable dataset (ArcGIS shapefile) Version 1.1. Centro Internacional de Agricultura Tropical (CIAT); Cali, Valle de Cauca, Colombia: 1997 Apr. Cobertura de América Latina. División Administrativa.

Kiszewski, Anthony; Spielman, Andrew; Malaney, Pia; Mellinger, Andrew; Sachs, Sonia Ehrlich; Sachs, Jeffrey. A Global Index Representing the Stability of Malaria Transmission. American Journal of Tropical Medicine and Hygiene. 2004; 70(5):486–498. [PubMed: 15155980]

Kofoid, Charles A.; Tucker, John P. On the Relationship of Infection by Hookworm to the Incidence of Morbidity and Mortality in 22,842 Men of the United States Army. American Journal of Hygiene. 1921 Jan; 1(1):79–117.

Lebergott, Stanley. Manpower in Economic Growth: the American Record since 1800. New York: McGraw-Hill Inc; 1964.

Lorentzen, Peter L.; McMillan, John; Wacziarg, Romain. Death and Development. CEPR Discussion Paper 5246, CEPR. 2005 Sep.

Lucas, Adrienne M. Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka. 2009 Feb. Unpublished manuscript.

Maxcy, Kenneth F. The Distribution of Malaria in the United States as Indicated by Mortality Reports. Public Health Reports. 1923 May; 38(21):1125–38. [PubMed: 19314879]

McCarthy, Desmond; Wolf, Holger; Wu, Yi. Working Paper 7541. National Bureau of Economic Research; Cambridge, Mass: 2000 Feb. The Growth Costs of Malaria.

Meltzer, David. Doctoral dissertation. University of Chicago; 1992 Dec. Mortality decline, the Demographic Transition, and Economic Growth.

Molineaux, Louis. The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. In: Wernsdorfer, WH.; McGregor, IA., editors. Malaria: Principles and Practice of Malariology. Vol. 2. Vol. 35. London, U.K: Churchill Livinstone; 1988. p. 913-998.

North Atlantic Population Project. NAPP: Complete Count Microdata, Preliminary Version 0.2. Minneapolis, Minn: Minnesota Population Center [distributor]; 2004. Computer file, http://www.nappdata.org [accessed November 12, 2005]

Oquist, Paul H, Jr. Doctoral dissertation. University of California; Berkeley. Photocopy: Ann Arbor, Mich: University Microfilms International; 1976. Violence, Conflict, and Politics in Colombia. 1978

Pesqueira, Manuel E. Programa de Erradicación del Paludismo en México. Boletin de la Officina Sanitaria Panamericana. 1957 Jun; XLII(6):537–41.

Poveda, German; Graham, Nicholas E.; Epstein, Paul R.; Rojas, William; Quiñones, Martha L.; Vélez, Iván Darío; Martens, Willem JM. Climate and ENSO Variability Associated with Vector-Borne Diseases in Colombia. In: Diaz, HF.; Markgraf, V., editors. El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts. Cambridge University Press; Cambridge, U.K. and New York, N.Y: 2000. p. 183-204.

Ruggles, Steven and Matthew Sobek. Integrated Public Use Microdata Series: Version 2.0. Minneapolis, Minn: Historical Census Projects, University of Minnesota. ; 1997. http:// www.ipums.umn.edu

Sachs, Jeffrey D. Working Paper 8119. National Bureau of Economic Research; Cambridge, Mass: 2001 Feb. Tropical Underdevelopment.

Sachs, Jeffrey D. Working Paper 9490. National Bureau of Economic Research; Cambridge, Mass: 2003 Feb. Institutions Don't Rule: Direct Effects of Geography on Per Capita Income.

Servicio Nacional de Erradicación de la Malaria (SEM). Plan de Erradicación de la Malaria, Vol. Tomos I and II. Bogotá, Colombia: Ministerio de Salud Pública; 1957.

Servicio Nacional de Erradicación de la Malaria (SEM). Technical report. Ministerio de Salud, Bogotá, Colombia: 1979 Apr. Evaluación del Programa de Malaria: Marco de referencia para el grupo de as-esorimiento nacional e internacional.

Sobek, Matthew; Ruggles, Steven; McCaa, Robert; King, Miriam; Levison, Deborah. Integrated Public Use Microdata Series-International: Preliminary Version 1.0. Minneapolis: Minnesota Population Center, University of Minnesota. Computer file, ; 2002. http://www.ipums.org/international [accessed February 21, 2005]

Thomas, Duncan; Frankenberg, Elizabeth; Friedman, Jed; Habicht, Jean-Pierre; Hakimi, Mohamme; Jaswadi, Nathan Jones; Pelto, Gretel; Sikoki, Bondan; Seeman, Teresa; Smith, James P.; Sumantri, Cecep; Suriastini, Wayan; Wilpo, Siswanto. Iron defficiency and the well-being of older adults: Early results from a randomized nutrition intervention. 2003 Apr. Unpublished manuscript

U.S. Bureau of the Census. Vital and Social Statistics. Vol. II. Washington: U.S. Government Printing Office; 1894. Compendium of the Eleventh Census: 1890.

U.S. Bureau of the Census. Vital Statistics of the United States. Washington: U.S. Government Printing Office; 1933.

US Office [Bureau] of Education. Annual Report of the Commissioner of Education. Washington D.C: Government Printing Office; 1894–1913.

Weil, David N. Accounting for the Effect of Health on Economic Growth. The Quarterly Journal of Economics. 2007 Aug; 122(3):1265–1306.

Williams, Ralph Chester. The United States Public Health Service, 1798–1950. Bethesda, Md: Commissioned Officers Association of the United States Public Health Service; 1951.

A Construction of the Cohort-Level Data

The micro data for the analysis are drawn primarily from the IPUMS data for the United States, Brazil, Colombia, and Mexico. For each country, these data are used to construct a pseudo-panel of average outcomes (principally income) by year of birth, area of birth, and census year. These average outcomes form the dependent variables used in the present study. Because these averages are constructed with differing degrees of precision due to differing numbers of observations in each cell, I also compute the square root of the cell sizes to use as weights when estimating equation 3 in the main text. To adjust for differences in units and composition across census years, the national mean is removed from each year and age cell. Income variables are transformed into natural logarithms.

A.1 Details for the United States Sample

The underlying sample used for the United States consists of native-born white males in the age range [35,55] in the 1900–1990 IPUMS microdata or in the 1880 microdata from the North Atlantic Population Project (NAPP, 2004). (These data were last accessed November 14, 2005.) This results in a data set with year-of-birth cohorts from 1825 to 1965. The original micro-level variables are defined as follows:

- **Occupational income score.** The occupational income score is an indicator of income by disaggregated occupational categories. It was calibrated using data from

the 1950 Census, and is the average by occupation of all reported labor earnings. See Ruggles and Sobek (1997) for further details.

- **Duncan socio-economic index.** This measure is a weighted average of earnings and education among males within each occupation. This measure serves to proxy for both the income and skill requirements in each occupation. It was similarly calibrated using data from the 1950 Census.

For the majority of the years of birth, I can compute average income proxies for all of the 50 states plus the District of Columbia. The availability of state-level malaria data and the control variables restricts the sample further to 46 states of birth. Alaska, Colorado, the District of Colombia, Hawaii, and Oklahoma are excluded because of missing data for at least one of the other independent variables. This leaves 46 states of birth in the base sample.

There are a number of cohorts born before 1885 for which as few as 37 states of birth are represented. For those born between 1855 and 1885, this appears to be due to small samples, because, while the NAPP data are a 100% sample for 1880, there are no microdata for 1890 and 1900 IPUMS data are only a 1% sample. On the other hand, for the 1843–1855 birth cohorts, all but two of the years have all 46 states represented. Nevertheless, even with the 100% sample from 1880, there are as many as six states per year missing for those cohorts born before 1843. A number of the territories (all of which would later become states) were being first settled by people of European descent during the first half of the 19th century, and it is quite possible that, in certain years, no one eligible to be enumerated was born in some territories. (Untaxed Indians were not counted in the censuses.) Note that I use the term state above to refer to states or territories. Territories were valid areas of birth in the earlier censuses, and are coded in the same way as if they had been states.

While this procedure generates an unbalanced panel, results are similar when using a balanced panel with only those states of birth with the maximum of 141 valid observations. A comparison of the cohort-specific estimates from the balanced and unbalanced panels shows high correlation (over 0.96, for example, in the case of the additional-controls specification for the occupational income score).

A.2 Details for the Brazilian Sample

The underlying sample used for Brazil consists of native males in the age range [15,55] in the 1960–2000 IPUMS microdata. (These data were last accessed April 7, 2006). This results in a data set with year-of-birth cohorts from 1905 to 1985.

State of birth is available for these samples. Brazilian states (and several territories that were to become states) were, by and large, consistently defined over the course of the sample. Those few that were not were merged together to reflect administrative divisions in the early 1950s. Specifically, I merged Rondônia into Guaporé, Roraima into Rio Branco, Tocantins into Goias, Fernando de Noronha into Pernambuco, Serra do Aimores into Minas Gerais, and Mato Grosso do Sul into Mato Grosso.

The original micro-level variables are as follows:

- **Literacy.** A binary variable individual measuring whether an individual can read and write at least a simple note.

- **Years of Schooling.** Numbers of years of education corresponding to highest grade completed. Non-numeric responses (e.g., "some secondary") are mapped onto the midpoints of the appropriate intervals.

- **Total Income.** Records the total personal income from all sources in the prior month. In the empirical work above, this variable is measured in natural logs. This variable is reported in income categories in the 1960 census, and their midpoints are used in translating the data into income.

- **Earned Income.** Records the personal income from their labor (wages, business, or farm) in the prior month. In the empirical work above, this variable is measured in natural logs.

A.3 Details for the Colombian Sample

The underlying sample used for Colombia consists of native males in the age range [15,60] in the 1973 and 1991 IPUMS microdata. (These data were last accessed April 10, 2006). This results in a data set with year-of-birth cohorts from 1918 to 1976.

Area of birth is available in these samples at the level of departamento and municipio. The departamento is a first-order administrative division, similar to a state, while the municipio is a second-order division, similar to a county in the United States. A cohort's municipio of birth is used in the present study to construct a proxy for childhood exposure to malaria. Colombia contains over one thousand municipios in the present day, but, to preserve confidentiality in the the IPUMS data, some of the smaller municipios are aggregated into larger groupings. This results in over 500 unique codes for area of birth, and I refer to these units simply as "municipios" in the text. Because municipal boundaries change over time, maps (SEM, 1957) and other administrative information (DANE, 2000) were used to relate data observed at various points in time onto the IPUMS recode of municipio.

The original micro-level variables are as follows:

- **Literacy.** A binary variable individual measuring whether an individual can read and write.

- **Years of Schooling.** Numbers of years of education corresponding to highest grade completed. Non-numeric responses (e.g., "some secondary") are mapped onto the midpoints of the appropriate intervals.

- **Industrial Income Score.** The industrial income score is an indicator of income by industry and class of worker. It was calibrated using data from the Brazilian and Mexican censuses for all available years. To remove census-year *times* country effects, the starting point for this variable is log total income after being projected onto year×country effects. These residuals are then averaged by industry and class of worker and matched onto the Colombian sample. Because of the way this score is constructed, the variable is measured in natural logs. (Total income is available in the 1973 Colombian census, but the range of years of birth that these data cover is too limited.)

A.4 Details for the Mexican Sample

The underlying sample used for Mexico consists of native males in the age range [15,60] in the 1960–2000 IPUMS microdata. (These data were last accessed April 7, 2006). This results in a data set with year-of-birth cohorts from 1905 to 1984.

State of birth is available for these samples. Mexican states (some of which were territories early on) were defined consistently throughout the sample period.

The original micro-level variables are as follows:

- **Literacy.** A binary variable individual measuring whether an individual can read and write.

- **Years of Schooling.** Numbers of years of education corresponding to highest grade completed. Non-numeric responses (e.g., "some secondary") are mapped onto the midpoints of the appropriate intervals.

- **Earned Income.** Records the personal income from their labor (wages, business, or farm) in the prior month. In the empirical work above, this variable is treated in natural logs. (Total income is available in certain years of the Mexican censuses, but the range of years of birth that these data cover is inappropriate for the analysis.)

B Sources and Construction of the Malaria Data

- **United States.** Malaria mortality expressed as a fraction of total mortality. This was measured in the 1890 Census and refers to the proceeding year. I normalize by total mortality in the state to filter any factor in the underreporting that is common to malaria and total mortality. These data were collected by Census and reported at the state level. (U.S. Bureau of the Census, 1894.)

  As a specification check, I use three alternative measures of pre-campaign malaria intensity. An additional measure of malaria mortality refers to 1919–21 and is drawn from Maxcy (1923), who surveyed state departments of health. Because the death-registration system was not yet completely operational at the time of the Maxcy's survey, these data are to be taken *cum grano salis*. No official measure of total mortality was available for the period, so I scale the data by population instead. The remaining two malaria variables are indices based on geographic and climatic factors. The first index (referred to as "Mellinger" in the text) of malaria ecology is computed using information on climate and local vectorial capacity. The construction of these data are described in Mellinger *et al.* (2004). The source data were provided as raster data in one-degree grids. A GIS program ("spatial analyst" within ArcView) was used to extract average malaria ecology by state. The second index is based on climatic and geographic factors and is due to Hong (2007), who calibrates his index using 19th-Century data on malaria morbidity in U.S. Army forts throughout the country. The data were made available to the author at the (1890) county level, and were aggregated to the state level on a 1890-population-weighted basis. The Census and Hong measures are used in the main analysis. The Maxcy and Mellinger measures are used as instruments for measurement error. Analysis with all four measures is available from the author upon request. (Maxcy, 1923; Andrew Mellinger, private communication, and author's calculations; Sok-Chul Hong, private communication, and author's calculations.)

- **Brazil.** An index of malaria ecology, computed using information on climate and local vectorial capacity. The construction of these data are described in Mellinger *et al.* (2004). The source data were provided as raster data in one-degree grids. A GIS program was used to extract average malaria ecology by state. (Andrew Mellinger, private communication, and author's calculations.)

- **Colombia.** Two measures of ecology are used, as well as one measure of morbidity. The Poveda measure is an index of malaria ecology based on climatic factors, described by Poveda, Graham, Epstein, Rojas, Quiñones, Darío Vélez, and Martens (2000).

  A map in that study displaying the computed survival probability of *p. vivax* (Fig 6.5) was digitized and fed into a GIS program, which was then used to construct

averages by municipio. The Mellinger measure of malaria ecology is the same as that used for Brazil, and was averaged by municipio in a GIS program (the "Spatial Analyst" toolbox within ArcView). Glenn Hyman of the Centro Internacional de Agricultura Tropical shared data on the Colombian municipio boundaries. Malaria cases notified per capita at the municipio level were drawn from the reports of the Servicio Nacional de Erradicación de la Malaria (SEM) and refer to 1956. (Poveda et al., 2000; Andrew Mellinger, private communication; Jonnes and Bell, 1997; SEM, 1957; and author's calculations.)

To account for measurement error in the above variables, I also construct climate-based instruments. The set of instruments consists of the municipio's temperature, altitude and the interaction of the two. The temperature and altitude data are from records prior to 1960 and reported by the Banco de la República (1964). Another proxy for climate is the fraction of each municipio within particular Holdridge climate zones. Those relevant for the areas under study are the following: cool temperate, warm temperate, subtropical dry, subtropical wet, tropical dry, and tropical wet. These data come from a GIS file provided by the Center for International Development at Harvard University, and were computed by municipio in a GIS program (the "spatial join" in ArcView). (Banco de la República, 1964; Gallup, Mellinger, and Sachs 1999b; and author's calculations.)

- **Mexico.** Malaria mortality by state, expressed in per-capita terms. (Pesqueira, 1957.)

C Control Variables

Control variables for the United States

- **Average wage, 1899.** I input the average monthly earnings (with board) for farm laborers by state in 1899. Various other wage measures are summarized by the same source, but are generally not available for a complete set of states. I transform this measure into natural logs. (Lebergott, 1964, Table A–24.)

- Dummy for being born in the South.

- **Doctors per capita, 1898.** Number of physicians per 1,000 inhabitants of each state. The primary source is listed as Polk's *Register of Physicians*, 1898. (Abbott, 1900).

- **State public-health spending, 1898.** Per capita appropriations, by state, for state boards of health in 1898. Primary sources include the annual reports of state boards of health, state appropriations laws, and correspondence with the secretaries of the boards of health. (Abbott, 1900.)

- **Infant mortality, 1890.** The estimates of infant mortality are constructed from published tabulations. Table 3 in Part III contains enumerated deaths of children under one year of age. I scale this number by the estimated birth rate (Part I, page 482) times the female population (Part I, Table 2). The rate from 1890 was used because child-mortality data are not available comprehensively for the years 1900–1932, during which time the death-registration system was established. The 1890 mortality data were collected by Census enumerators. (U.S. Bureau of the Census, 1894.)

- **Hookworm Infection.** Computed from examinations of army recruits. (Kofoid and Tucker, 1921)

- **Log change in School Term Length, c. 1902–1932.** Average length of school term, in weeks. (Annual reports of the federal Commissioner of Education, U.S. Office [Bureau] of Education, 1905–1932.)

- **Log change in Pupil/Teacher Ratio, c. 1902–1932.** Average attendance divided by number of teachers. (Annual reports of the federal Commissioner of Education, U.S. Office [Bureau] of Education, 1905–1932.)

- **Adult literacy rate.** These data were compiled at the state level and come from the 1910 Census. Adult literacy refers to males of voting age. (ICPSR #3.)

- **Population urban.** From Census tabulations measuring the population residing in metro areas in 1910. (ICPSR #3)

- **Fraction black.** From tabulations of the 1910 Census. (ICPSR #3)

- **Male unemployment rate.** From tabulations of the 1930 Census. (ICPSR #3)

Control variables for the Brazilian states

- **Region dummies.** North (Norte and Nordeste) and South (Centro-Oeste, Sudeste, and Sul).

- **Population Density.** Population per square kilometer in 1950. (IBGE, 1950 and 1951.)

- **Infant mortality.** Number of infant deaths in the municipio of the state capital, scaled by the estimated birth rate, which is computed from data for the whole state. (IBGE, 1951.)

- **Log of Electricity Capacity.** Measured *circa* 1950. Original data in kilowatts. (IBGE, 1950.)

- **Fraction of population economically active.** Measured for population ten years and older for 1950. (IBGE, 1950.)

- **Shares of labor force by sector.** Fraction of economically active population in each of the following sectors: agriculture, extractive industries, manufacturing, transportation, and services. Measured for population ten years and older for 1950. (IBGE, 1950.)

Control variables for Colombian municipios

- **Region dummies.** The regions are as follows: Central, Bogota, Pacifico Norte, Eje Cafetero, Andina Norte, Andina Sur, Pacifico Sur, Caribe, Orinoquia, and Amazonia.

- **"La Violencia".** A qualitative variable (ranging from 1 to 3) indicating the intensity of violence in the Colombian civil war known locally at "La Violencia". The data are taken from Oquist, who classified conflict intensity decomposed by municipio and sub-period: before 1955, when the violence was largely in population centers, and 1955 and after, when the conflict was more likely to take place in the countryside. (Oquist, 1976.)

- **High Concentration** "**Minifundista".** Binary variable indicating the presence of small-land holders or minifundistas, as opposed to large land holders or urban areas. The reference period is the 1950s, although land-holding patterns were persistent historically. To construct municipio-level data, the map was digitized and georeferenced. Digital data on municipio boundaries, provided under special

agreement from the Centro Internacional de Agricultura Tropical (CIAT), was overlaid on the map and municipios were coded dichotomously as indicated by the map. The municipio boundaries of the CIAT data refer to 1993, and therefore these mapped back onto 1950s entities. (Banco de la República, 1964 (map 57); Jonnes and Bell, 1997; DANE, 2000; author's calculations.)

- **Coffee-growing Region.** Binary variable indicating the presence of coffee cultivation. The reference period is 1960. Municipio-level data were created using the process described above for the "minifundista" variable. (Banco de la República, 1964, map 38.)

- **Coal Mining Region.** Dummy indicating the presence of actively exploited coal deposits, *circa* 1960. Municipio-level data were created using the process described above for the "minifundista" variable. (Banco de la República, 1964, map 22.)

- **Expansion of Ranching.** Areas identified for possible expansion of ranching in 1960. Municipio-level data were created using the process described above for the "minifundista" variable. (Banco de la República, 1964, map 55.)

- **Infrastructure/Market Access.** An index variable for the ease of transport to major markets or seaports from the area, based on infrastructure in *circa* 1960. Six (ordered) categories are used, following the map's categorization. Municipio-level data were created using the process described above for the "minifundista" variable. (Banco de la República, 1964.)

- **Level of development.** An index variable for the general level of economic development of the area ("nivel de vida"), *circa* 1960. Six (ordered) categories are used, following the map's categorization. Municipio-level data were created using the process described above for the "minifundista" variable. (Banco de la República, 1964, map 59.)

- **Manufacturing employment per capita.** Computed by municipio from the 1945 Colombian census of manufacturing. (Dirección Nacional de Estadística, 1947.)

- **Disease controls.** The fractions of territory within each municipio in which transmission of the following diseases occurs: leishmaniasis, yellow fever, hookworm, and non-hookworm helminth diseases. The first two categories are vector-born diseases and could themselves have been affected by the campaign. The category of non-hookworm helminths represents an aggregate of numerous types of helminths. The underlying geographic data are defined with a fairly broad brushstroke, and as a result this is almost a dichotomous variable by municipio. These data come from a GIS file provided by the Center for International Development at Harvard University, and were computed by municipio in a GIS program (the "spatial join" in ArcView). (Gallup, Mellinger, and Sachs, 1999b; and author's calculations.)
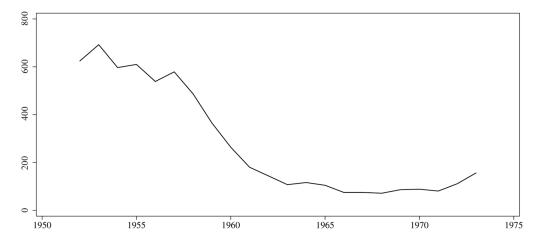
Control variables for the Mexican states

- **Region dummies.** "Norte", "Centro" and "Sur", as per the divisions in the 1960 *Anuario Estadístico*. (Dirección General de Estadística, 1960.)

- **Population Density.** Population per square kilometer in 1950. (Dirección General de Estadística, 1952a and 1952b.)

- **Infant mortality.** Rate per 1,000 births. Data refer to 1950. (Coordinación General de los Servicios Nacionales de Estadística, Geografía e Informatica, 1981.)

- **Log of Electricity Capacity.** Measured *circa* 1950. Original data in kilowatts. (Dirección General de Estadística, 1952b.)

- **Fraction of pop economically active.** Measured for population 12 years and older for 1950. (Dirección General de Estadística, 1952b.)

- **Shares of labor force by sector.** Fraction of economically active population in each of the following sectors: agriculture, extractive industries, manufacturing, transportation, and services. Measured for population 12 years and older for 1950. (Dirección General de Estadística, 1952b.)

Panel A: Large Decline in Malaria Following Onset of Spraying Campaign

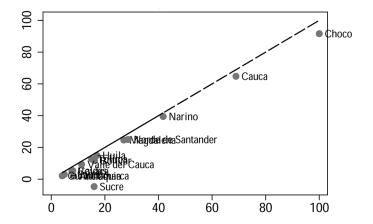Panel B: Highly Infected Areas Saw Greater Declines in Malaria

Figure 1.
Malaria Incidence Before and After the Eradication Campaign, Colombia
Notes: Panel A reports time-series data on notified cases of malaria for Colombia (SEM, 1979) for various years. Panel B shows the post-campaign decline in malaria mortality versus pre-campaign levels across Colombian departmentos. The *y* axis displays the estimated decrease in malaria mortality post-intervention. The *x* axis is the pre-campaign malaria mortality rate. (The 45-degree line represents complete eradication.) All variables are expressed per 100K population. SEM (1957) and the Colombian *Anuario de Salubridad* (DANE, 1968–70) are the sources for the departamento data. The malaria-eradication campaign in Colombia (like Mexico and Brazil) started mass spraying around 1957. (The time-series data required for this figure were not available in comparable form for the other three countries.)
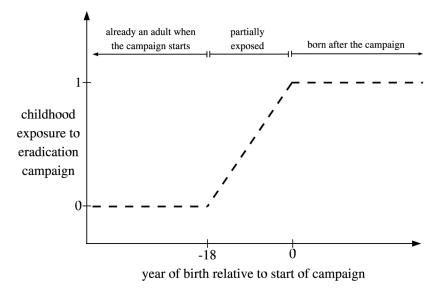
Figure 2.
Childhood Exposure to Eradication Campaign
Notes: This graph displays on the fraction of childhood that is exposed to a hypothetical (and instantaneous) campaign as a function of year of birth minus the start year of the campaign.

Figure 3.
Cross-Cohort Differences in Income versus Malaria
Notes: Each panel plots the cross-cohort change in income versus malaria for the four countries studied. The *y*-axes are the changes in the indicated income proxy from cohorts born well before the campaign (before 1895 in the U.S., and before 1940 in Brazil, Colombia, and Mexico) to cohorts born after the campaign (after 1920 in the U.S., and after 1957 elsewhere). The *x*-axis plots the malaria proxy for each country. Appendices A and B describe, respectively, the outcome variables and the malaria measure. Both variables are residuals from having projected the original data onto the controls from the "full controls" specification in the text and Appendix C. The dashed line is the best-fit regression line. State labels are left-justified on the corresponding coordinates.

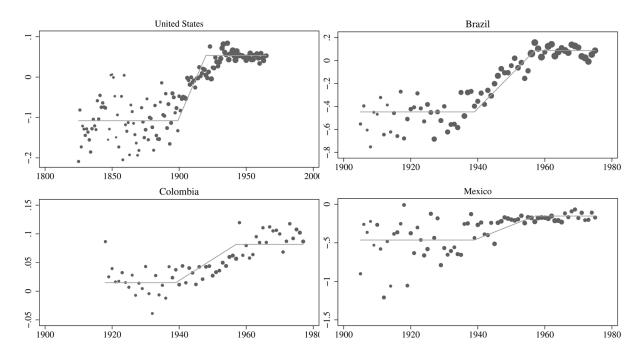**Figure 4.**
Cohort-Specific Relationships: Income and Pre-Campaign Malaria
Notes: These graphics summarize regressions of income on measures of malaria prior to eradication. The y axis for each graphic plots the estimated coefficient on pre-campaign malaria in one's place of birth. Income is measured in logs and the malaria proxy is scaled by the gap between the 95th and 5th percentile areas. The x axis is the cohort's year of birth. For each year-of-birth cohort, OLS regressions coefficients are estimated using the cross section of areas of birth and census years. Each cohort's point estimate is marked with a dot, and the size of the dots are proportional to the inverse of the standard error on that coefiicient. The dashed lines measure the approximate number of years of potential childhood exposure to the malaria-eradication activities. The regressions also include cohort-specific estimates of the mean-reversion variable, region dummies, and the 'full controls' specified in the text. These control variables are described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure.

Table 1

Cross-Cohort Differences and Malaria: United States

	Malaria Mortality (Fraction of Total), 1890		Malaria Ecology (Hong)	
Dependent Variable:				
Occupational Income Score	X		X	
Duncan's Socioeconomic Index		X		X
Additional Controls:	*Panel A: Alternative Control Sets*			
Basic Specification Only	0.112 *** (0.039)	0.134 ** (0.065)	0.236 *** (0.032)	0.219 *** (0.053)
Health	0.100 *** (0.038)	0.144 ** (0.067)	0.225 *** (0.031)	0.280 *** (0.048)
Education	0.136 *** (0.041)	0.131 ** (0.062)	0.219 *** (0.027)	0.206 *** (0.055)
Other	0.094 ** (0.044)	0.115 * (0.063)	0.204 *** (0.029)	0.178 *** (0.068)
Full Controls	0.110 ** (0.049)	0.172 * (0.094)	0.215 *** (0.049)	0.265 *** (0.096)
Instrumental Variables:	*Panel B: Estimates using Two-Stage Least Squares*			
The Other Malaria Proxy	0.142 *** (0.054)	0.175 ** (0.088)	0.207 *** (0.060)	0.244 ** (0.106)
Average Temperature and Altitude	0.154 * (0.083)	0.209 ** (0.104)	0.138 ** (0.059)	0.174 ** (0.075)
All of the Above Instruments	0.149 *** (0.054)	0.192 ** (0.095)	0.164 *** (0.052)	0.185 *** (0.071)
Sample/Specification:	*Panel C: Migration*			
Movers	0.111 ** (0.048)	0.136 * (0.074)	0.292 *** (0.042)	0.367 *** (0.070)
Nonmovers	0.107 ** (0.045)	0.165 * (0.100)	0.193 *** (0.040)	0.185 * (0.096)
Sample w/ parents' nativity data (exposed = born 1905–15)	0.064 (0.071)	0.220 * (0.116)	0.387 *** (0.092)	0.355 *** (0.123)
Sample with native-born fathers (exposed = born 1905–15)	0.069 (0.051)	0.232 * (0.128)	0.217 *** (0.079)	0.334 ** (0.159)
Use malaria of father's birthstate (exposed = born 1905–15)	0.121 (0.082)	0.215 * (0.126)	0.305 ** (0.124)	0.405 * (0.212)
Sample:	*Panel D: By Region*			
South + Bordering States (N=21)	0.100 *** (0.038)	0.173 *** (0.038)	0.208 ** (0.086)	0.203 (0.236)
Rest of the country (N=25)	0.183 (0.276)	0.134 (0.408)	0.162 (0.212)	0.281 (0.338)

Notes: This table reports estimates of equation 1 using OLS and 2SLS. The units of observation are U.S. states. The dependent variables are as indicated in the column headings and are defined as cross-cohort differences between exposed and unexposed cohorts. Robust (Huber-White) standard errors in parentheses, and the square root of the cell sizes are used to construct weights for the observations. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of constant term suppressed. Unexposed cohorts are those born before 1890 and fully exposed cohorts are those born after 1920 (except as noted in Panel C). Cohorts are determined based on state of birth. The universe for the base sample consists of the native-born white male population between the ages of 35 and 55 (15–55 for literacy) in the 1880–2000 census microdata from the IPUMS and NAPP databases. The specification for the basic results includes the malaria variable, a dummy for Southern birthplace, and the Lebergott (1964) measure of average unskilled wage in the state of birth in 1899. The health controls (row 2, Panel A) include fertility and infant mortality in 1890; late-1910s hookworm infection; state public-health spending and the number of doctors per capita in 1898; and the fraction of recruits rejected from service for health reasons by WWI-era Army physicians. The education controls (row 3, Panel A) consist of the 1910 adult literacy rate, and the logarithmic change (*circa* 1902–32) of teacher salaries, pupil/teacher ratios. The other controls (row 4, Panel A) are comprised of male unemployment rates from 1930, the 1910 fraction black, and the 1910 fraction living in urban areas. The "full controls" specification includes all of the above control variables simultaneously. For Panel C, state-level averages from microdata are computed separately according the indicated characteristics. Movers and nonmovers are defined by comparing an observation's state of residence versus state of birth. Data on father's state of birth is only available up through the 1940 Census, so the exposed (post-campaign) group includes cohorts with partial exposure (born in 1905–1915, inclusive), and the coefficient on exposure is adjusted to be in the same units (fraction of childhood potentially exposed) as the rest of the table. The bordering states in Panel D include DE, IL, IN, MD, MO, NM, OH, OK, and PA.

Table 2

Cross-Cohort Differences and Malaria: Brazil and Mexico

Dependent Variables: Differences across Cohorts in…	Brazilian States (N=24)				Mexican States (N=32)		
	Literacy	Education	Log Total Income	Log Earned Income	Literacy	Education	Log Earned Income
Specification:	*Panel A: Estimates using Ordinary Least Squares*						
Basic	0.063 (0.063)	0.555 (0.607)	0.351 ** (0.173)	0.267 ** (0.131)	0.116 *** (0.032)	0.058 (0.298)	0.292 *** (0.112)
Include Infant Mortality	0.063 (0.063)	0.576 (0.581)	0.366 ** (0.147)	0.262 * (0.136)	0.119 *** (0.032)	0.138 (0.237)	0.286 ** (0.112)
Include Sectorial Shares	0.131 *** (0.042)	1.288 ** (0.597)	0.434 ** (0.183)	0.283 *** (0.094)	0.032 (0.039)	−0.234 (0.247)	0.196 (0.135)
Full Controls	0.147 *** (0.042)	0.995 ** (0.487)	0.393 ** (0.178)	0.283 * (0.147)	0.035 (0.035)	−0.247 (0.260)	0.254 * (0.148)
Specification:	*Panel B: Estimates using Two-Stage Least Squares (Temperature and Altitude Instruments)*						
Basic	0.225 (0.215)	−1.356 (2.162)	0.649 * (0.335)	0.434 (0.335)	0.128 ** (0.058)	0.112 (0.648)	0.494 ** (0.196)
Full Controls	0.215 * (0.120)	0.257 (0.979)	0.785 * (0.414)	0.497 (0.330)	0.048 (0.042)	−0.234 (0.510)	0.398 ** (0.176)

Notes: This table reports estimates of malaria in equation 1 using OLS and 2SLS. The units of observation are Brazilian and Mexican states. The dependent variables are as indicated in the column headings and are defined as cross-cohort differences between exposed and unexposed cohorts. Robust (Huber-White) standard errors in parentheses, and the square root of the cell sizes are used to construct weights for the observations. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of additional estimates is suppressed. Unexposed cohorts are those born before 1940 and fully exposed cohorts are those born after 1957. Cohorts are determined based on state of birth. The universe for the base sample consists of the native-born male population between the ages of 25 and 55 (15–55 for literacy) in the 1960–2000 census microdata from the IPUMS. The malaria measure for Brazil is Mellinger's ecology variable, while for Mexico it is malaria mortality circa 1950 (Pesqueira, 1957). All regressions include dummies for region of birth, population density, and log electrical capacity as a proxy for economic development. The regression in row 2 of Panel A includes infant mortality in 1950. The controls in row 3 of Panel A are 1950 composition of the labor force: the fraction of the population that is economically active, as well as employment shares in agriculture, extractive industries, manufacturing, transportation, and services. The "full controls" specification includes all of the above controls. Excluded instruments in Panel B are state-of-birth average temperature, altitude, and the interaction of the two. Appendices A and B describe, respectively, the outcome variables and malaria measures. The additional controls are described in the text and Appendix C.

Table 3

Cross-Cohort Differences and Malaria: Colombia

Dependent Variables; Differences across Cohorts in…	Malaria Ecology (Poveda)			Malaria Ecology (Mellinger)			Cases Notified, 1955, per 100K Pop.		
	Literacy	Years of Schooling	Income Index	Literacy	Years of Schooling	Income Index	Literacy	Years of Schooling	Income Index
Additional Controls:									
Panel A: Alternative Controls									
None (Basic Specification)	0.035 *** (0.013)	0.168 * (0.088)	0.065 *** (0.011)	0.071 *** (0.016)	0.064 (0.108)	0.048 *** (0.014)	0.018 * (0.010)	0.077 (0.067)	0.026 *** (0.008)
Conflict	0.032 *** (0.012)	0.175 * (0.090)	0.063 *** (0.011)	0.068 *** (0.016)	0.068 (0.110)	0.046 *** (0.014)	0.016 (0.010)	0.078 (0.067)	0.024 *** (0.008)
Economic Activity	0.008 (0.010)	0.194 ** (0.089)	0.057 *** (0.012)	0.043 *** (0.013)	0.156 (0.110)	0.039 *** (0.014)	0.008 (0.008)	0.071 (0.067)	0.022 *** (0.006)
Other Diseases	0.024 * (0.013)	0.180 ** (0.089)	0.065 *** (0.012)	0.058 *** (0.016)	0.057 (0.114)	0.042 *** (0.015)	0.012 (0.010)	0.067 (0.067)	0.020 ** (0.008)
Full Controls	0.006 (0.011)	0.165 * (0.095)	0.064 *** (0.013)	0.046 *** (0.015)	0.076 (0.117)	0.034 ** (0.015)	0.008 (0.008)	0.041 (0.067)	0.016 *** (0.006)
Instrumental Variables:									
Panel B: Alternative Instrument Sets									
Temperature, Altitude, and their Interaction	0.037 ** (0.018)	0.372 *** (0.136)	0.092 *** (0.017)	0.067 * (0.036)	0.766 *** (0.268)	0.170 *** (0.037)	0.055 (0.035)	0.679 ** (0.284)	0.159 *** (0.047)
The Other Two Malaria Proxies	0.126 *** (0.032)	0.113 (0.190)	0.084 *** (0.026)	0.082 ** (0.029)	0.390 * (0.203)	0.149 *** (0.028)	0.200 *** (0.069)	0.145 (0.312)	0.131 *** (0.051)
Holdridge Climate Zone	0.045 ** (0.021)	0.303 * (0.159)	0.102 *** (0.020)	0.082 ** (0.037)	0.593 ** (0.248)	0.124 *** (0.035)	0.006 (0.059)	−0.483 (0.508)	0.088 (0.061)
All of the Above Instruments	0.049 *** (0.017)	0.323 *** (0.122)	0.092 *** (0.016)	0.074 *** (0.026)	0.516 *** (0.184)	0.120 *** (0.025)	0.059 ** (0.026)	0.131 (0.179)	0.075 *** (0.026)

Notes: This table reports estimates of malaria in equation 1 using OLS and 2SLS for the indicated dependent variables. The dependent variables are as indicated in the column headings and are defined as cross-cohort differences between exposed and unexposed cohorts. The units of observation are Colombian municipios. The malaria variables are as indicated in the column headings. Robust (Huber-White) standard errors in parentheses, and the square root of the cell sizes are used to construct weights for the observations. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of additional estimates is suppressed. Unexposed cohorts are those born before 1940 and fully exposed cohorts are those born after 1957. Cohorts are determined based on municipio of birth. The universe for the base sample consists of the native-born male population between the ages of 25 and 55 (15–55 for literacy) in the 1973 and 1993 census microdata from the IPUMS. All regressions include dummies for region of birth, log population density and a measure of the general level of economic development, *circa* 1960. Row 2 of Panel A adds controls for the intensity of the 1950s civil war in Colombia and the concentration of land holdings. Row 3 of Panel A includes controls for manufacturing employment per capita, proximity to major markets, and dummies for being in the coffee, coal-mining, and cattle-ranching areas. The "other diseases" specification includes the fraction (circa 1950) of each area in which the following diseases occur: leishmaniasis, yellow fever, hookworm, and non-hookworm helminth diseases. The final row of Panel A reports estimates controlling for all of the above controls simultaneously. Excluded instruments for 2SLS in Panel B are indicated in the row label. Appendices A and B describe, respectively, the outcome variables and malaria measures. The additional controls and instruments are described in the text and Appendix C.

NIH-PA Author Manuscript   NIH-PA Author Manuscript   NIH-PA Author Manuscript

Table 4

Panel Estimates of the Effect of Childhood Exposure

Degree of Polynomial Trend for Year of Birth:	(1)	(2)	(3)	(4)	(5)	(6)
	Mean Reversion and Region Controls			Additional Controls		
	0	1	2	0	1	2
Specification:						
Panel A: United States, Log Occupational Income Score						
Baseline	0.132 *** (0.030)	0.115 *** (0.031)	0.131 *** (0.025)	0.120 *** (0.024)	0.098 *** (0.035)	0.116 *** (0.027)
Post-1920 Break in Birthplace Time Trend	0.082 *** (0.015)	0.094 *** (0.020)	0.105 *** (0.024)	0.073 *** (0.020)	0.080 *** (0.021)	0.085 *** (0.021)
Allow for Birthplace x Time Effects	0.103 *** (0.026)	0.110 *** (0.030)	0.123 *** (0.023)	0.089 *** (0.030)	0.092 *** (0.033)	0.108 *** (0.025)
Drop early census years (<1930)	0.106 *** (0.021)	0.105 *** (0.017)	0.032 ** (0.015)	0.096 *** (0.014)	0.109 *** (0.023)	0.032 * (0.019)
Add region x year x YOB effects	0.131 *** (0.030)	0.116 *** (0.029)	0.131 *** (0.024)	0.123 *** (0.025)	0.102 *** (0.034)	0.119 *** (0.027)
Panel B: Brazil, Log Total Income						
Baseline	0.287 ** (0.134)	0.249 ** (0.102)	0.269 *** (0.103)	0.341 *** (0.106)	0.335 *** (0.085)	0.403 *** (0.093)
Allow for Birthplace x Time Effects	0.134 (0.088)	0.214 * (0.120)	0.202 * (0.115)	0.136 * (0.070)	0.294 *** (0.104)	0.311 *** (0.110)
Drop earliest census year (1960)	0.267 *** (0.099)	0.122 (0.116)	0.133 (0.117)	0.308 *** (0.062)	0.347 *** (0.053)	0.349 *** (0.069)
Add region x year x YOB effects	0.332 ** (0.135)	0.200 * (0.117)	0.246 ** (0.121)	0.366 *** (0.110)	0.324 *** (0.091)	0.411 *** (0.092)
Panel C: Colombia, Log Industrial Income Score						
Baseline	0.052 *** (0.010)	0.046 *** (0.010)	0.046 *** (0.010)	0.054 *** (0.010)	0.048 *** (0.010)	0.048 *** (0.010)
Allow for Birthplace x Time Effects	0.047 *** (0.009)	0.046 *** (0.010)	0.046 *** (0.010)	0.050 *** (0.010)	0.048 *** (0.010)	0.049 *** (0.010)
Add region x year x YOB effects	0.099 *** (0.019)	0.076 *** (0.019)	0.077 *** (0.019)	0.081 *** (0.020)	0.068 *** (0.020)	0.068 *** (0.020)
Panel D: Mexico, Log Earned Income						
Baseline	0.214 ** (0.091)	0.007 (0.059)	0.026 (0.049)	0.157 * (0.085)	0.155 *** (0.048)	0.157 *** (0.047)
Allow for Birthplace x Time Effects	0.137 *** (0.035)	0.013 (0.057)	0.022 (0.053)	0.146 *** (0.025)	0.146 *** (0.063)	0.152 ** (0.060)
Drop earliest census year (1960)	0.104 ** (0.042)	0.021 (0.033)	0.110 (0.144)	0.129 *** (0.047)	0.104 ** (0.042)	0.085 (0.097)
Add region x year x YOB effects	0.354 *** (0.110)	−0.069 (0.079)	−0.037 (0.060)	0.203 * (0.110)	0.104 (0.066)	0.121 ** (0.055)

Notes: Each cell reports estimates, from a separate regression, of the childhood-exposure variable times pre-campaign malaria intensity using OLS and 2SLS. The outcome variables are the income proxies indicated the panel heading. The malaria proxies and control variables, which enter the specification interacted with potential childhood exposure to the anti-malaria campaign, are described in Appendices B and C. These parameters are estimated from three-dimensional cohort/panel data structured by year of birth, census year, and area of birth, the construction of which is described in Appendix A. The regression uses equation 2, and the table reports estimates of . The columns marked "mean reversion and region controls" use the basic controls sets, while the "additional controls" columns use the "full controls" specifications. These controls are described in the tables above, as well as the text and Appendices. Standard errors, shown in parentheses, are clustered on area of birth. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Observations are weighted by the square root of the size of the cell used to construct the cohort×census-year average. Reporting of additional terms is suppressed.

Table 5

Approximate Effects on Adult Income per Probability of Childhood Malaria Infection

Dependent Variables:	United States		Brazil		Colombia	Mexico
	Occupational Income Score	Duncan's Index	Log Total Income	Log Earned Income	Industrial Income Score	Log Earned Income
Estimates:						
Reduced-form Differences; 95/5 percentile comparison	0.14	0.18	0.37	0.27	0.28 (adjusted)	0.26
Maximal Endemicity (approx. Malaria Infection Rate)	Mesoendemic (0.3)		Hyperendemic (0.625)		Hyperendemic (0.625)	Hyperendemic (0.625)
Income Effect per Probability of Childhood Infection	0.47	0.60	0.59	0.45	0.45	0.41

Notes: The reduced-form differences are taken from Tables 1, 2, and 3, for the indicated outcomes. Reported are averages of the OLS point estimates, divided by the difference between 95th and 5th percentile malaria intensity. Maximal endemicity levels and approximate malaria-infection rates are according to Molineaux (1988, p. 988 and Figure 35.10). Note that these numbers refer to pre-eradication malaria burdens for children. The effect on adult income per probability of childhood infection is the reduced-form difference divided by the estimated pre-eradication infection rate for malaria.