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

Hoyt Bleakley†

August 17, 2007

Abstract

This study considers the malaria-eradication campaigns in the United States (circa 1920), and in Brazil, Colombia and Mexico (circa 1955) in order 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 changes coincided with childhood exposure to the campaigns rather than to pre-existing trends.

Keywords: Malaria, returns to health, eradication campaigns.
JEL codes: I12, J24, 010, H43.

*This study subsumes two earlier draft papers, "Malaria and Human Capital: Evidence from the American South" and "Malaria Eradication in Colombia and Mexico: A Long-Term Follow-up." The author gratefully acknowledges the contribution of many in preparing this study. Funding support came from the University of California Pacific Rim Research Program, the UCSD Faculty Senate's Committee on Research, the UCSD Hellman Faculty Fellowship, and the Graduate School of Business of the University of Chicago. This paper was partially prepared while the author was visiting the Center for US/Mexican Studies at UCSD and the Universidad de los Andes. Imelda Flores, Dr. Mauricio Vera, and Dr. Victor Olano of Colombia's Instituto Nacional de Salud provided useful guidance in interpreting the Colombian malaria data. Glenn Hyman of the Centro Internacional de Agricultura Tropical shared data on the Colombian municipio boundaries. Andrew Mellinger provided the malaria-ecology raster data. Librarians from UC-San Diego, la Universidad Nacional de Colombia, la Fundación Santa Fe de Bogotá, and the University of Chicago have consistently gone the extra mile in aiding my search for data. I have also benefited from the comments of colleagues too numerous to mention, and seminar participants at UC-San Diego's Center for US/Mexican Studies, Universidad de los Andes, Yale University, the University of Chicago, Columbia University, MIT, the NBER Cohort Studies group, Princeton University, the Stanford SITE conference, Harvard Program on the Global Demography of Aging, the University of Illinois at Chicago, University College London, the University of Bristol, and the CID/Gates Foundation Conference on "Health Improvements for Economic Growth". Jennifer Baca, Rebeca Mohr, Michael Pisa, Tareq Rashidi, Lisandra Rickards, and Raghavan Selvara all provided able research assistance.

†Assistant Professor of Economics, Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL, 60637. Telephone: (773) 834-2192. Electronic mail: bleakley[at]chicagogsb[dot]edu

1 Introduction

The disease known as malaria, a scourge of mankind through 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 cut through this Gordian knot of 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 U.S., 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 U.S. 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.) Additionally, sufficient time has passed that we can evaluate the long-term consequences of eradication.

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. Moreover, the reductions in disease burden occur in the space of a few years, and resulted from critical innovations to knowledge and spending. These innovations came largely from outside the studied areas. This latter fact mitigates the usual concern about policy endogeneity.

[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), Acemoglu and Johnson (2006), and Weil (2007) for evidence on the broader health/income link across countries.

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. 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 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. This mechanism would drive increases in years of schooling, and presumably literacy and income would follow. Finally, malaria's possible effect on contemporaneous wages implies that an additional channel is via parental income.

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 area 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 from all four countries are used to construct data on a panel of cohorts by birth year and birthplace. Pre-campaign malaria intensity across areas is estimated from a variety of sources. (The identification strategy is further described in Section 2.4, while the data construction is detailed in Section 3 and the appendices.)

Cohorts more exposed to the eradication efforts went on to earn higher incomes and have higher literacy rates as adults. Mixed results are found for years of schooling, consistent with the economic theory of schooling. In Section 4, I present estimates of regressions of adult outcomes on pre-campaign malaria for each year of birth in the sample. Graphs of the cohort-specific coefficients reveal a shift in the malaria-income relationship that coincides with childhood exposure to the eradication efforts. Furthermore, statistical tests tend to favor this childhood-exposure hypothesis over pre-existing time-series processes.

These results are not sensitive to accounting for a variety of alternative hypotheses. In Section 5, I show that this shift was systematic across areas of these countries, and not due to a few outliers. Moreover, I obtain essentially similar estimates of malaria coefficients even when controlling for different indicators of health and economic development. In Section 6.2, I present evidence 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 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 10 to 40 percent. 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 suggest persistent childhood malaria infection reduces adult income 40 to 60 percent. In Section 6.3, I show that these estimates explain a modest fraction of the gap between the areas under study and more developed economies.

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. The main 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 U.S., circa 1920

The turn of the 20th century saw considerable advances in the scientific understanding of the disease. Doctor 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. (later Sir) 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.

The U.S. 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 the acquisition of Cuba and of the Panama Canal Zone. Early in the occupation of Cuba, the U.S. 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.[2]

The progress made by U.S. Army doctors against malaria 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.[3] 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.

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

[2]It is doubtful that the construction of the Canal would have been economically feasible were it not for these sizable innovations to knowledge. The following anecdote is illustrative of the primitive state of medical knowledge about malaria just a few years earlier:

> And all the while, in the lovely gardens surrounding the hospital, thousands of ring-shaped pottery dishes filled with water to protect plants and flowers from ants provided perfect breeding grounds for mosquitoes. Even in the sick wards themselves the legs of the beds were placed in shallow water, again to keep the ants away, and there were no screen in any of the windows or doors. Patients, furthermore, were placed in the wards according to nationality, rather than by disease, with the results that every ward had its malaria and yellow-fever cases. As Dr. Gorgas was to write, had the French been consciously trying to propagate malaria and yellow fever, they could not have provided conditions better suited for the purpose. (McCullough, 1977)

History records that the French effort to build a canal across the isthmus did indeed fail, in part because of malaria. Moreover, the American effort to build the Canal did not get off the ground until malaria was under control.

[3]Williams (1951) presents a thorough history of the U.S. Public Health Service.

in all but the most peripheral areas of the region (Williams, 1951).

During this period, the South experienced a substantial decline in malaria. Malaria mortality per capita is seen in Panel A of Figure 1. Apart from a hiccup in the first years of the Depression, the region saw a drop of around 50 percent in the 15 years after WWI.

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.[4] 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 (transmitted by lice) and later malaria. 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 a worldwide campaign to eradicate malaria in the late 1940s and early 1950s. 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.[5] Its purpose was not to kill every mosquito in the land, but rather to interrupt the transmission of malaria for long enough that 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) all mounted malaria eradication campaigns, and all saw large declines in malaria prevalence. Panel B of Figure 1 shows malaria cases per capita in Colombia. 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.

[4]The historical narrative on the worldwide campaign is drawn from Harrison (1978). For a review of health programs in Latin America prior to the 1950s, see Cueto (1994, 2004) and Quevedo (2005).

[5]The effect of DDT was sufficiently persistent after spraying that only 1-3 applications per year were mandated.

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 U.S. 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. This is a further advantage of examining these anti-malaria campaigns: enough time has passed since their inception that we can assess their long-term consequences.

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 U.S. 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. (See Figure 1.) 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 supposed effect of reducing infection rates *more* in highly infected areas than in areas with moderate infection rates. The decline in malaria incidence as a function of intensity prior to the eradication campaign is found in Figure 2. 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 across areas in the countries where data are available.

[6] Humid areas with slow-moving water were the preferred habitat for mosquitoes, the vector that transmits malaria.

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 3. 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. In reality, the campaigns took effect over the better part of a decade. This implementation delay would tend to elongate the line in Figure 3 (if measured as *effective* childhood exposure to the campaign), inasmuch as being born at the time of the commencement of the campaign does not actually bring the full benefits of the eradication efforts. Note, however, that the left-hand side of the curve (representing those already in adulthood before the campaign started) would remain unchanged. In any case, the *effects* of partial childhood exposure might not take on the straight-line form suggested in the graph, depending on the relative importance of malaria infection at various points in childhood. For example, if the most sensitive years are in adolescence (when the child is near the margin of staying in school or entering the workforce), the line in Figure 3 would need to be adjusted to rise more sharply for those cohorts in adolescence at the time of the campaign. If, in contrast, the most deleterious point in the life cycle to have malaria is as an infant, the impact-adjusted line in the figure would rise most sharply for those born just before the campaign. For the moment, I maintain the uniform (agnostic) weighting of childhood exposure, but I discuss in Section 6.2 what we learn on this point from the empirical results.

2.5 Related Literature

Over a century has passed since the birth of the interdisciplinary field of malariology, and the associated literature has grown to Brobdingnagian proportions since that time. A thorough reference is provided by Wernsdorfer and McGregor (1988), who edited an encyclopedic tome on medical and social aspects of the disease. There is also an excellent survey by Nájera, Liese, and Hammer (1992).

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?[7] For example,

[7] A related approach is to measure and then value the excess mortality from malaria in the framework of compensating and equivalent variation. A number of studies treat the effect of malaria on other aspects of mortality. There was a considerable debate in the social sciences about the role of the anti-malaria campaign in reducing general mortality in one episode: Sri Lanka, *circa* 1950. Gray (1974), in a study attempting to rectify the competing positions, finds that a substantial fraction of the mortality decline in Sri Lanka coincided with the malaria-eradication program. More recently, Acemoglu and Johnson (2006) show that the decline in malaria in the 1950s increased life expectancy

Conly (1975) presents such an analysis in Paraguay, while Bonilla Castro, Kuratomi, Rodríguez, and Rodríguez (1991) perform this exercise in Colombia. These studies also consider spillovers within the household (e.g., parents' caring for sick children). 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. (See Malaney, 2003, for a thorough discussion.) 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. Utzinger, Tozan, Doumani, and Singer (2001) argue, for example, that the control of malaria transmission was a key factor in the development of Zambian copper mining. Furthermore, as mentioned above, the control of malaria in the Panamian isthmus was crucial for the successful completion of the canal. 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.[8]

Several studies also consider the role of broadly defined health in Latin America. In the interest of space, I mention only a few examples here. Lopez Alonso and Porras Condey (2003) and Meisel (2004) use data on stature to analyze changes in health across birth cohorts in Mexico and Colombia, respectively. Sanchez and Nuñez (2000) and Mendoza and Rosas (2004) consider the role of geographic and health factors in explaining cross-municipio income differences within Colombia at a point in time. Additionally, using local resources in childhood as an instrument for adult height, Ribero and Nuñez (2000) analyze the effect of health endowments on income in Colombia, while Miller (2005) studies the impact of family planning in that same country.

The present study is closely related to several recent studies. Bleakley (2007a) finds that hookworm eradication in the U.S. South was followed by an increase in school attendance and literacy. Furthermore, using a retrospective/cohort design similar to the present study, he finds evidence that childhood exposure to the hookworm-eradication campaign increased adult income. Using a database of Union Army veterans, Hong (2007) considers the effect of early-life exposure to malaria on later-life health outcomes in 19th century America. Lucas (2005) shows that women born after malaria eradication in Sri Lanka completed more years of schooling, suggesting that returns to education rose faster than child wages in that episode. Similarly, Cutler, Fung, Kremer, and Singhal (2007) find malaria exposure in early life raised educational attainment in India. None

in malarious countries. Their study does not, however, analyze the impact of malaria on productivity separately from other mortality risks.

[8]This point is made by Barlow (1967), Wernsdorfer and Wernsdorfer (1988), and Gallup and Sachs (2001), but none of these studies evaluates the empirical magnitude of this channel.

consider the direct impact of childhood malaria on income.[9]

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 U.S. and several other countries (Ruggles and Sobek (1997); Sobek *et al.* (2002)). I analyze the census data from the U.S., 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 area of birth to conduct the analysis, which is therefore an intention-to-treat design. For the U.S., 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 U.S. counties.

For the United States, the base sample consists of native-born white 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 1965.[10] 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 due to Duncan (1961) 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.

[9]There are a few recent exceptions to this characterization. One is by Barreca (2007) who studies the effect on later-life outcomes of *in utero* exposure to short-term fluctuations in malaria caused by within-year variation in rainfall and temperature in the U.S. He finds positive results for education, but the results for adult income are noisy. In addition, interannual variation in malaria is likely to have economic impacts that differ from those due long-term changes induced by eradication. Another example is the aforementioned study by Cutler *et al.* (2007), which considers wages tangentially, although their sample size is considerably smaller than that of the present study, and the income results are imprecise.

[10]The choice of the lower age bound of 35 for the U.S. 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).

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 benchmarked from the Mexican and Brazilian data.

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 $\times$ census year $\times$ place of birth. (Cohorts can appear in multiple censuses in this pooling strategy, so the resulting dataset is not, strictly speaking, a panel.) In Section 4, I consider how cross-area outcomes change by year of birth, so the panels are constructed with year of birth $\times$ area of birth $\times$ census year as the units of observation. For Section 5, I compare two groups—cohorts born well before or just after the campaign—so the averages by period of birth are computed accordingly.

Malaria data are drawn from a variety of sources. U.S. data are reported by the Census (1894), Maxcy (1923), and later in the *Vital Statistics* (Census, 1933). Mexican data are drawn from Pesqueira (1957) and from 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 Mellinger *et al.* (2004) and Poveda *et al.* (2000). The ecology data were matched with states and municipios using a geographic information system (GIS). 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 for each country in Appendix C. Additionally, Appendix C–1 reports correlations of the malaria measures with each control variable. The main results that follow from these correlations are that malarious areas in these countries were (i) less developed and (ii) closer to the equator (except in Colombia).

4 Cohort-specific Results

The shift in the malaria-income relationship coincides with childhood exposure to the eradication efforts. This can be seen graphically in this section. I also provide statistical tests that indicate the break is indeed coincident with exposure to eradication rather than with some simple time-series process. 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:

$$Y_{jkt} = \beta_k \, \mathrm{M}_j + \delta_k + X_j \, \Gamma_k + \nu_{jkt} \tag{1}$$

in which M_j is the pre-campaign malaria intensity in area of birth j, β_k is year-of-birth-specific coefficient on malaria, X_j is a vector of other state-of-birth controls, and δ_k and Γ_k are cohort-specific intercept and slope coefficients. I estimate this equation using OLS for each year of birth k, thus generating a series of estimates across cohorts. (The cohort outcomes are estimated with differing degrees of precision, so the square root of the cell sizes is used as a weight when estimating equation 1.) This specification allows one to examine how the relationship between income and pre-eradication malaria ($\hat{\beta}_k$) 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.)

I start with a simple graphical analysis using the flexible specification (equation 1) for cross-cohort comparison. Figures 4 and 5 display plots of the estimated β_k, for the various income measures and 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.

Results for the U.S. are shown in Figure 4, which displays the coefficient on state-of-birth 1890 malaria mortality for each year of birth. The top row of this figure is denoted "basic specification", and presents estimates of β_k in equation 1 that are produced controlling flexibly for being born in the South and 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 U.S.: the occupational income score and the Duncan socioeconomic indicator.

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. Observe that U.S. 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[11] to the malaria-eradication campaign. (The line is rescaled such that pre-1895 and post-1925 levels match those of the $\hat{\beta}_k$. The exposure line is not rescaled in the x dimension.)

The estimates in Figure 4 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 malariousity in one's birthplace and adult income.[12] 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, which is counterfactual. As seen in Figure 1, there was considerable progress against malaria in the 1920s, but it did take a number of years for these declines to be realized.

These results are not sensitive to including a variety of additional, state-of-birth controls. These results are found in the bottom row of Figure 4. In addition to the South dummy and wage variable, the summarized regressions flexibly control for health conditions and educational resources. The health controls include 1890 infant mortality; late-1910s hookworm infection; and state public-health spending and the number of doctors per capita in 1898. The education-related controls are as follows: the 1910 adult literacy rate, and the logarithmic change (*circa* 1902–32) of pupil/teacher

[11] Specifically, the formula is $\text{Exp}_k = \max(\min(21, k - (1920 - 21)), 0)$, which treats 1920 as an approximate start date for exposure. Because the campaigns had their effect over a decade or more, the childhood-exposure measure represents an optimistically fast guess.

[12] 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 sample variance in the estimates of cohort-level means. See Appendix Figure A–1.

ratios and school term lengths. Moreover, the regressions include the male unemployment rates from 1930, the 1910 fraction black, and the 1910 fraction living in urban areas. (Appendix C has details on these variables. Section 5 below considers the sensitivity of these results to the choice of control sets. Further, Appendix D contains plots of the year-of-birth-specific coefficients on each of the control variables.) The broad shape of these curves is similar in the top and bottom rows of the figure. If anything, adding these controls actually increases the gap between exposed and unexposed cohorts.

Formal statistical tests indicate that the shift in the income/malaria ecology relationship coincided with exposure to malaria eradication, rather than with some polynomial trend or autoregressive process. This can be seen by treating the estimated β_k as a time series and estimating the following regression equation:

$$\hat{\beta}_k = \alpha \ \text{Exp}_k + \sum_{i=1}^{n} \gamma_n k^n + \Phi(L)\hat{\beta}_k + \text{constant} + \epsilon_k^{ts} \tag{2}$$

in which Exp_k is exposure to the malaria-eradication campaign (defined above), the k^n terms are nth-order trends, and $\Phi(L)$ is a distributed-lag operator. To account for the changing precision with which the generated observations are estimated, observations are weighted by the inverse of the standard error for $\hat{\beta}_k$. Table 1 reports estimates of equation 2 under a variety of order assumptions about trend and autoregressive processes. The dependent variables are the cohort-specific regression estimates that are shown in the figure above.[13] Panel A of Table 1 contains estimates for the United States. For each income variable, I present estimates of α, the effect of childhood exposure to the campaign. The "basic" and "additional controls" specifications as the same as those in Figure 4. For the occupational income score, the estimates on the exposure term are qualitatively similar across specifications. When the Duncan socioeconomic index is used instead, there is evidence of a slight downward trend, but estimates of the exposure coefficient are stable once this is accounted for.

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

[13] A word about the standard errors. Simply treating the $\hat{\beta}_k$ as data might result in incorrect inferences for two reasons: (1) the $\hat{\beta}_k$ are in fact estimates, and (2) the various year-of-birth cohorts within a given state of birth are not likely to be independent observations. To deal with these problems, the point estimates and standard errors throughout Table 1 are computed by 1,000 iterations of a block bootstrap, where state of birth is used as the clustering variable. This procedure does indeed result in standard errors that are around 30% larger than the uncorrected errors. As a check on this, I consider a complementary approach using the full dataset: I estimate the interaction of pre-campaign malaria and potential childhood exposure in one step, but adjust the standard errors for clustering at the state-of-birth level. I adopt this strategy in Appendix E. The point estimates are similar to those reported here. Relative to Table 1, the standard errors using this latter method tend to be larger for Brazil and the U.S. and smaller for Colombia and Mexico, but, in any case, are never different enough to affect the inferences here.

unit increase in the independent variable. Recall that the malaria measure is renormalized by the difference across the 95th and 5th percentile states.[14] 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 10-15% more than the previous generation, relative to the benchmark of cohorts in malaria-free states.

In the Latin American countries, childhood exposure to the malaria-eradication campaign is associated with higher income as well. To see this, I estimate equation 1 using the census data from Brazil, Colombia, and Mexico. Figure 5 plots the estimated income-malaria relationship for each year-of-birth cohort (the $\hat{\beta}_k$, as before). The analysis is conducted separately for each country (across the columns) and for two different specifications (shown by row). Potential childhood exposure to the eradication efforts is plotted as a dashed line in the figure. The start of large-spraying spraying of DDT was in the mid-to-late 1950s, so 1957 is chosen as an approximate start date. The range of years of birth is narrower for these countries (70 instead of 140 for the U.S.) because the available census data span fewer years.

Graphical results for Brazilian states are shown in the first column of Figure 5. Because data on pre-eradication malaria prevalence is limited, I use an index of malaria ecology (based on Mellinger *et al.*, 2004). The outcome variables is the logarithm of total income. The basic specification, like for the U.S., includes flexible controls for region and a proxy for pre-campaign income. The log of population density and log electrical capacity *circa* 1950 are used as a control for the pre-program level of development.[15] The specification with additional controls from *c.* 1950. On the employment side, I include 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. On the health side of things, infant mortality in 1950 is used to control for possible catch-up in general health. Recall that in equation 1 these control variables also enter into the specification very flexibly: estimated coefficients are computed for each year of birth. (Sources and definitions of these variables are found in the appendices. Moreover, Appendix D plots the estimated coefficients on the control variables for the augmented specification.)

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

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

[15]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 (2006), gross domestic product per capita and electricity production per capita (in logs) have a correlation coefficient of 0.83 across countries.

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. This pattern of coefficients is quite similar whether one considers the basic or augmented specification (top or bottom row of the figure).

The pattern of coefficients for 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 1. The first two rows of this panel contain estimates of childhood exposure from equation 2 using the two series displayed for Brazil in Figure 5. Point estimates are broadly consistent across each row. Estimates of the exposure effect are statistically significant even when controlling for up to a second-degree polynomial trend, but standard errors become quite large when including both a quadratic trend and second-order autoregression. (It bears mentioning the span of years in the Latin American data is much shorter than the range for the U.S., so horse-racing the exposure variable with second-degree trends and autoregressions is a more difficult test to pass.) The third row repeats the estimates above, but uses the log of earned income as the outcome variable.

Graphs of the malaria coefficients for Colombia are found in the second column of Figure 5. Results from two specifications are presented. The basic specification contains log population density and a measure of the general level of economic development, *circa* 1960, from the Colombian Banco de la República (1964). The "additional controls" specification includes birth-region dummies, along with a variety of supplementary variables. The additional economic controls are manufacturing employment per capita, proximity to major markets, and dummies for being in the coffee, coal-mining, and cattle-ranching areas. To control for health differences across areas, I 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-born diseases and would themselves have been affected by the campaign. I return to this issue below.) Finally, I include measures of the intensity of violence in the Colombian civil war known locally at "La Violencia" and a proxy for land concentration. The unit for area of birth for this analysis is the municipio, and an ecological index is used for malaria intensity. Log income is proxied by a score based on industry and class of worker, calibrated using data from Brazil and Mexico.

Results for Colombia point to an effect of childhood exposure to malaria on adult income. The estimated malaria coefficients tend to be higher for those fully exposed to the eradication campaign. Moreover, the series of coefficients shift across cohorts with increasing childhood exposure to the eradication efforts. Estimates of equation 2 for Colombia are found in Panel C of Table 1. The Colombian data span 15 fewer years than the Brazilian and Mexican samples, and results are less robust to the inclusion of second-order trends, particularly for the full-controls specification. One anomaly, however, for the Colombian results is that the coefficients for the earlier cohorts tend to be positive.

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 third column of Figure 5. The outcome variable is log earned income. The basic specification includes mean-reversion controls (logs of population density and electrical capacity), while the "additional controls" include variables similar to those used for Brazil (region dummies, sectoral shares, infant mortality, etc.) 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. Estimates of equation 2, summarized in Panel D of Table 1, confirm this positive relationship between exposure and income. Nevertheless, in the microdata, sample sizes for the earlier cohorts are quite small, and the imprecision with which the pre-1940 coefficients are estimated complicates time-series analysis for the Mexican case. In Section 5 below, to ameliorate this problem, I consider estimates in the next section based on pooling the data across broader ranges of birth years.

These estimates suggest quite similar reduced-form impacts 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 U.S. South., which is consistent with the likelihood that pre-eradication malaria infection rates were lower in the Southern U.S. 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.

Although the estimates above are constrained to work through a particular type of exposure, the results are not sensitive to relaxing this assumption. The first issue is the importance of cumulative versus contemporaneous exposure to malaria. The estimates from the literature suggest

relatively small effects on output of episodes of malarial fever, but persistent morbidity is likely to depress labor productivity contemporaneously. This effect operates as a function of time (i.e., pre/post campaign), so I implement a simple fix to purge the U.S. data of contemporary impacts of eradication. Specifically, I project the income data onto dummies for state of birth interacted with post campaign (in terms of time not cohort). These results are seen in the third and sixth row of Table 1, Panel A. (This projection is included in the estimator that is fed to the bootstrap routine, so the standard errors are adjusted for the additional step.) The estimates on childhood exposure are qualitatively similar to those reported elsewhere in Panel A. Note, on the other hand, 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. I therefore re-estimate the exposure coefficients excluding the 1960 census data, and report results in the third rows of Panel B and D. The estimates are qualitatively similar. The second issue is how to define the end of childhood; i.e., what is the appropriate age at which early-life malaria-exposure effects go to zero? An age around 20 years is a logical choice, given that educational investments and physical growth have mostly decelerated by then. As a check, I estimate smoothed versions of the $\hat{\beta}_k$, but this time organizing the data by age at first exposure to the campaign. For all four countries, cohorts that were first exposed in their twenties have similar outcomes to those exposed at older ages, which suggests that the cutoff around 20 years old is not a bad approximation. (See Appendix F for these results.) For two of the countries (Brazil and Colombia), the smoothed representation passes through the mean of those exposed at age 30+ at precisely 21 years. On the other hand, curves for Mexico and the U.S. reach the mean of those with later-life exposure by the late twenties.

Literacy increases with childhood exposure to the campaign, consistent with the central hypothesis of the present study. Cohort-specific results are seen in the first row of Figure 6. For all three Latin American countries, coefficients are generally negative for the earlier cohorts, and move closer to zero with greater childhood exposure to the anti-malaria campaign. The correspondence with childhood exposure is most evident in the graph for Brazil, and the exposure function is related to the estimated coefficients in a statistically significant way for all three countries. (Literacy and years of schooling are not available over the appropriate range of census years for the United States, so I cannot conduct this analysis for that country.)

Cohort-specific results for education yield mixed results across the Latin American countries. These estimates are seen in the second row of Figure 6. Estimates for the early cohorts hover around zero for Brazil and Colombia, but are generally negative for Mexico. Greater childhood exposure to the campaign is associated with more years of school in Brazil and Colombia, but fewer years of school in Mexico. Finally, the latter cohorts in Brazil and Mexico exhibit partial reversion to the levels seen for cohorts with no childhood exposure to the campaign, while the

Colombian series does not exhibit this property. These inconsistent results across countries and cohorts are not, however, inconsistent with economic theory of schooling, which suggests a first-order condition in which individuals compare returns to schooling with the opportunity cost of schooling (the childhood wage comprising an important part of the latter). Childhood malaria might depress both the return to education and the return to working. And, moreover, whether it depresses wages or returns to schooling more might depend on institutional and sectoral factors that differ across areas, or on constraints that are (differentially) relaxed over time. Therefore, the theory does not give a sharp prediction about the sign of the effect of the anti-malarial campaign on years of schooling.

5 Pre/Post Comparisons

I obtain similar results using a simple pre/post comparison. Once again, I compare outcomes across cohorts while separating along two lines: (i) by year of birth relative to the campaign and (ii) by the degree of pre-campaign malaria intensity in the area of birth. In contrast with the methodology of Section 4, however, 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 in this section.) Therefore, for each area 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. This permits the use of simple scatter plots, the analysis of which indicates that the results do not arise from the undue influence of a few observations. I also show that the results are generally robust to the inclusion of alternative controls, migration corrections and to the use of instruments to correct for measurement error. A further advantage of this long differencing is to reduce the bias in inference stemming from higher-frequency serial correlation.[16]

The basic equation to be estimated is

$$\Delta Y_{j,t} \;=\; \beta M_{j,t-1} \;+\; X_{j,t-1}\Gamma \;+\; \alpha \;+\; \varepsilon_{j,t} \tag{3}$$

in which Y is some socioeconomic outcome for state or municipio j. The time subscript t refers to a year of birth following the malaria-eradication campaign, while $t-1$ indicates being born (and having become an adult) prior to the advent of the campaign. The pre-program malaria incidence is $M_{j,t-1}$, the X variables are a series of controls, and α is a constant term. The parameter of interest is β. This parameter can be thought of as coming from a reduced-form equation.[17] (This

[16] While this methodology does not account for pre-existing trends, these were considered above. It was shown in most cases that the inclusion of linear trends had only modest effects on the estimates of childhood exposure to the eradication campaigns.

[17] The model can be motivated in the following manner. For an individual i, born in area j, with year-of-birth t,

equation is estimated with least squares, and, as above, the square root of the cell sizes are used to construct weights to account with the different precisions with which cohort means are estimated.)

5.1 United States

Areas in the U.S. 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 2. The first row of Panel A contain estimates for the basic specification of equation 3, which includes a dummy for being born in the South plus the natural logarithm of state unskilled wages in 1899 from Lebergott (1964). 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 2, Panel A contains these results. The second row controls for additional state-of-birth-level measures of health, including fertility, infant mortality, and the proportion of deaths from various childhood diseases 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 and total school expenditures. 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 (which corresponds to the set of controls used to estimate equation 1 in Section 4 above). Finally, the upper left graph in Figure 7 displays

---

we start with an individual-level model with individual infection data and linear effects of malaria:

$$Y_{ijt} = \alpha M_{ijt} + \delta_j + \delta_t + \tilde{\varepsilon}_{ijt}$$

where M_{ijt} is a measure of childhood malaria infection. The data do not contain both childhood malaria infection data and adult income, and moreover the research design is fundamentally at the period-of-birth $\times$ area-of-birth level, so I rewrite the equation above in aggregate form:

$$Y_{jt} = \tilde{\alpha} M_{jt} + \delta_j + \delta_t + \tilde{\varepsilon}'_{jt}.$$

I partition the cohorts into those born after the advent of the campaign and those who were already adults by the time the campaign started. I then difference the model along these lines, and take $M_{i,t-1}$ as an instrument for the decline in malaria following eradication. (This instrument follows the logic of Figure 2, which is notionally a first-stage relationship.) The resulting reduced form of this system is equation 3. Alternatively, one could have written the individual-level model with separate terms for individual and aggregate infection variables, the latter of which reflecting some spillover from peer infection to own income. But both of these effects would be subsumed into the $\tilde{\alpha}$ coefficient on the ecological infection rate, and it is this composite coefficient that I seek to measure in the present study.

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 of the over-identifying restrictions fails to reject the null of identical parameter estimates in the second stage.

Incorporating migration into the analysis does not materially alter the results. These results, estimated using the "full controls" specification from Panel A, are seen in Panel C of Table 2. The first two rows decompose the results by residence in one's state of birth. 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"). (To compute these estimates, I construct two cohort-level datasets, one for movers and another for nonmovers.) I also re-do the analysis using the state of birth of the individual's mother (results for father'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). In view of the year-of-birth-specific results above, I expect this intermediate group to exhibit a weaker response to the campaign than the fully exposed group. In the third row of Panel C, I restrict the sample to those with native-born mothers and use the redefined treatment group, but otherwise I replicate the specification from above in that the malaria variable is based on one's own state of birth. As expected, the coefficients are smaller in magnitude than those from Panel A. Next, I assign malaria based on mother's state of birth, and repeat the analysis. These estimates, found in the fourth row of Panel C, are similar to their baseline in the third row.

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 2. 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.

5.2 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 3 reports the estimates for these two countries for a variety of 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 as a control for mean reversion. The next row includes estimates controlling for infant mortality as well. The third row includes controls for the sectoral composition of the labor force (at a one-digit level). For the final row of Panel A, all the controls listed in Appendix C are included as regressors. 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 education are mixed, as per theoretical ambiguity. The components of malaria and income that are orthogonal to these controls are plotted in Figure 7. 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: in the range of 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 4. Specifications with alternative sets of controls are shown in Panel A, The basic specification (seen in the first row) includes malaria intensity and dummies for region of birth. In addition to the Poveda measure of malaria ecology from above, I also consider alternative measures of the disease: the Mellinger variable described above and malaria cases registered by the Colombian eradication campaign in 1955 (SEM, 1957). The second row adds controls for conflict,[18] 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 (and therefore related to conflict arising during land reform). The third row reports estimates controlling for various types of economic activity: coffee-growing, mining, ranching, and manufacturing; as well as population density and indices describing the general level of development and quality of infrastructure. For the fourth row, controls for the presence of various helminth and (non-malarial) vector-borne diseases are included. The final row of the panel reports estimates controlling for all

[18]See Garcia Montalvo and Reynal Querol (2006) for evidence on the importance of conflict in worsening the malaria situation in a panel of countries.

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 7.

When correcting for measurement error, estimates of the malaria coefficient rise substantially, which suggests attenuation bias in the OLS estimates.[19] 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 produced with all of the above-mentioned instrumental variables. Hausman tests do not reject the overidentification restrictions at conventional confidence levels.

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 0.09–0.14 log points of income.

6 Interpretation

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

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

[19] An additional measure from the SEM is the fraction of cases of malaria due to *falciparum* (the more lethal strain) rather than *vivax* (the high morbidity strain). I have experimented with using this measure in an interactive model to estimate effects by strain of malaria. The estimates on *falciparum* are typically lower than on *vivax*, consistent with the maintained interpretation that the effects identified in the present study work through childhood morbidity. However, the estimates that attempt to separately identify the two strains are quite imprecisely determined, suggesting a large measurement-error problem in the discrimination between these two strains by municipio. Understanding differences in the response to different strains of the disease remains an important topic for future research.

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 U.S. 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.[20]

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 4 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 over-estimate of the full income effect, since the index effectively double counts schooling. I report it

[20]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.

nevertheless for completeness.

6.2 Mechanisms

Schooling. Formal education had an important, but by no means dominant, role in the Latin-American results. (Lack of data prevent doing this computation for the U.S.) 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.[21]

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.

Increases in the return to schooling may have contributed to the rise in income following the eradication campaigns, although the estimates of this channel are too imprecise to make definitive statements. To investigate this, I estimate Mincerian returns to schooling (RTS) for each cohort (again defined by year of birth $\times$ state of birth) for Mexico and Brazil. This procedure generates a panel data set of RTS, which I then analyze in a like manner to Section 4.[22] For Mexico, estimates for the reduced-form effect of exposure on RTS hover around zero, but are imprecisely determined. When comparing the most to least malarious areas, the estimated differential effect has a 95% confidence interval of ± 3 percentage points. At the upper extreme of this confidence band, it is possible to account, via increasing RTS $\times$ average years of schooling, for all of the effect

[21] 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.

[22] This test has certain limitations. This procedure, by comparing individuals with different terminal levels of attainment, estimates the average marginal effect of schooling in the sample (and how it changes following the anti-malaria campaign eradication). Absent distortions, the standard economic theory of schooling suggests that this effect should look like an interest rate, and that we should instead look for an effect along the quantity margin. (I.e., even if less exposure to malaria increased the return to schooling *ceteris paribus*, the quantity of schooling would adjust upwards until the return to additional schooling was back down to the rate of interest on borrowing.) However, the sign of the quantity response of schooling to childhood malaria is not informative about the structural RTS because malaria also affects childhood wages. Moreover, it might be that the intervention had large effects on *inframarginal* returns to school investments. For example, the benefit might mostly accrue to elementary-school children without altering the marginal return at higher levels of education. A final problem is that the returns are estimated from the cross section, and no attempts are made to purge these estimates of endogeneity bias.

of childhood malaria exposure on income. On the other hand, because of the imprecision of the estimates, changes in RTS could equally well account for none or even -100% of the estimated above. For Brazil, estimates of the effect of childhood malaria on RTS are also imprecise, such that the estimates can similarly account for both the observed effect on income as well as zero.

Labor-market experience. I consider the quantity of time worked and return to labor-market experience as possible mechanisms. Information on time worked are not available for a long enough range of cohorts is these data, so I employ an alternative strategy. The analysis for the U.S. was based on occupational indices of income calibrated to median total income in 1950. Using this same methodology, I constructed on occupational index of *hours*, which is constructed from the each occupation's median number of weeks worked last year $\times$ usual hours worked. Using the methodology of Section 4, I treat this as an outcome variable and find that childhood exposure to the anti-malaria campaign shifts workers into occupations with fewer, not more, average hours. This suggests that childhood malaria's impact on income was via the hourly wage rather than via labor supply. (This conclusion is confirmed by constructing occupational indices of hourly wages rather than total income.)

On the other hand, increasing returns to labor-market experience was a modest component of the income results above. I repeat the analysis of Section 4, but do so separately for each age from 25–55. The estimated effect of childhood exposure is significantly differently from zero for men in their late twenties. This effect rises by about 30% by the time the worker reaches 55. Taken together, these facts suggest that childhood malaria depressed the return to labor-market experience, but that this mechanism accounts for about a third of the total effect on income.

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 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 U.S. 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%.

The timing of childhood exposure. The definition of childhood exposure that I used above assumed an effect of malaria infection that was uniform across youth. As seen above and in Appendix F, 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 those same cohorts were adolescents at the time of the campaign, this suggests that the result is not driven exclusively by adolescent exposure.

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 4. 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 U.S. 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. (Bleakley, 2007b, attempts to estimate across-cohort spillovers effects for

malaria and hookworm eradications in the Southern U.S. and finds them to be positive, although considerably smaller than the main effect of childhood exposure to eradication.)

6.3 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.[23] 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 U.S.

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 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 those 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.

[23] The approximate income figures cited in this subsection are from Barro and Sala-i-Martin (1999). The infection rates are estimated from Molineaux (1988).

7 Conclusions

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

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 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 data.

In both absolute terms and relative to the comparison group of non-malarious areas, cohorts born after eradication had higher income and literacy as adults than the preceding generation. Mixed results are found for years of education, consistent with the economic theory of schooling (which compares returns with opportunity costs).

Another result of the present study is finding similar estimates of the effect of childhood malaria exposure on adult income across the four countries considered. This fact is remarkable particularly given the composition of the sample: one developed country and three others 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. The results indicate potentially large benefits of interventions against malaria (especially the *vivax* strain) in developing 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.

References

Abbott, Samuel W. (1900). "The Past and Present Condition of Public Hygiene and State Medicine in the United States." in H. B. Adams, ed., *Monographs of American Social Economics*, Vol. XIX, Commonwealth of Massachusetts. Prepared for the United States Commission to the Paris Exposition of 1900.

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

——— , ——— , and James A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review.* December **91** (5), 1369–401.

Agudelo, Saúl Franco (1990). *El Paludismo en America Latina* Medicina Social, Guadalajara, Jalisco: Editorial Universidad de Guadalajara.

Alsan, Marcella, David E. Bloom, David Canning, and Dean Jamison (2006). "The Consequences of Population Health for Economic Performance." PGDA Working Papers 1306, Program on the Global Demography of Aging October. available at http://ideas.repec.org/p/gdm/wpaper/1306.html.

Banco de la República (1964). *Atlas de Economía Colombiana*, Bogotá: Departamento de Imprenta del Banco de la República.

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

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

Barro, Robert J. and Xavier Sala-i-Martin (1999). *Economic Growth*, Cambridge: The MIT Press.

Bleakley, C. Hoyt (2007a). "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics.* February **122** (1), 73–117.

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

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

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

Brinkley, Garland L. (1994). "The Economic Impact of Disease in the American South, 1860-1940." Doctoral dissertation, University of California, Davis.

Central Intelligence Agency (2006). *CIA World Fact Book*, Washington: Central Intelligence Agency, June 29. Web page: http://www.cia.gov/cia/publications/factbook/.

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

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

Cueto, Marcos (2004). *El valor de la salud: historia de la Organización Panamericana de la Salud*, Washington, D.C.: Organización Panamericana de la Salud.

——— , ed. (1994). *Missionaries of science: the Rockefeller Foundation and Latin America*, Bloomington, Ind.: Indiana University Press.

Cutler, David, Winnie Fung, Michael Kremer, and Monica Singhal (2007). "Mosquitoes: The Long-Term Effects of Malaria Eradication in India." May. Unpublished manuscript.

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

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

——— (1968-70). *Anuario de Salubridad*, Bogotá, Colombia: Departamento Administrativo Nacional de Estadística.

——— (2000). *División Político-Administrativa de Colombia*, Bogotá, Colombia: DANE.

Dirección General de Estadística (1952a). *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.

——— (1952b). *Compendio Estadistico*, Mexico, D.F.: Secretaria de la Economía Nacional.

——— (1960). *Anuario Estadístico 1958-1959*, Mexico, D.F.: Secretaria de Industria y Comercio.

Dirección Nacional de Estadística (1947). *Primer censo industrial de Colombia, 1945*, Vol. 1-10, Bogotá: Imprenta Nacional.

Duncan, Otis Dudley (1961). "A Socioeconomic Index for All Occupations." in A. J. Reiss, ed., *Occupations and Social Status*, The Free Press New York pp. 109–138.

Gallup, John Luke and Jeffrey D. Sachs (2001). "The Economic Burden of Malaria." *The American Journal of Tropical Medicine & Hygiene.* January/February **64** ((1, 2)S), 85–96.

———, Andrew Mellinger, and Jeffrey D. Sachs (1999a). "Geography and Economic Development." Working Paper 1, Center for International Development at Harvard University, Cambridge, Mass. March.

———, ———, and ——— (1999b). "GIS Data." Machine-readable dataset (Arc/info interchange file), Center for International Development (CID) at Harvard University. http://www2.cid.harvard.edu/ciddata/Geog/gisdata.html.

Garcia Montalvo, José and Marta Reynal Querol (2006). "Fighting against malaria: prevent wars while waiting for the miraculous vaccines." *Review of Economics and Statistics.* forthcoming.

Gray, R. H. (1974). "The Decline of Mortality in Ceylon and the Demographic Effects of Malaria Control." *Population Studies.* July **28** (2), 205–229.

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

Hong, Sok Chul (2007). "Health and Economic Burden of Malaria in Nineteenth-Century America." Doctoral dissertation, University of Chicago. (proposed June 2006).

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

——— (1956). *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.

Inter-university Consortium for Political and Social Research (1984). *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, http://www.icpsr.org. [Accessed Sept. 20, 2002.].

International Health Board (1919). *Annual Report*, New York: Rockefeller Foundation.

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

Kofoid, Charles A. and John P. Tucker (1921). "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.* January **1** (1), 79–117.

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

Lopez Alonso, Moramay and Raúl Porras Condey (2003). "The Ups and Downs of Mexican Economic Growth: The Biological Standard of Living and Inequality, 1870-1950." *Journal of Economics and Human Biology.* **1** (2).

Lucas, Adrienne M. (2005). "Economic Effects of Malaria Eradication: Evidence from the Malarial Periphery." October. Unpublished manuscript.

Malaney, Pia (2003). "Micro-Economic Approaches to Evaluating the Burden of Malaria." Working Paper 99, Center for International Development at Harvard University, Cambridge, Mass. April.

Maxcy, Kenneth F. (1923). "The Distribution of Malaria in the United States as Indicated by Mortality Reports." *Public Health Reports.* May **38** (21), 1125–38.

McCarthy, Desmond, Holger Wolf, and Yi Wu (2000). "The Growth Costs of Malaria." Working Paper 7541, National Bureau of Economic Research, Cambridge, Mass. February.

McCullough, David (1977). *The Path Between the Seas: The Creation of the Panama Canal, 1870-1914*, New York: Simon and Schuster.

Meisel, Adolfo and Margarita Vega (2004). "A Tropical Success Story: A Century of Improvements in the Biological Standard of Living, Colombia 1910-2002." Borradores de Economía 289, Banco de la República, Bogotá, Colombia.

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

Mendoza, Juan and Andrés Rosas (2004). "Geography, Health and Development: Evidence from Colombia." Unpublished manuscript, Brown University March.

Miller, N. Grant (2005). "Contraception as Development? New Evidence from Family Planning in Colombia." Working Paper 11704, National Bureau of Economic Research, Cambridge, Mass.

Molineaux, Louis (1988). "The epidemiology of human malaria as an explanation of its distribution, including some implications for its control." in W. H. Wernsdorfer and I. A. McGregor, eds., *Malaria: Principles and Practice of Malariology*, Vol. 2, London, U.K.: Churchill Livinstone, chapter 35, pp. 913–998.

Nájera, José A., Bernhard H. Liese, and Jeffrey Hammer (1992). "Malaria: New Patterns and Perspectives." Technical Paper 183, The World Bank, Washington, D.C.

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

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

Pesqueira, Manuel E. (1957). "Programa de Erradicación del Paludismo en México." *Boletin de la Oficina Sanitaria Panamericana.* June **XLII** (6), 537–41.

Poveda, German, Nicholas E. Graham, Paul R. Epstein, William Rojas, Martha L. Quiñones, Iván Darío Vélez, and Willem J. M. Martens (2000). "Climate and ENSO Variability Associated with Vector-Borne Diseases in Colombia." in H. F. Diaz and V. Markgraf, eds., *El Niño and the Southern Oscillation : Multiscale Variability and Global and Regional Impacts*, Cambridge University Press Cambridge, U.K. and New York, N.Y. pp. 183–204.

Quevedo, Emilio, Catalina Borda, Juan Carlos Eslava, Claudia Mónica García, María de Pilar Guzmán, Paola Mejía, and Carlos Ernesto Moguera (2004). *Café y Gusanos, Mosquitos y Petróleo: El tránsito desde higiene hacia la medicina tropical y la salud pública en Colombia, 1873-1953*, Bogotá, Colombia: Universidad Nacional de Colombia/Instituto de Salud Pública/Centro de Historia de la Medicina.

Ribero, Rocío and Jairo Nuñez (2000). "Adult Morbidity, Height, and Earnings in Colombia." in W. D. Savedoff and T. P. Schultz, eds., *Wealth from Health: Linking Social Investments to Earnings in Latin America*, Washington, D.C.: Inter-American Development Bank.

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

Sachs, Jeffrey D. (2001). "Tropical Underdevelopment." Working Paper 8119, National Bureau of Economic Research, Cambridge, Mass. February.

——— (2003). "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income." Working Paper 9490, National Bureau of Economic Research, Cambridge, Mass. February.

Sanchez Torres, Fabio and Jairo Nuñez (2000). "Geograpy and Economic Development: A Municipal Approach for Colombia." *Revista de Planeación y Desarollo.* July-December **XXXI** (3-4).

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

——— (1979). "Evaluación del Programa de Malaria: Marco de referencia para el grupo de asesorimiento nacional e internacional." Technical report, Ministerio de Salud, Bogotá, Colombia. April.

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

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

U.S. Bureau of the Census (1894). "Compendium of the Eleventh Census: 1890." in "Vital and Social Statistics," Vol. II, Washington: U.S. Government Printing Office.

——— (1933). *Vital Statistics of the United States*, Washington: U.S. Government Printing Office.

US Office [Bureau] of Education (1894-1913). *Annual Report of the Commissioner of Education*, Washington D.C.: Government Printing Office.

Utzinger, Jürg, Yesim Tozan, Fadi Doumani, and Burton H. Singer (2001). "The economic payoffs of integrated malaria control in the Zambian copperbelt between 1930 and 1950." *Tropical Medicine and International Health.* August **7** (8), 657–677.

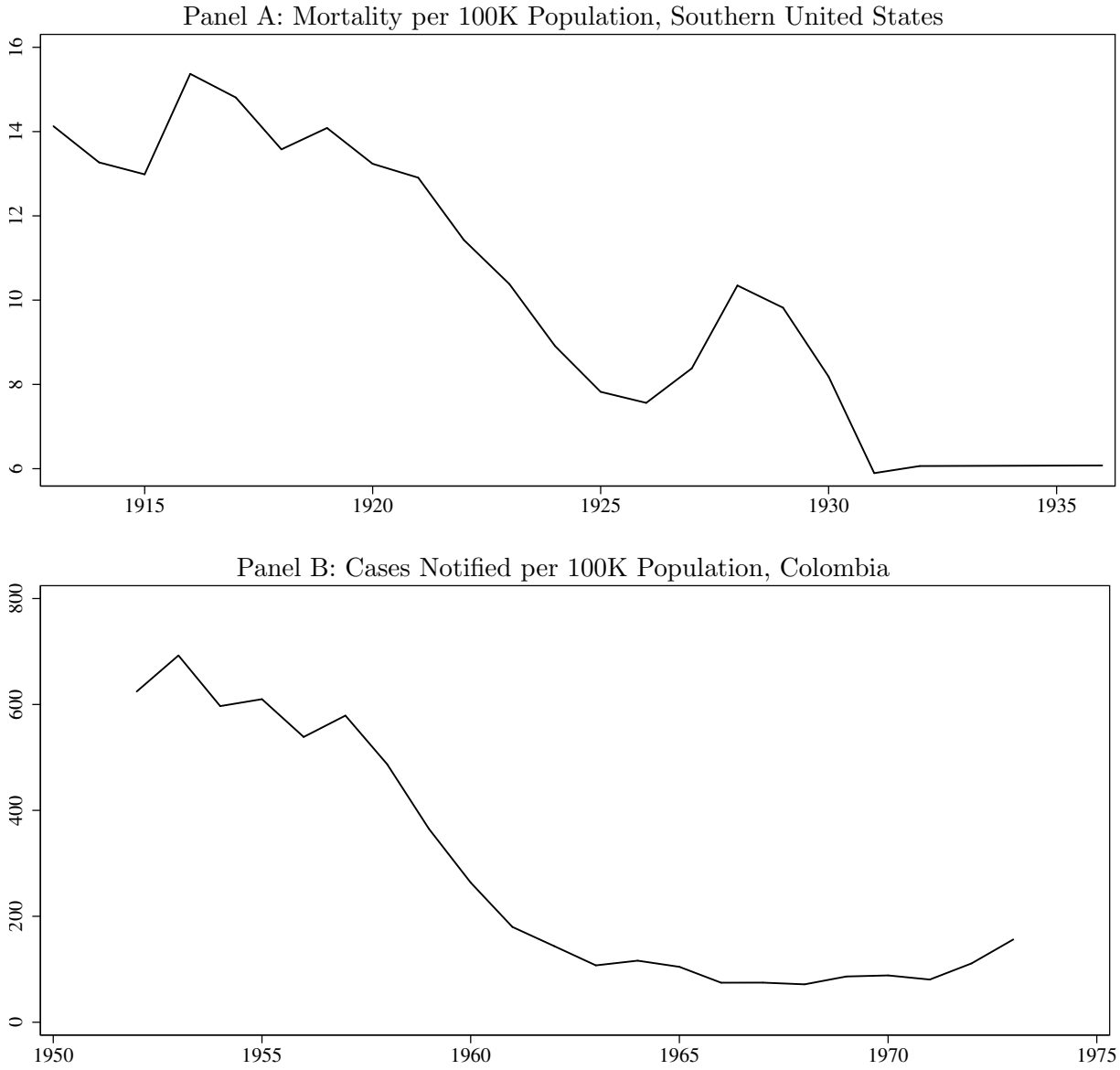
Weil, David N. (2007). "Accounting for the Effect of Health on Economic Growth." *The Quarterly Journal of Economics.* August **122** (3), 1265–1306.

Wernsdorfer, Gunther and Walter H. Wernsdorfer (1988). "Social and economic aspects of malaria and its controls." in W. H. Wernsdorfer and I. A. McGregor, eds., *Malaria: Principles and Practice of Malariology*, Vol. 2, London, U.K.: Churchill Livinstone, chapter 47, pp. 1421–1472.

Wernsdorfer, Walter H. and Ian A. McGregor (1988). *Malaria: Principles and Practice of Malariology*, London, U.K.: Churchill Livinstone.

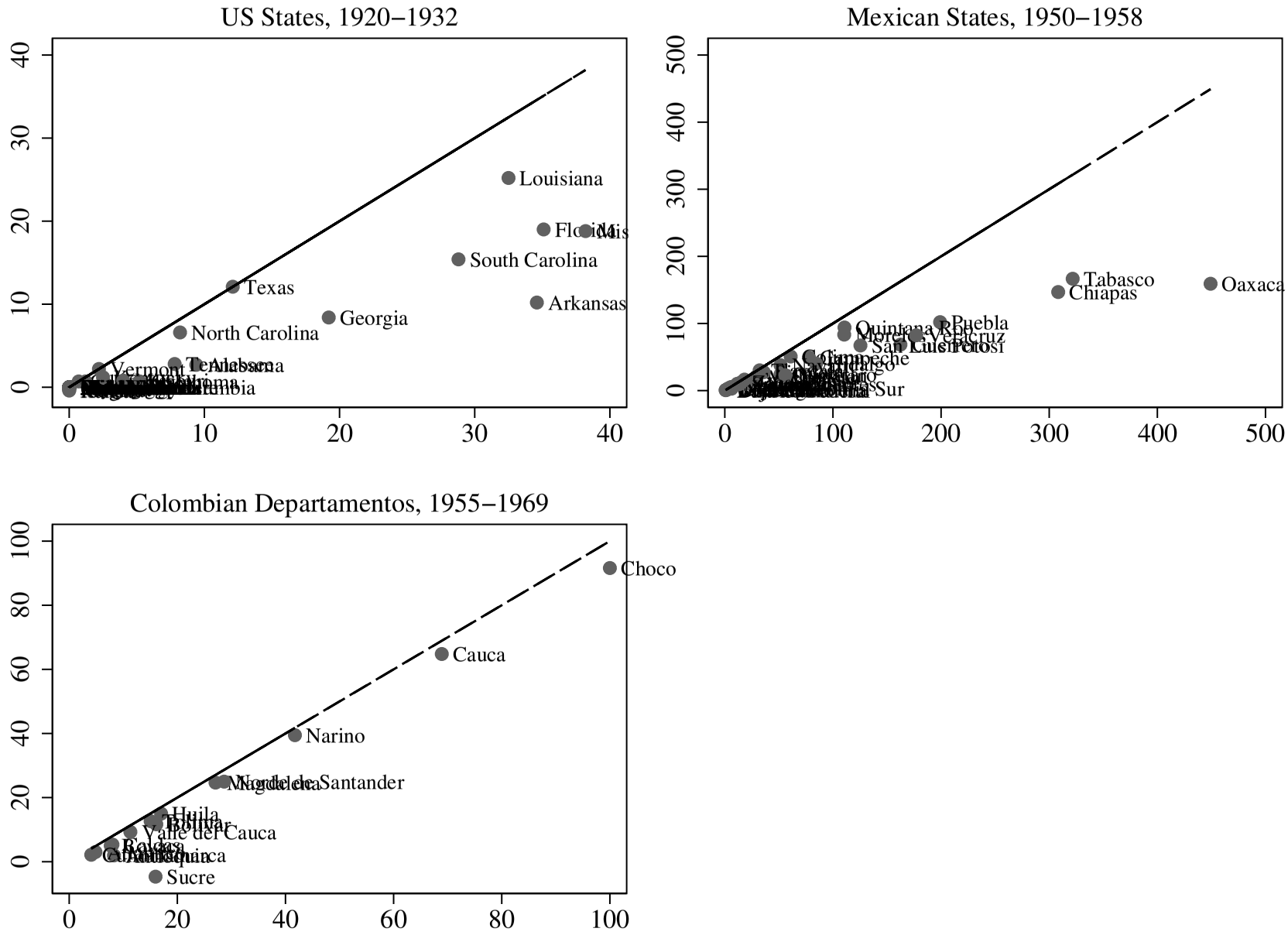
Williams, Ralph Chester (1951). *The United States Public Health Service, 1798-1950*, Bethesda, Md.: Commissioned Officers Association of the United States Public Health Service.

Figure 1: Malaria Incidence Before and After the Eradication Campaigns



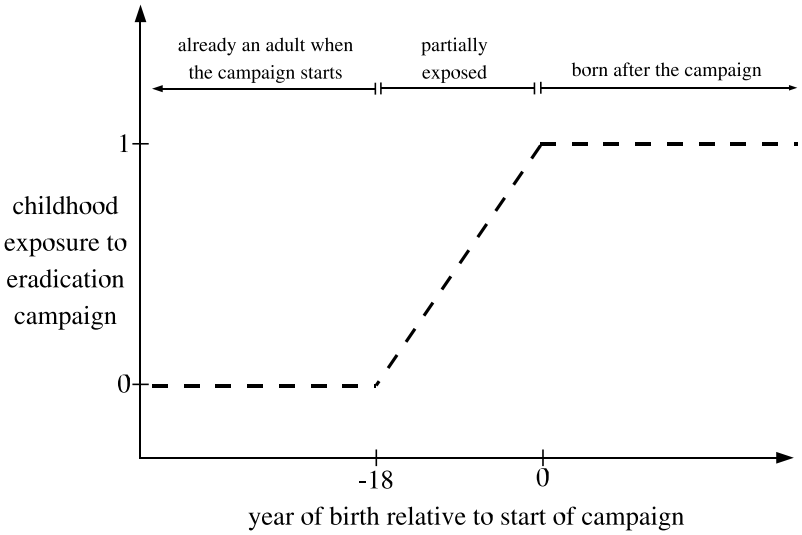
Notes: Panel A plots the estimated malaria mortality per capita for the Southern region and bordering states. Because the death registration system was being phased in over the period, a regression model with state fixed effects is used to control for sample changes, and the time series is constructed from the year dummies in the regression, normalized to match the end-of-period data when all states were represented. (Census Bureau *Vital Statistics*, various years, and author's calculations.) Panel B reports data on notified cases of malaria for Colombia (SEM, 1979).

Figure 2: Highly Infected Areas Saw Greater Declines in Malaria



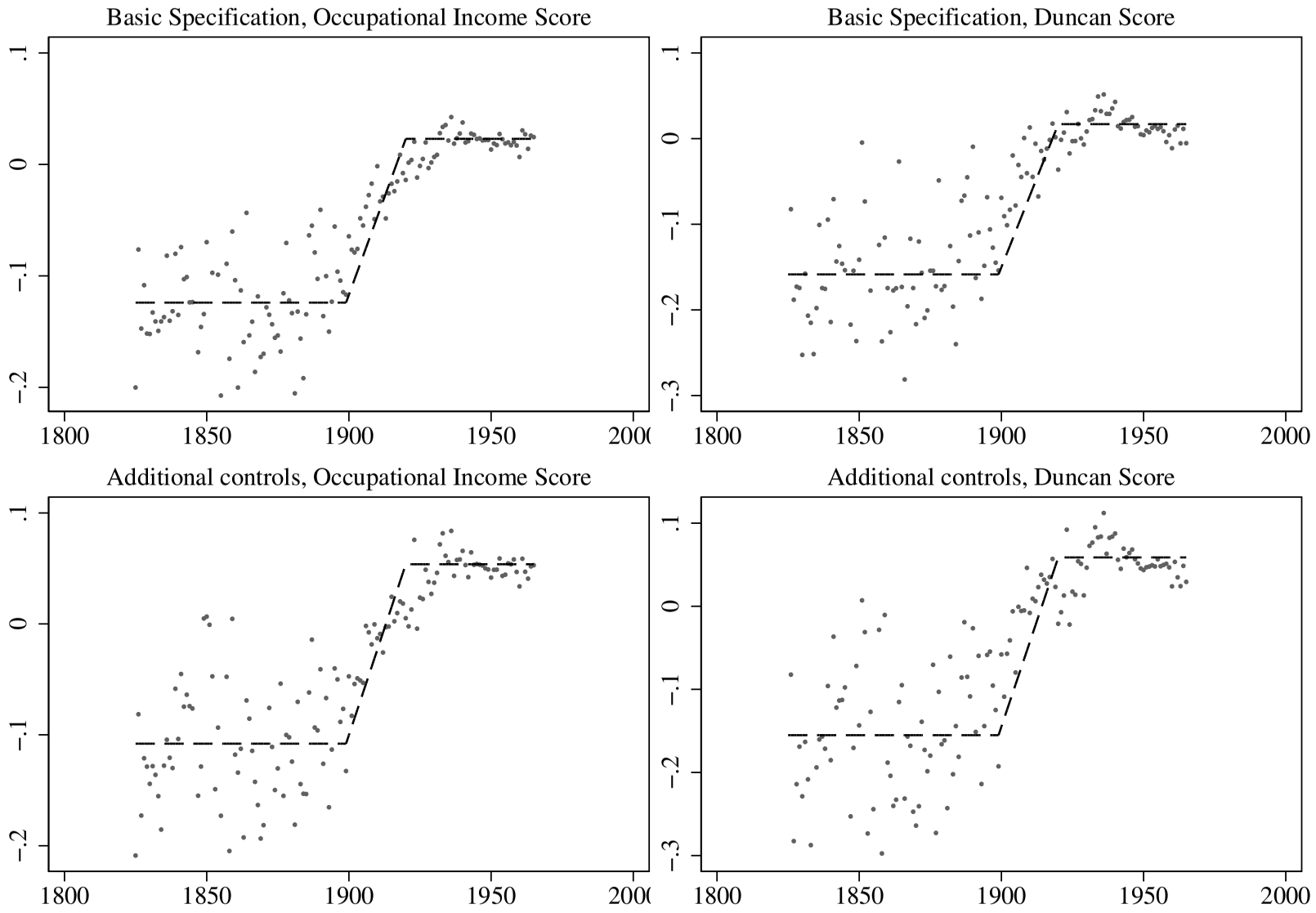
Notes: The y axis displays the estimated decrease in malaria mortality post-intervention. The x axis is the pre-campaign malaria mortality rate. The 45-degree line represents complete eradication. Both variables are expressed per 100,000 population. United States data are reported in Maxcy (1923) and *Vital Statistics* (Census, 1933). Mexican data are drawn from Pesqueira (1957) and from the Mexican *Anuario Estadístico* (Dirección General de Estadística, 1960). SEM (1957) and the Colombian *Anuario de Salubridad* (DANE, 1968-70) are the sources for the Colombian data.

Figure 3: Childhood Exposure to Eradication Campaign



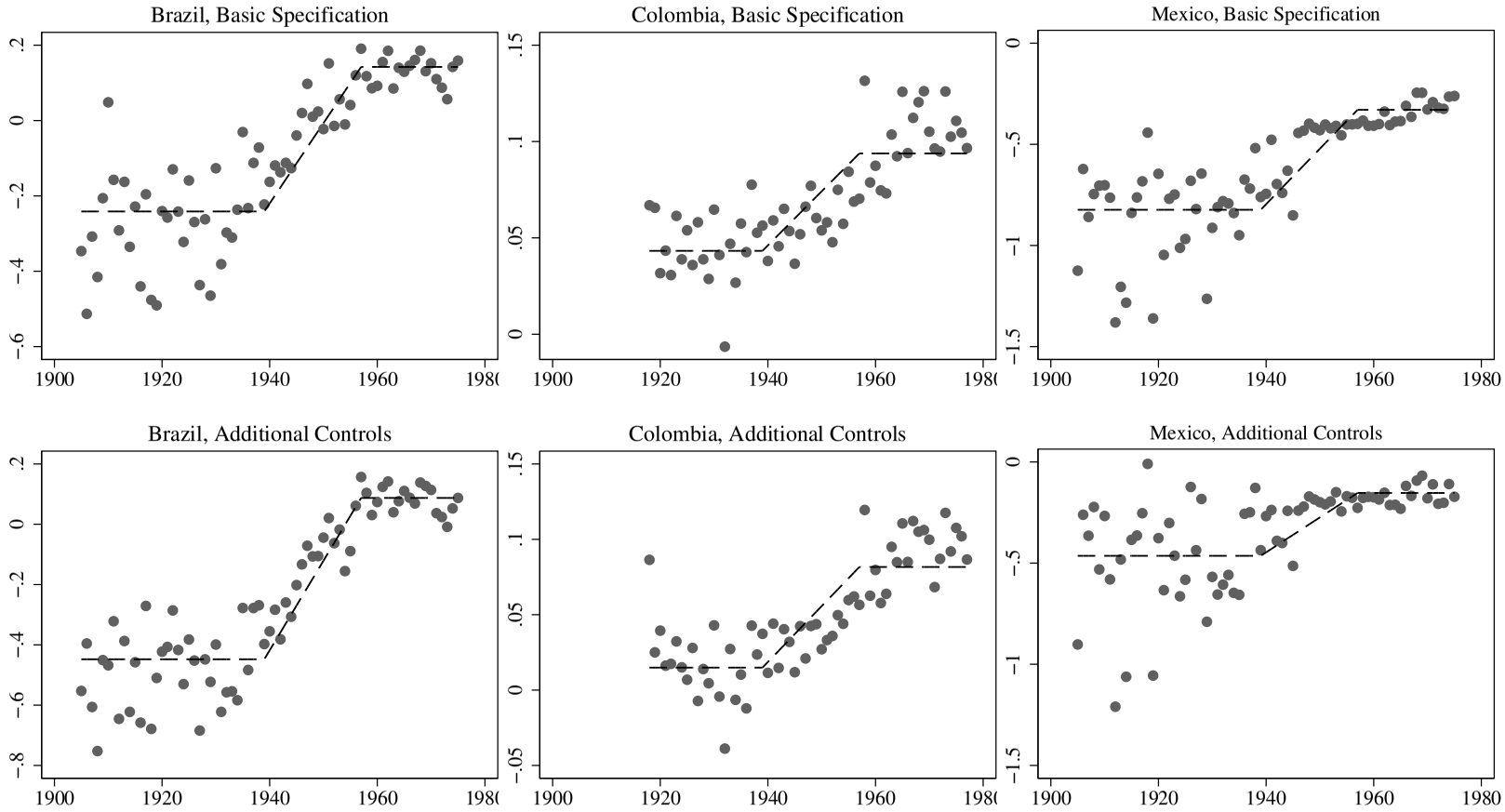
Notes: This graph displays on the fraction of childhood that is exposed to a hypothetical (and instantaneous) campaign as a function of year of birth minus the start year of the campaign.

Figure 4: Cohort-Specific Relationship: Income across States in the U.S.



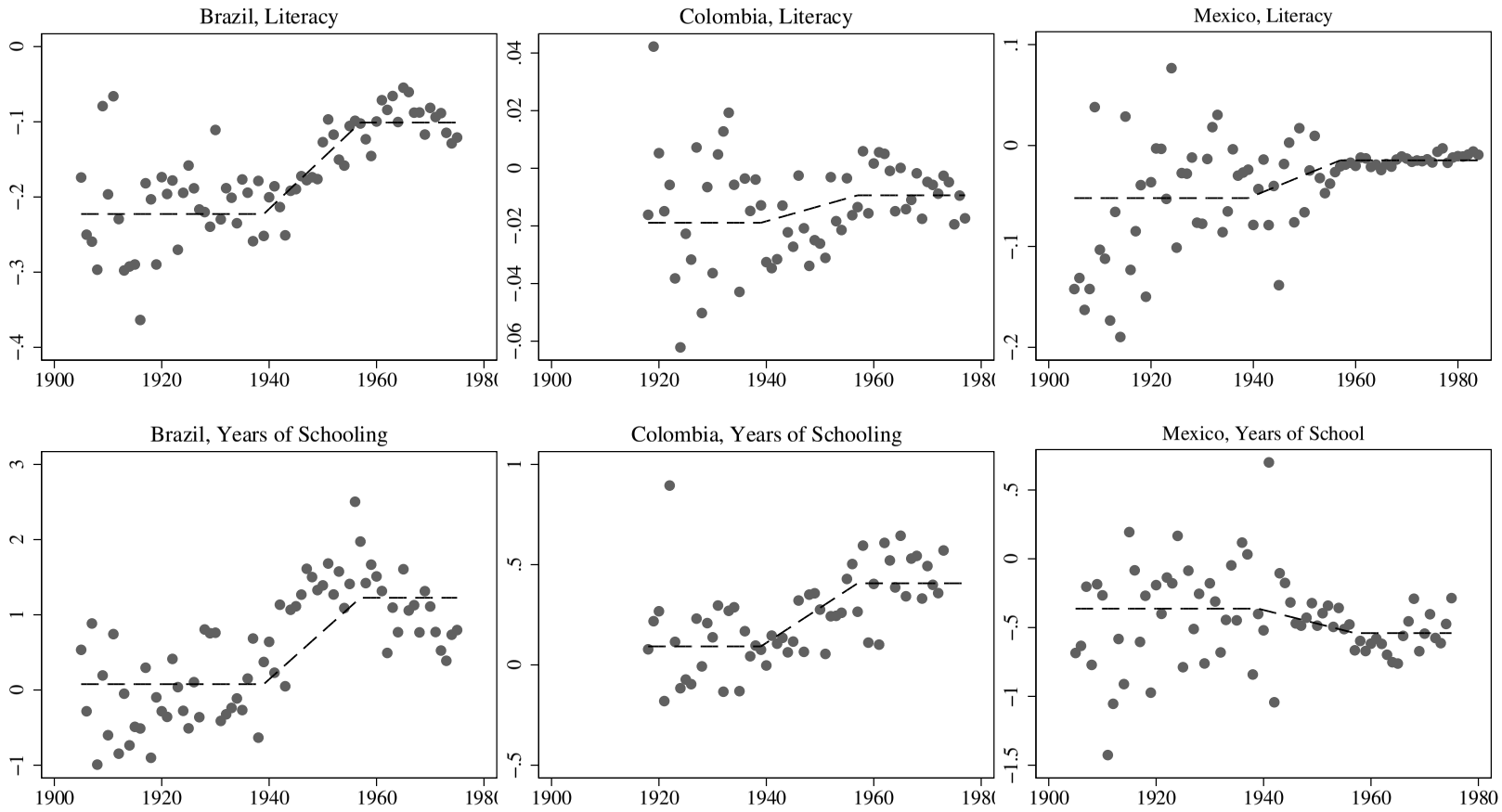
Notes: These graphics summarize regressions of income proxies on pre-eradication malaria-mortality rates (measured by the Census in 1890). The y axis for each graphic plots the estimated cohort-specific coefficients on the state-level malaria measure (malaria mortality / total mortality in 1890). Income is measured in logs and the malaria proxy is scaled by the gap between the 95th and 5th percentile states in 1890. The x axis is the cohort's year of birth. Each cohort's point estimate is marked with a dot. The dashed lines measure the approximate number of years of potential childhood exposure to the malaria-eradication activities in the South. For each year-of-birth cohort, OLS regressions coefficients are estimated using the cross section of states of birth and census years. In the "basic" specification, the income proxy is regressed onto malaria, Lebergott's (1964) measure of 1899 wage levels, and a dummy for being born in the South. The "additional controls" specification includes the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure. The cohort-specific coefficients on the additional controls are plotted in Appendix D.

Figure 5: Cohort-Specific Relationship: Income in Brazil, Colombia, and Mexico



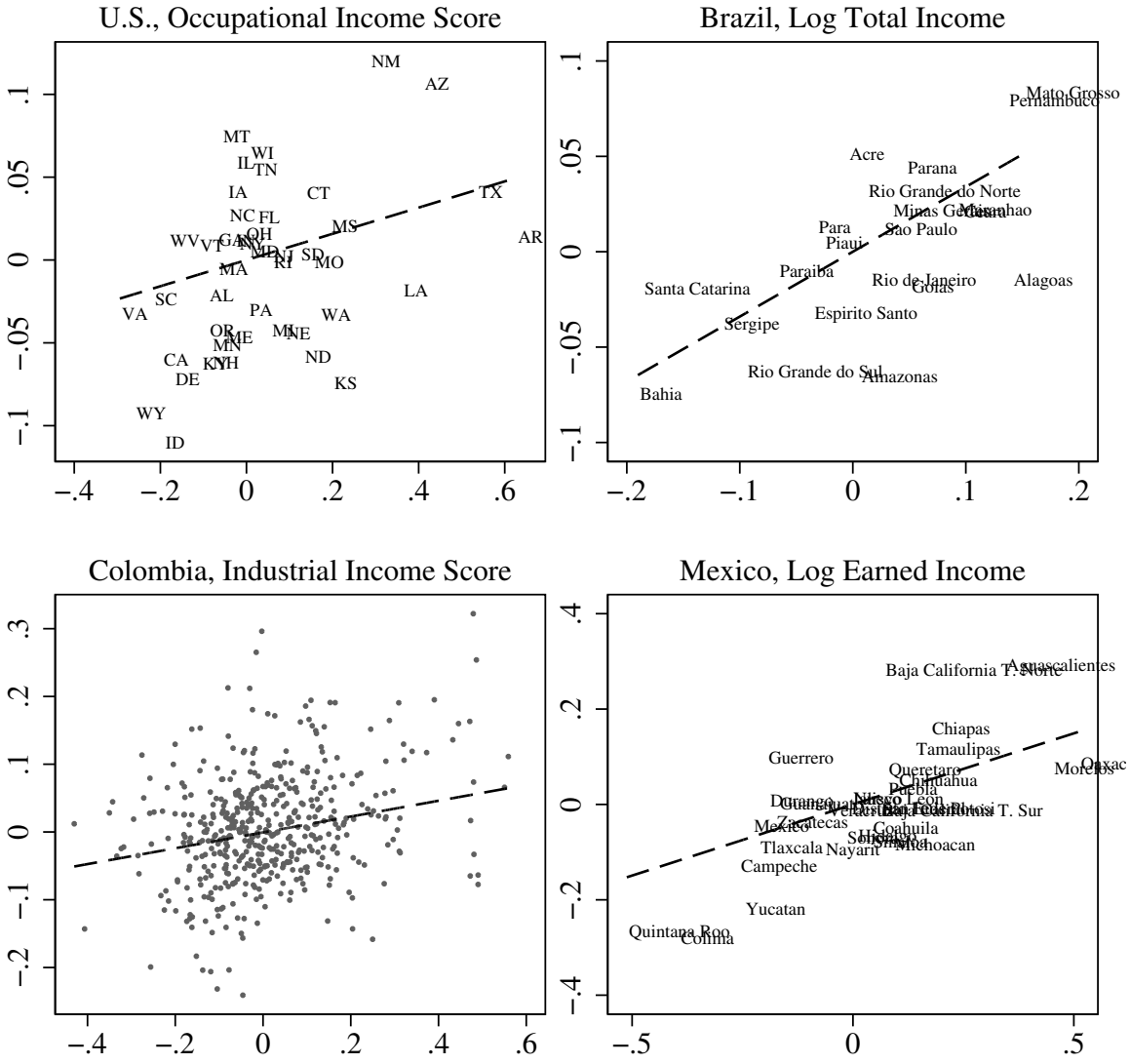
Notes: These graphics summarize regressions of the indicated income variables on measures of malaria prevalence. The y axis for each graphic plots the estimated cohort-specific coefficients on the state-level malaria measure. Income is measured in logs and the malaria proxy is scaled by the gap between the 95th and 5th percentile areas. The x axis is the cohort's year of birth. Each cohort's point estimate is marked with a dot. The dashed lines measure the approximate number of years of potential childhood exposure to the commencement of the malaria-eradication campaign in Latin America (approximately 1957). For each year-of-birth cohort, OLS regressions coefficients are estimated using the cross section of states (Brazil and Mexico) or municipios (Colombia) of birth and census years. For the "basic" specification, income is regressed onto malaria, region dummies, and an proxy for pre-campaign state income (population density (all 3 countries), log of electrical capacity (Brazil and Mexico), and a heuristic indicator of "nivel de vida" (Colombia)). The "additional controls" specification also includes the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure. The cohort-specific coefficients on the additional controls are plotted in Appendix D.

Figure 6: Cohort-Specific Relationship: Human Capital in Brazil, Colombia, and Mexico



Notes: These graphics summarize regressions of the indicated human-capital variables on measures of malaria prevalence. The y axis for each graphic plots the estimated cohort-specific coefficients on the state-level malaria measure. 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. Each cohort's point estimate is marked with a dot. The dashed lines measure the approximate number of years of potential childhood exposure to the commencement of the malaria-eradication campaign in Latin America (approximately 1957). For each year-of-birth cohort, OLS regressions coefficients are estimated using the cross section of states (Brazil and Mexico) or municipios (Colombia) of birth and census years. For the "basic" specification, income is regressed onto malaria, region dummies, and an proxy for pre-campaign state income (population density (all 3 countries), log of electrical capacity (Brazil and Mexico), and a heuristic indicator of "nivel de vida" (Colombia)). The "full controls" specification also includes the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure.

Figure 7: 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 "additional 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.

Table 1: Malaria Eradication: Childhood Exposure versus Alternative Processes

Degree of Polynomial-Trend Control:		0	1	0	1	2	0	2
Degree of Autoregressive Process:		0	0	1	1	0	2	2
Specification:	Outcome:							
				Panel A: United States				
Basic	Occupational Income Score	0.124 ***	0.109 ***	0.104 ***	0.094 ***	0.109 ***	0.093 ***	0.082 **
		(0.004)	(0.009)	(0.015)	(0.019)	(0.008)	(0.030)	(0.036)
Additional controls	Occupational Income Score	0.061 ***	0.150 ***	0.128 ***	0.101 ***	0.131 ***	0.120 ***	0.080 *
		(0.006)	(0.012)	(0.011)	(0.026)	(0.011)	(0.027)	(0.047)
Birthstate x census>1920	Occupational Income Score	0.071 ***	0.150 ***	0.133 ***	0.099 ***	0.131 ***	0.026 **	0.100 **
		(0.005)	(0.010)	(0.009)	(0.022)	(0.008)	(0.015)	(0.040)
Basic	Duncan's Index	0.162 ***	0.126 ***	0.138 ***	0.113 ***	0.139 ***	0.121 **	0.114 **
		(0.007)	(0.015)	(0.022)	(0.031)	(0.014)	(0.050)	(0.060)
Additional controls	Duncan's Index	0.088 ***	0.184 ***	0.058 ***	0.154 ***	0.172 ***	0.041	0.113
		(0.010)	(0.018)	(0.018)	(0.044)	(0.017)	(0.030)	(0.079)
Birthstate x census>1920	Duncan's Index	0.099 ***	0.181 ***	0.067 ***	0.159 ***	0.168 ***	0.053 **	0.139 **
		(0.007)	(0.014)	(0.012)	(0.031)	(0.013)	(0.023)	(0.063)
				Panel B: Brazil				
Basic	Log Total Income	0.184 ***	0.220 ***	0.164 ***	0.197 **	0.277 ***	0.122	0.205
		(0.020)	(0.048)	(0.047)	(0.092)	(0.048)	(0.087)	(0.620)
Additional controls	Log Total Income	0.348 ***	0.437 ***	0.308 ***	0.405 ***	0.486 ***	0.268 *	0.417
		(0.019)	(0.050)	(0.082)	(0.128)	(0.048)	(0.160)	(1.896)
Additional controls	Log Earned Income	0.297 ***	0.459 ***	0.345 ***	0.520 **	0.432 ***	0.308	0.368
		(0.042)	(0.110)	(0.117)	(0.260)	(0.138)	(0.224)	(2.069)
Additional controls, drop 1960 census	Log Total Income	0.226 ***	0.133 **	0.190 ***	0.088	0.201 ***	0.132	0.161
		(0.023)	(0.061)	(0.058)	(0.120)	(0.055)	(0.125)	(0.714)
				Panel C: Colombia				
Basic	Industrial Income Score	0.036 **	0.041 **	0.036 ***	0.034 **	0.031 **	0.032 **	0.036 **
		(0.015)	(0.016)	(0.014)	(0.014)	(0.014)	(0.013)	(0.018)
Additional controls	Industrial Income Score	0.063 ***	0.047 **	0.053 ***	0.025 **	0.032 **	0.037 **	0.021 **
		(0.019)	(0.023)	(0.018)	(0.018)	(0.020)	(0.016)	(0.020)
				Panel D: Mexico				
Basic	Log Earned Income	0.253 ***	0.162 *	0.269 ***	0.135	0.077	0.191 *	-0.001
		(0.057)	(0.068)	(0.094)	(0.169)	(0.052)	(0.108)	(0.535)
Additional controls	Log Earned Income	0.231 ***	0.155 *	0.250 **	0.105	0.068	0.211	0.059
		(0.071)	(0.084)	(0.118)	(0.162)	(0.074)	(0.187)	(0.805)
Additional controls, drop 1960 census	Log Earned Income	0.385 ***	0.176 *	0.365 ***	0.142	0.176 *	0.360	0.076
		(0.043)	(0.099)	(0.132)	(0.203)	(0.105)	(0.311)	(1.511)

Notes: This table reports estimates of the childhood-exposure variable in equation 2 using OLS. The outcome variables used to construct the time series of $\hat{\beta}_k$ are as indicated in each row. Estimates are computed by 1000 iterations of a block bootstrap (at the level of state or municipio of birth) of the two-step procedure implied by equations 1 and 2. Standard errors in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Observations for equation 2 are weighted by the inverse of the coefficient's standard error. Reporting of additional terms suppressed.

Table 2: 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 ***	0.134 **	0.236 ***	0.219 ***
	(0.039)	(0.065)	(0.032)	(0.053)
Health	0.100 ***	0.144 **	0.225 ***	0.280 ***
	(0.038)	(0.067)	(0.031)	(0.048)
Education	0.136 ***	0.131 **	0.219 ***	0.206 ***
	(0.041)	(0.062)	(0.027)	(0.055)
Other	0.094 **	0.115 *	0.204 ***	0.178 ***
	(0.044)	(0.063)	(0.029)	(0.068)
Full Controls	0.110 **	0.172 *	0.215 ***	0.265 ***
	(0.049)	(0.094)	(0.049)	(0.096)
Instrumental Variables:	*Panel B: Estimates using Two-Stage Least Squares*			
The Other Malaria Proxy	0.142 ***	0.175 **	0.207 ***	0.244 **
	(0.054)	(0.088)	(0.060)	(0.106)
Average Temperature and Altitude	0.154 *	0.209 **	0.138 **	0.174 **
	(0.083)	(0.104)	(0.059)	(0.075)
All of the Above Instruments	0.149 ***	0.192 **	0.164 ***	0.185 ***
	(0.054)	(0.095)	(0.052)	(0.071)
Sample/Specification:	*Panel C: Migration*			
Movers	0.111 **	0.136 *	0.292 ***	0.367 ***
	(0.048)	(0.074)	(0.042)	(0.070)
Nonmovers	0.107 **	0.165 *	0.193 ***	0.185 *
	(0.045)	(0.100)	(0.040)	(0.096)
Sample with native-born mothers (exposed = born 1905-15)	0.035	0.116 *	0.108 ***	0.167 **
	(0.026)	(0.064)	(0.039)	(0.079)
Use malaria of mother's birthstate (exposed = born 1905-15)	0.061	0.107	0.153 **	0.203 *
	(0.041)	(0.079)	(0.062)	(0.106)
Sample:	*Panel D: By Region*			
South + Bordering States (N=21)	0.100 ***	0.173 ***	0.208 **	0.203
	(0.038)	(0.038)	(0.086)	(0.236)
Rest of the country (N=25)	0.183	0.134	0.162	0.281
	(0.276)	(0.408)	(0.212)	(0.338)

Notes: This table reports estimates of equation 3 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 population between the ages of 25 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. Appendices A and B describe, respectively, the outcome variables and malaria measures. The additional controls are described in the text and Appendix C. 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 mother's state of birth is only available up through the 1940 Census, so the exposed (post-campaign) group includes mostly cohorts with partial exposure to 1920. The bordering states in Panel D include DE, IL, IN, MD, MO, NM, OH, OK, and PA.

Table 3: Cross-Cohort Differences and Malaria: Brazil and Mexico

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

Notes: This table reports estimates of malaria in equation 3 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 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. 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 4: 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 ***	0.168 *	0.065 ***	0.071 ***	0.064	0.048 ***	0.018 *	0.077	0.026 ***
	(0.013)	(0.088)	(0.011)	(0.016)	(0.108)	(0.014)	(0.010)	(0.067)	(0.008)
Conflict	0.032 ***	0.175 *	0.063 ***	0.068 ***	0.068	0.046 ***	0.016	0.078	0.024 ***
	(0.012)	(0.090)	(0.011)	(0.016)	(0.110)	(0.014)	(0.010)	(0.067)	(0.008)
Economic Activity	0.008	0.194 **	0.057 ***	0.043 ***	0.156	0.039 ***	0.008	0.071	0.022 ***
	(0.010)	(0.089)	(0.012)	(0.013)	(0.110)	(0.014)	(0.008)	(0.067)	(0.006)
Other Diseases	0.024 *	0.180 **	0.065 ***	0.058 ***	0.057	0.042 ***	0.012	0.067	0.020 **
	(0.013)	(0.089)	(0.012)	(0.016)	(0.114)	(0.015)	(0.010)	(0.067)	(0.008)
Full Controls	0.006	0.165 *	0.064 ***	0.046 ***	0.076	0.034 **	0.008	0.041	0.016 ***
	(0.011)	(0.095)	(0.013)	(0.015)	(0.117)	(0.015)	(0.008)	(0.067)	(0.006)
Instrumental Variables:	*Panel B: Alternative Instrument Sets*								
Temperature, Altitude, and their Interaction	0.037 **	0.372 ***	0.092 ***	0.067 *	0.766 ***	0.170 ***	0.055	0.679 **	0.159 ***
	(0.018)	(0.136)	(0.017)	(0.036)	(0.268)	(0.037)	(0.035)	(0.284)	(0.047)
The Other Two Malaria Proxies	0.126 ***	0.113	0.084 ***	0.082 ***	0.390 *	0.149 ***	0.200 ***	0.145	0.131 ***
	(0.032)	(0.190)	(0.026)	(0.029)	(0.203)	(0.028)	(0.069)	(0.312)	(0.051)
Holdridge Climate Zone	0.045 **	0.303 *	0.102 ***	0.082 **	0.593 **	0.124 ***	0.006	-0.483	0.088
	(0.021)	(0.159)	(0.020)	(0.037)	(0.248)	(0.035)	(0.059)	(0.508)	(0.061)
All of the Above Instruments	0.049 ***	0.323 ***	0.092 ***	0.074 ***	0.516 ***	0.120 ***	0.059 **	0.131	0.075 ***
	(0.017)	(0.122)	(0.016)	(0.026)	(0.184)	(0.025)	(0.026)	(0.179)	(0.026)

Notes: This table reports estimates of malaria in equation 3 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 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. Excluded instruments for 2SLS in Panel B are municipio-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 and instruments are described in the text and Appendix C.

Table 5: Approximate Effects on Adult Income per Probability of Childhood Malaria Infection

	United States		Brazil		Colombia	Mexico
Dependent Variables:	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 2, 3, and 4, for the indicated outcomes. Reported are averages of the OLS point estimates, multiplied by the difference between 95th and 5th percentile malaria intensity. For the United States, these numbers are also normalized by the average value of the relevant income proxy for white males born in the South between 1875 and 1895. 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.

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 1 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. The weights are based on analysis by Duncan (1961) who regressed a measure of perceived prestige of several occupations on its average income and education. 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. (See Appendix Figure A–1.) 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. See Appendix Figure A–2 for sample statistics by year of birth.

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. See Appendix Figure A–3 for sample statistics by year of birth.

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

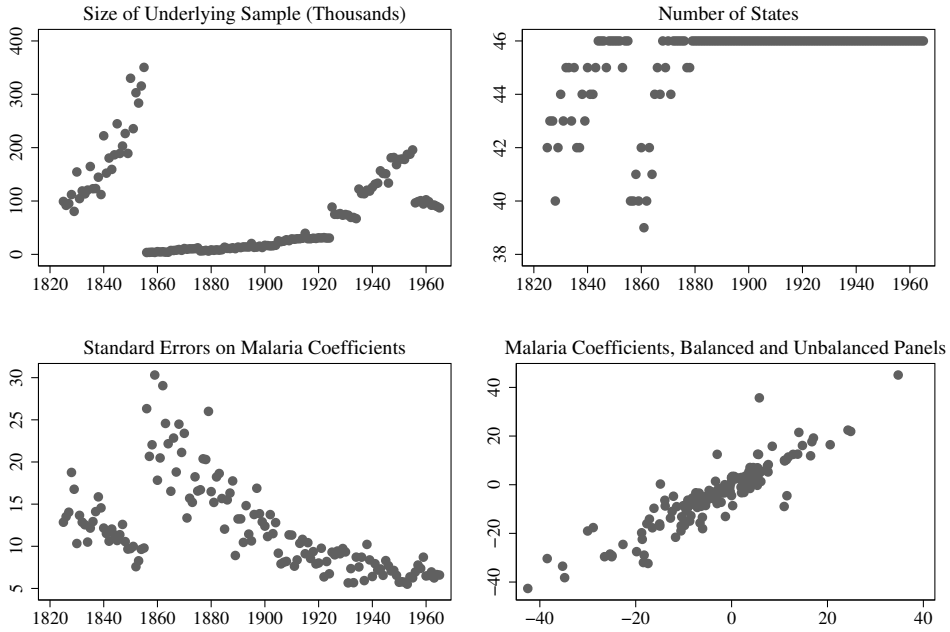
The underlying sample used for Mexico consists of native males in the age range [15,60] in the 1960–2000 IPUMS microdata. (These data were last accessed April 7, 2006.) This results in a data set with year-of-birth cohorts from 1905 to 1984. See Appendix Figure A–4 for sample statistics by year of birth.

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:

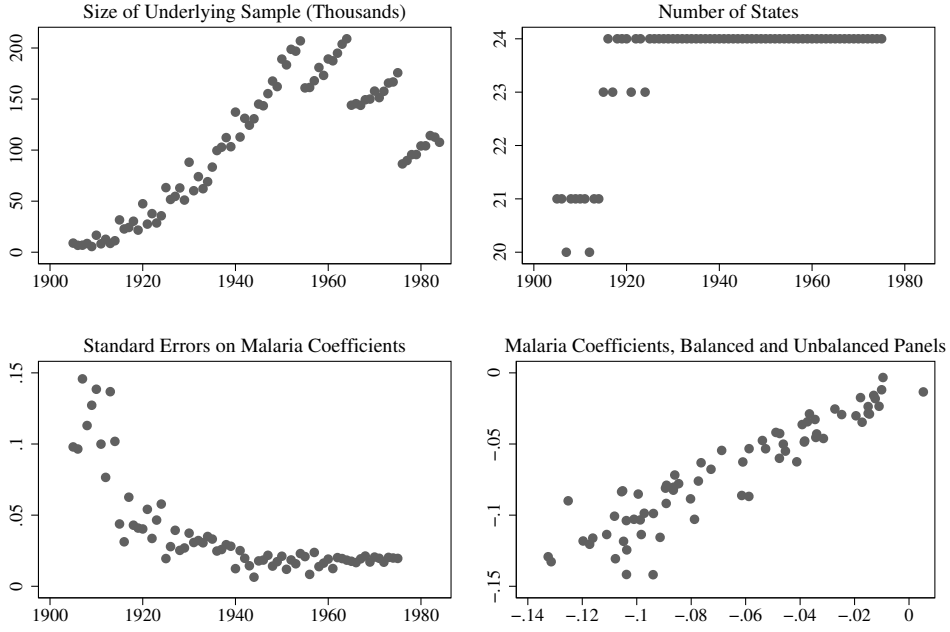
- **Literacy.** A binary variable individual measuring whether an individual can read and write.
- **Years of Schooling.** Numbers of years of education corresponding to highest grade completed. Non-numeric responses (e.g., "some secondary") are mapped onto the midpoints of the appropriate intervals.
- **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.)

Appendix Figure A – 1: Sample Statistics for the U.S. Sample



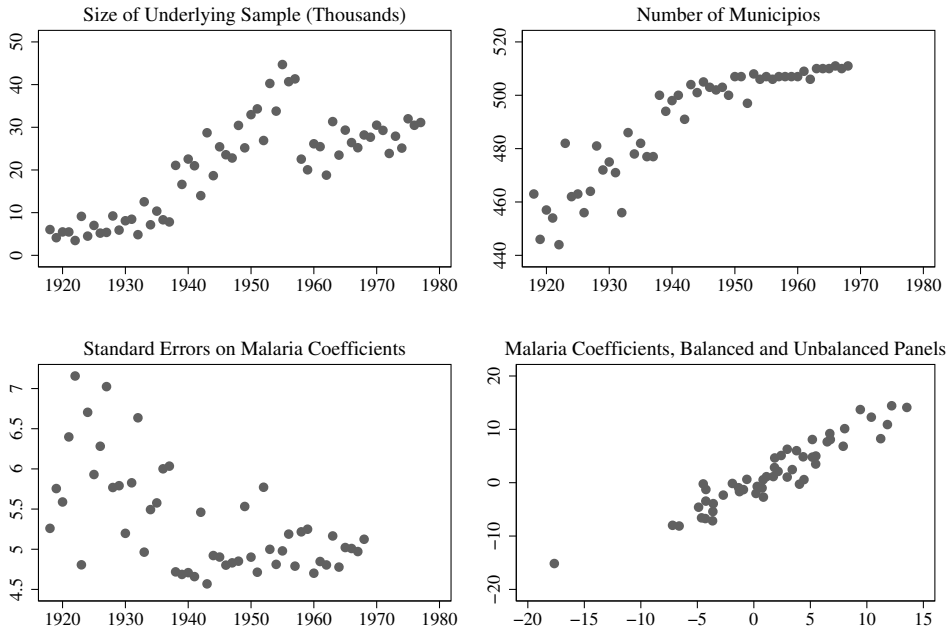
These graphs report additional summary statistics by year of birth for the $\hat{\beta}_t$ reported in Figure 4 in the subplot labeled "Additional controls, Occupational Income Score."

Appendix Figure A – 2: Sample Statistics for the Brazilian Sample



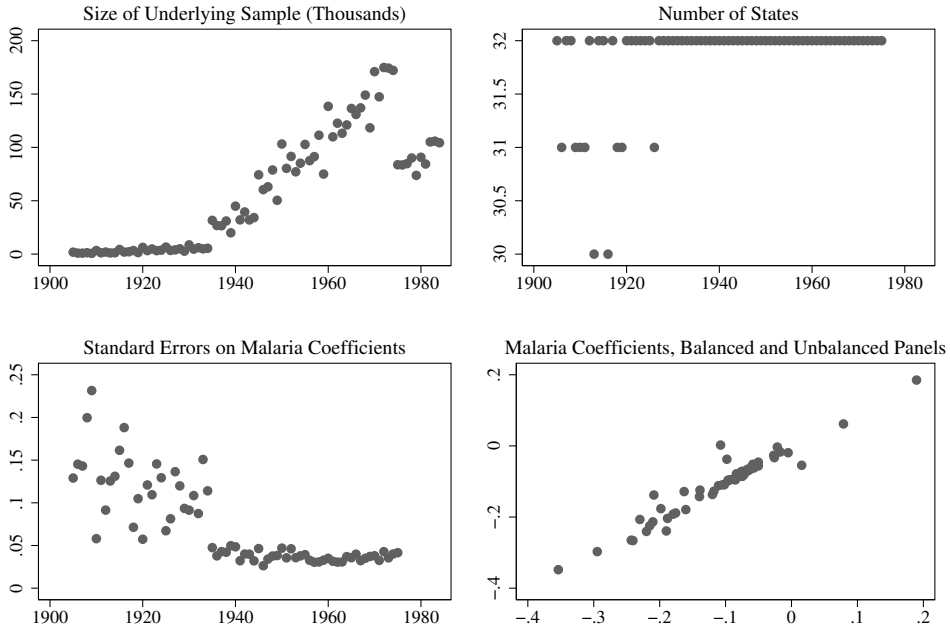
These graphs report additional summary statistics by year of birth for the $\hat{\beta}_t$ reported in Figure 5 in the subplot labeled "Brazil, Additional Controls".

Appendix Figure A – 3: Sample Statistics for the Colombian Sample



These graphs report additional summary statistics by year of birth for the $\hat{\beta}_t$ reported in Figure 5 in the subplot labeled "Colombia, Additional Controls".

Appendix Figure A – 4: Sample Statistics for the Mexican Sample



These graphs report additional summary statistics by year of birth for the $\hat{\beta}_t$ reported in Figure 5 in the subplot labeled "Mexico, Additional Controls".

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.)

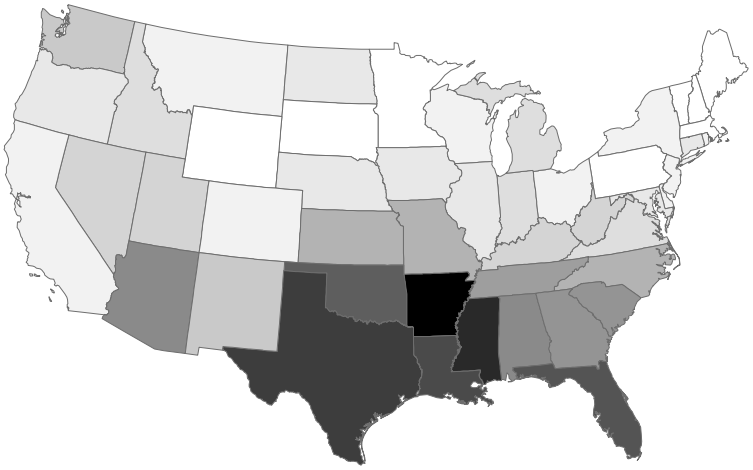
- **Colombia.** Two measures of ecology are used, as well as one measure of morbidity. The Poveda measure is an index of malaria ecology based on climatic factors, described by Poveda, Graham, Epstein, Rojas, Quiñones, Darío Vélez, and Martens (2000).

  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.)

Figure B – 1: Malaria Intensity by State in the United States

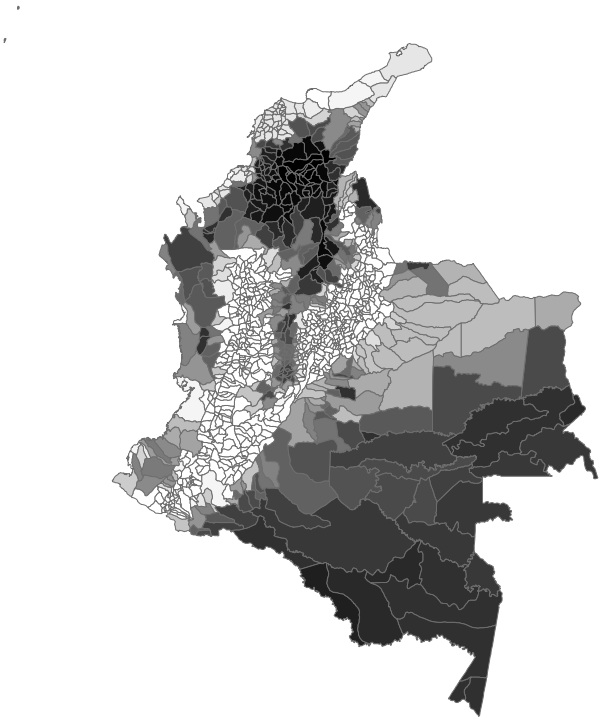


Notes: Displays a map of the ratio of malaria mortality to total mortality by state *circa* 1890. Source: U.S. Bureau of the Census (1894). Darker colors indicate more malaria.

Figure B – 2: Malaria Intensity by State in Brazil

Notes: Displays a map of an index of malaria ecology as constructed by Mellinger *et al.* (2004). Darker colors indicate climatic and geographic conditions more conducive to the transmission of malaria.

Figure B – 3: Malaria Intensity by Municipio in Colombia



Notes: Displays a map of an index of malaria ecology as constructed by Mellinger *et al.* (2004). Darker colors indicate climatic and geographic conditions more conducive to the transmission of malaria.

Figure B – 4: Malaria Intensity by State in Mexico

Displays a map of malaria mortality per capita, *circa* 1950. Source: Pesqueira (1957). Darker colors indicate more malaria.

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:

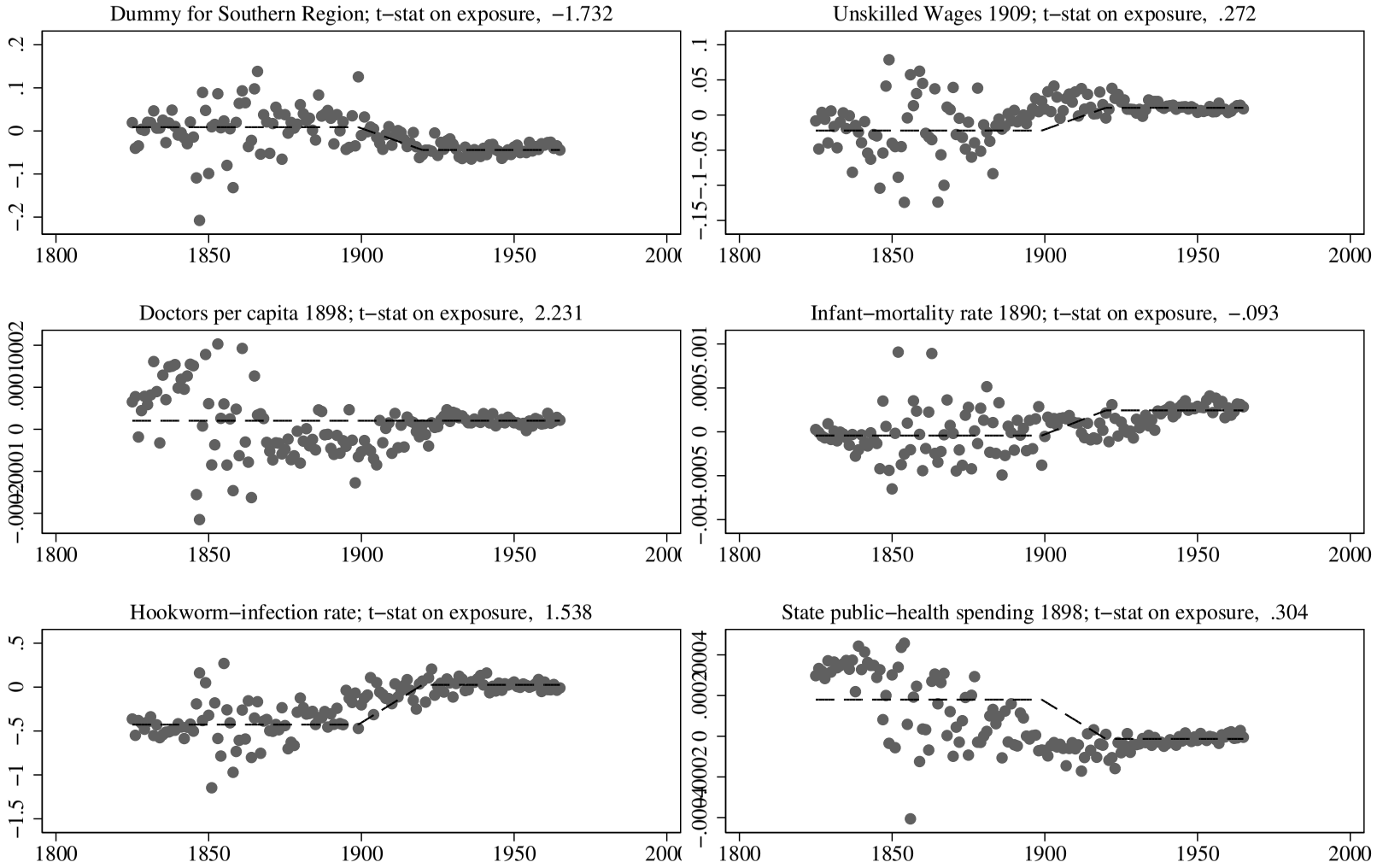
- **Region dummies.** "Norte", "Centro" and "Sur", as per the divisions in the 1960 *Anuario Estadístico*. (Dirección General de Estadística, 1960.)
- **Population Density.** Population per square kilometer in 1950. (Dirección General de Estadística, 1952a and 1952b.)
- **Infant mortality.** Rate per 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.)

Appendix Table C – 1: Correlations between Malaria Proxy and Control Variables

Correlation coefficient:	Simple	Partial		Simple	Partial		Simple	Partial		Simple	Partial
Country:	United States			Mexico			Brazil			Colombia	
Control variables:			Control variables:						Control variables:		
South	0.73 *** {0.00}	0.11 {0.56}	North	-0.36 ** {0.05}	0.27 {0.26}				Access to principal markets	-0.14 *** {0.00}	-0.07 {0.15}
Average unskilled wage	-0.62 *** {0.00}	0.19 {0.29}	South	0.69 *** {0.00}	0.57 *** {0.01}		-0.83 *** {0.00}	-0.19 {0.57}	"Nivel de Vida"	-0.37 *** {0.00}	-0.15 *** {0.00}
Infant mortality	-0.44 *** {0.00}	0.06 {0.75}	Population density	-0.04 {0.83}	0.00 {0.99}		-0.41 * {0.06}	-0.31 {0.36}	Manufacturing empl.	-0.05 {0.24}	0.07 {0.13}
Population urban	-0.63 *** {0.00}	-0.08 {0.67}	Infant mortality	-0.18 {0.34}	0.16 {0.49}		0.38 * {0.08}	0.14 {0.68}	Coffee-growing Region	-0.38 *** {0.00}	-0.18 *** {0.00}
Adult literacy rate	-0.70 *** {0.00}	0.10 {0.59}	Fraction economically active	0.39 ** {0.03}	0.43 * {0.06}		-0.46 ** {0.03}	-0.22 {0.51}	Coal-mining Region	-0.22 *** {0.00}	0.07 {0.13}
Doctors per capita	-0.30 ** {0.05}	-0.28 {0.12}	Emp. share, agriculture	0.44 ** {0.01}	0.22 {0.34}		0.36 {0.10}	-0.26 {0.45}	Expansion of Ranching	0.27 *** {0.00}	0.20 *** {0.00}
State public-health spending	0.00 {0.99}	0.01 {0.97}	Emp. share, extractive industries	-0.27 {0.14}	0.17 {0.47}		0.35 {0.11}	-0.26 {0.45}	High concentration "Minifundista"	-0.56 *** {0.00}	-0.34 *** {0.00}
Hookworm infection rate	0.68 *** {0.00}	-0.07 {0.70}	Emp. share, manufacturing	-0.26 {0.15}	0.29 {0.22}		-0.64 *** {0.00}	-0.18 {0.59}	"La Violencia", early	-0.07 {0.10}	0.03 {0.48}
Fraction black	0.77 *** {0.00}	0.43 ** {0.01}	Emp. share, transportation	-0.46 *** {0.01}	0.19 {0.43}		-0.57 *** {0.01}	-0.13 {0.71}	"La Violencia", late	-0.11 ** {0.02}	0.02 {0.66}
Unemployment 1930	-0.59 *** {0.00}	-0.01 {0.95}	Emp. share, services	-0.33 * {0.07}	0.17 {0.48}		-0.51 ** {0.01}	0.20 {0.56}	Non-hookworm helminth present	-0.28 *** {0.00}	-0.18 *** {0.00}
Change in school term length	0.42 *** {0.00}	0.17 {0.36}	Emp. share, other	-0.52 *** {0.00}	0.05 {0.83}		-0.57 *** {0.01}	-0.47 {0.15}	Hookworm present	0.31 *** {0.00}	0.41 *** {0.00}
Change in pupil/teacher ratio	0.32 ** {0.03}	0.20 {0.26}	Electrical Capacity	-0.22 {0.23}	0.26 {0.27}		-0.75 *** {0.00}	-0.75 *** {0.01}	Leishmaniasis present	-0.09 ** {0.05}	0.13 *** {0.00}
									Yellow fever present	0.13 *** {0.00}	-0.08 * {0.08}

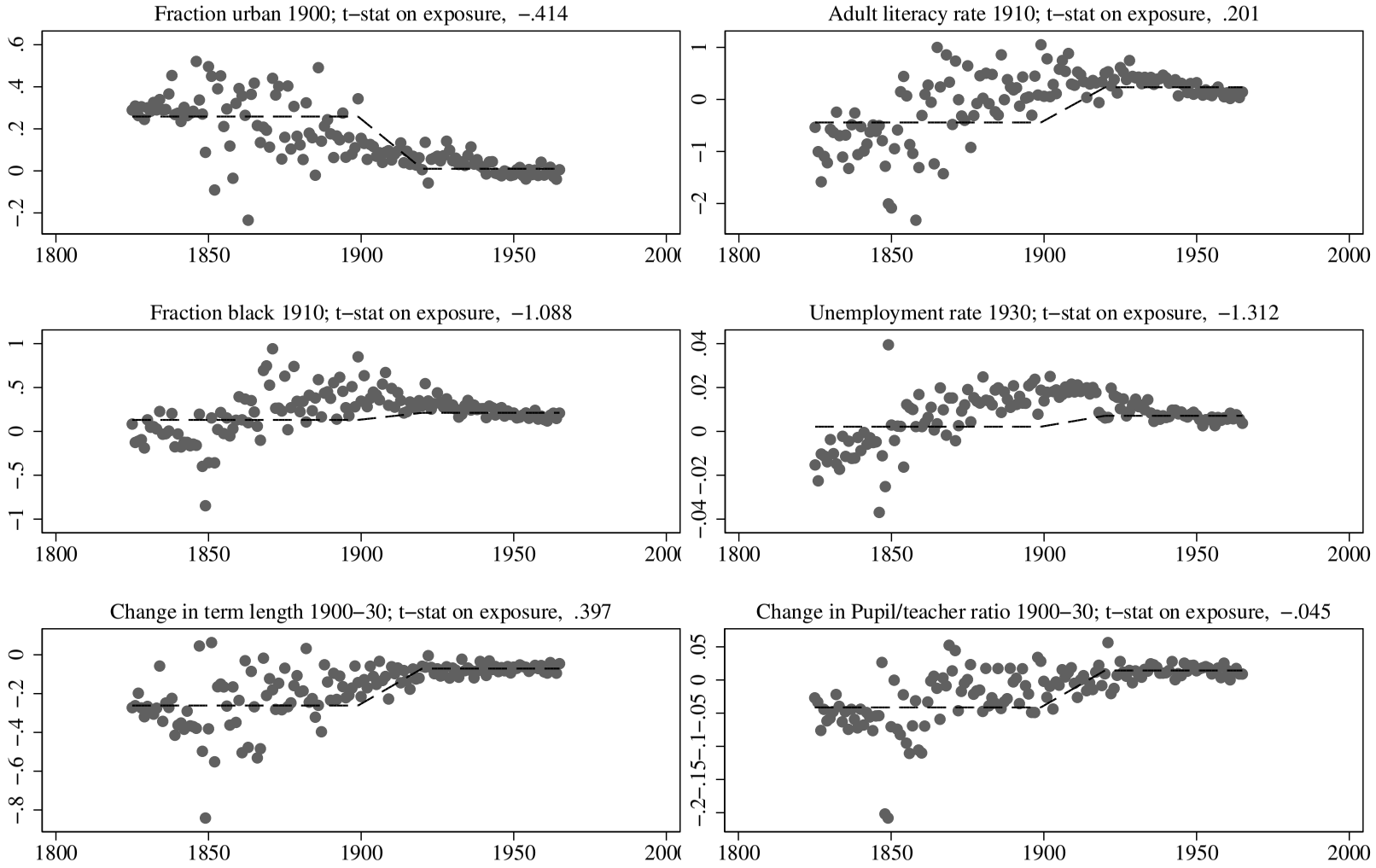
Notes: This table reports the simple and partial correlations between the malaria proxy (described in Appendix B) and the several control variables (described in Appendix C) for each country. Reporting of the ten region dummies for Colombia is suppressed, but these are included in the computation of the partial correlations. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. For comparability with the rest of the study, observations are weighted as in Tables 2, 3, 4.

(not for publication) Appendix Figure D – 1: Estimates on Controls, United States



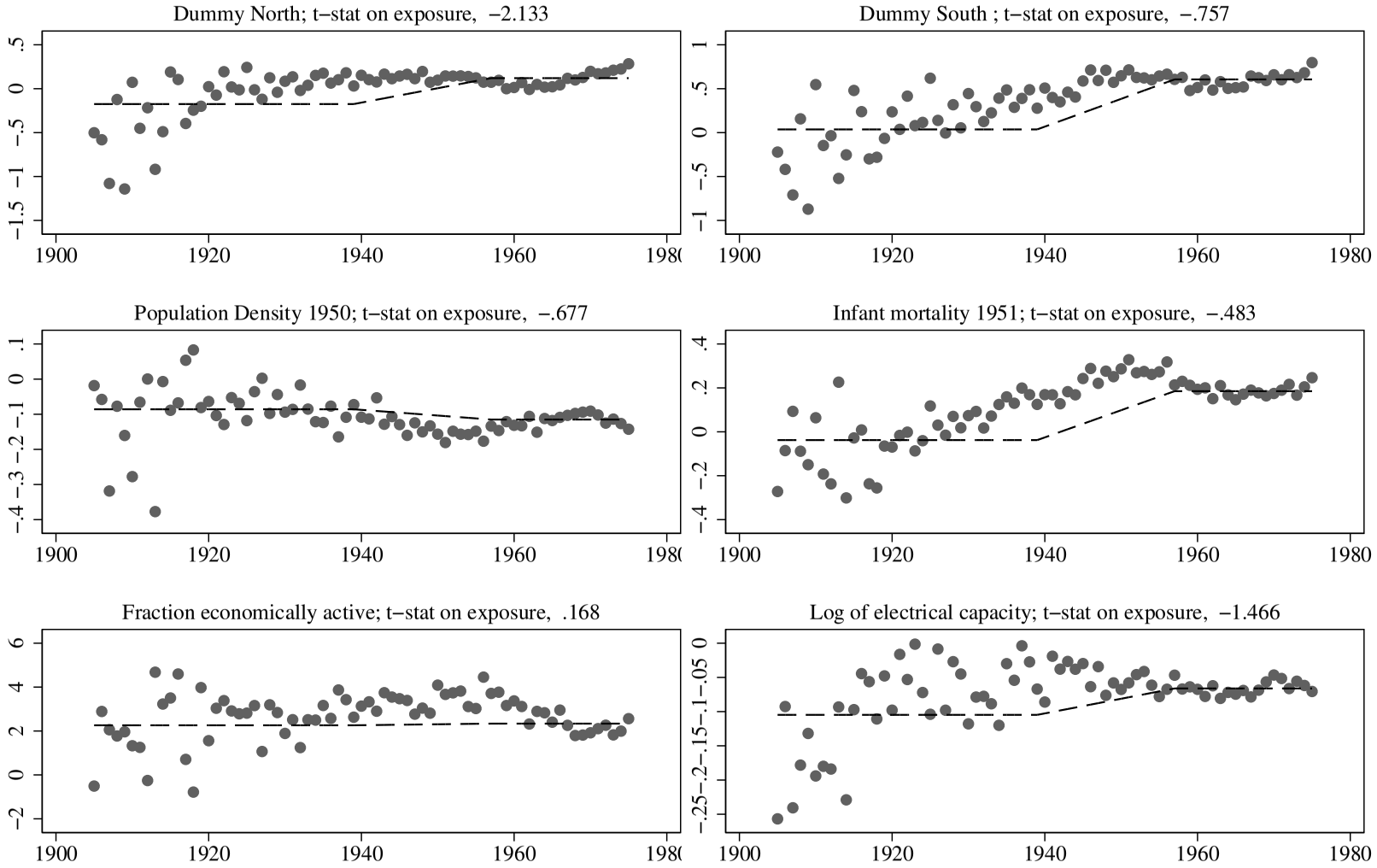
Note: Figure continues on next page.

(not for publication) Appendix Figure D – 1 (Continued): Estimates on Controls, United States



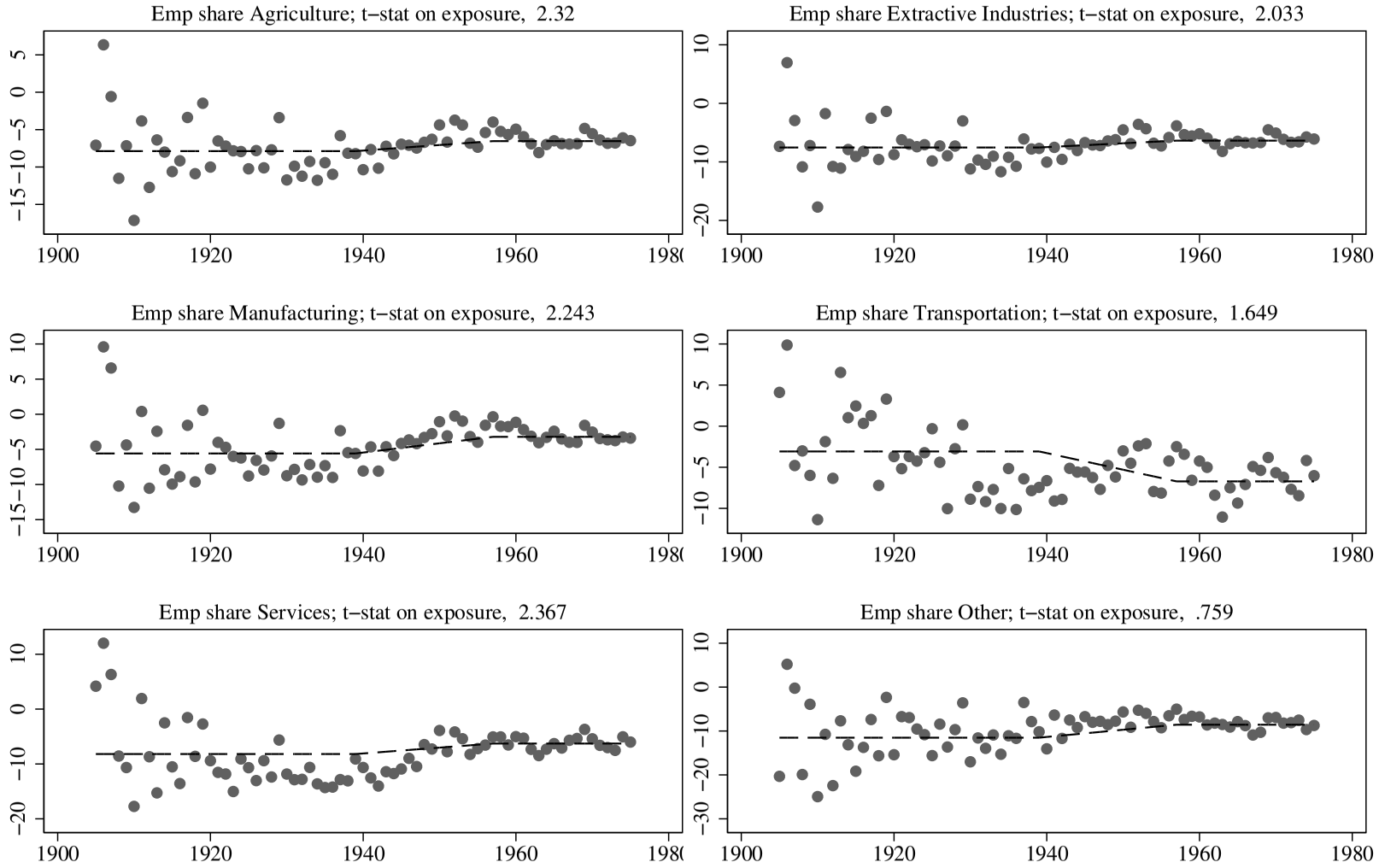
Notes: This figure plots the cohort-specific estimates on the controls variables from equation 1. Variables are defined in Appendix C.

(not for publication) Appendix Figure D – 2: Estimates on Controls, Brazil



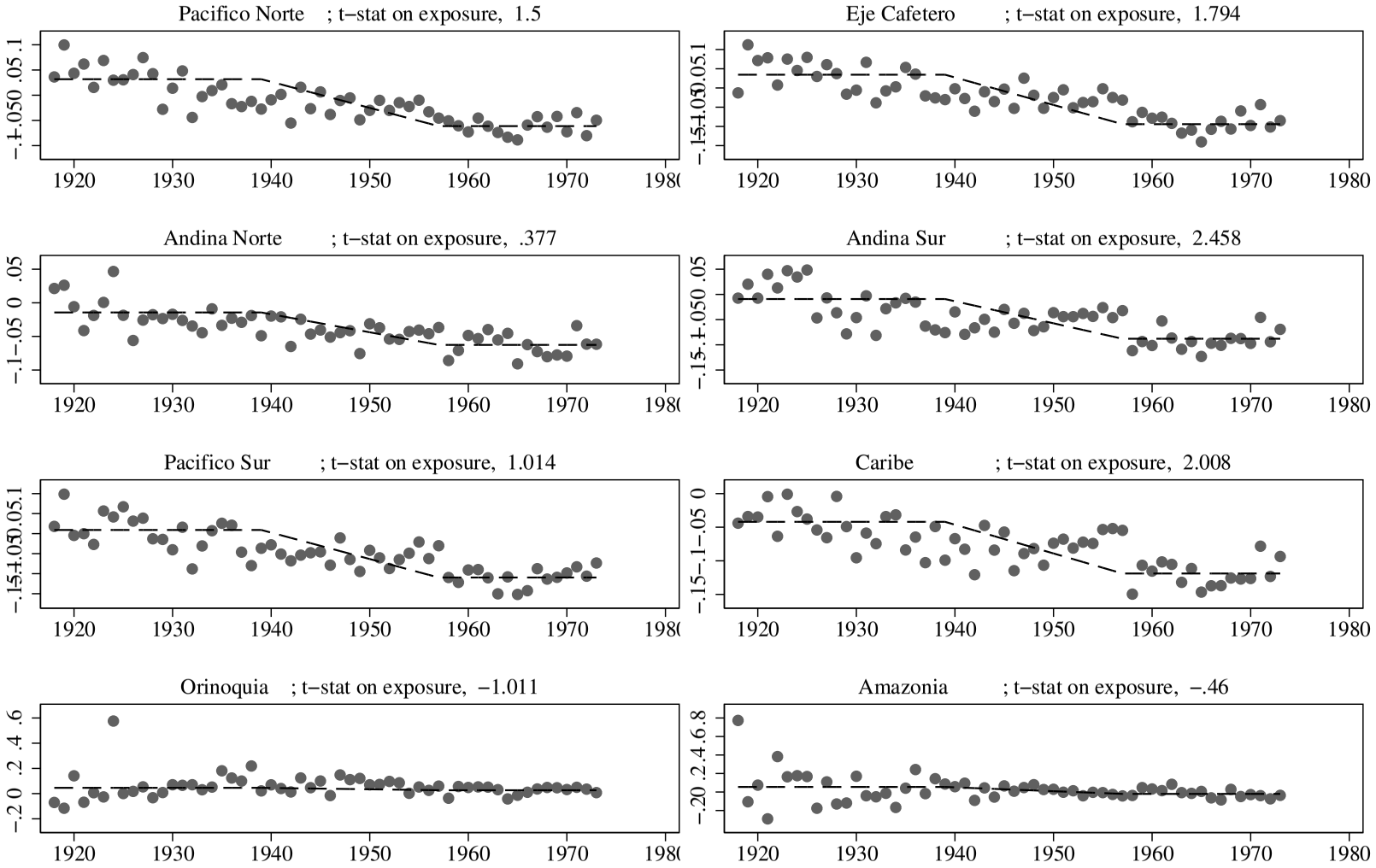
Note: Figure continues on next page.

(not for publication) Appendix Figure D – 2 (Continued): Estimates on Controls, Brazil



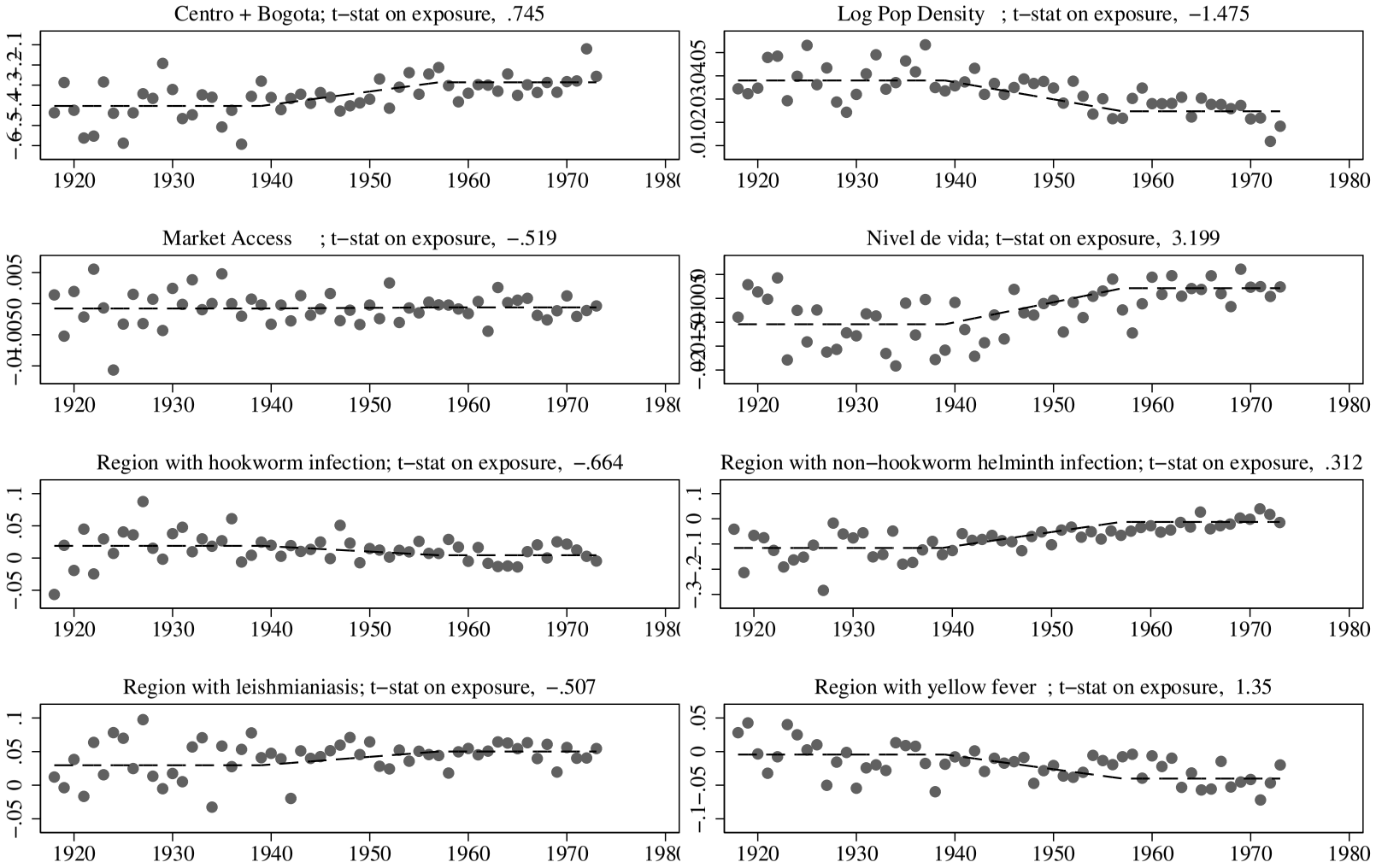
Notes: This figure plots the cohort-specific estimates on the controls variables from equation 1. Variables are defined in Appendix C.

(not for publication) Appendix Figure D – 3: Estimates on Controls, Colombia



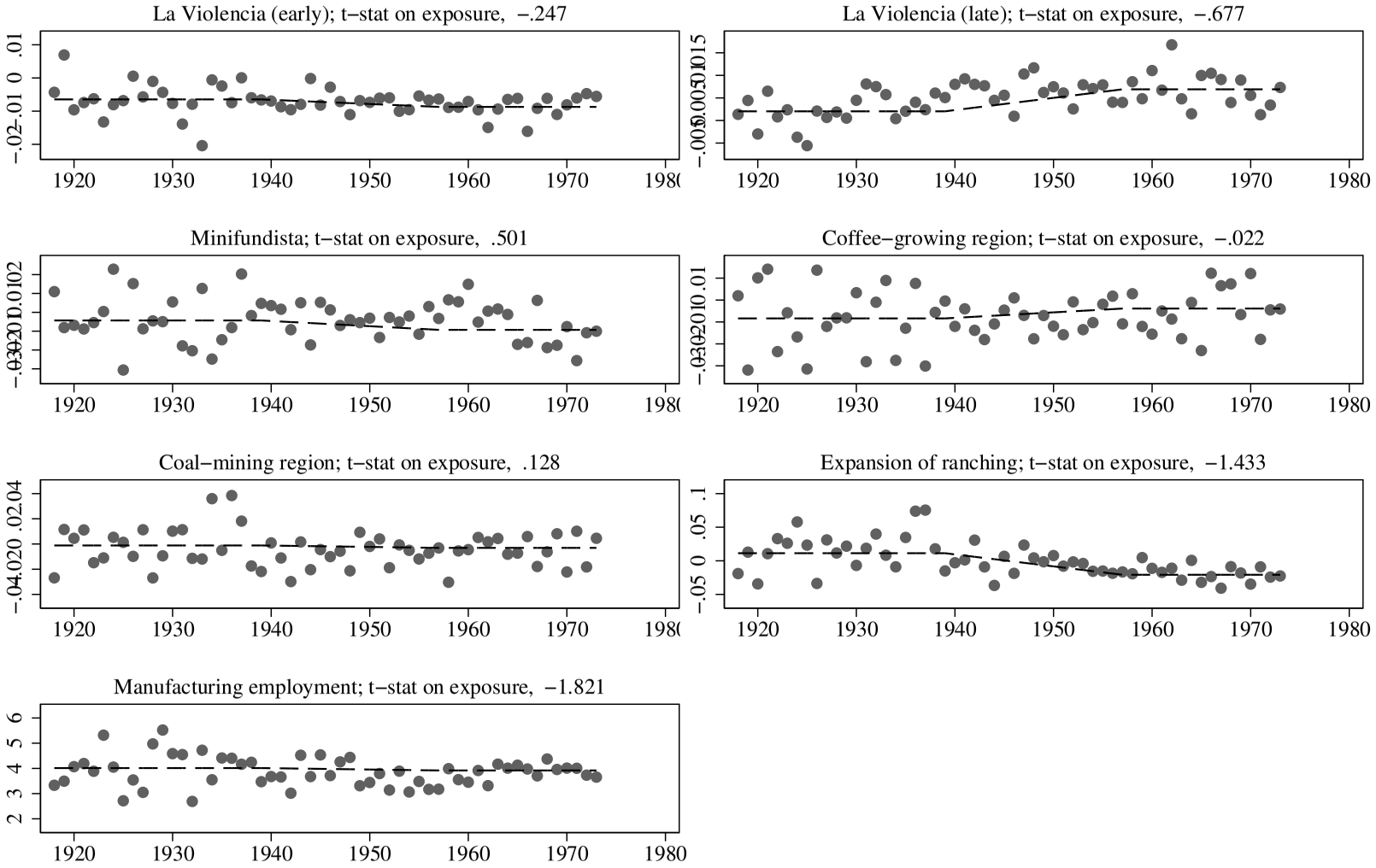
Note: Figure continues on next page.

(not for publication) Appendix Figure D – 3 (Continued): Estimates on Controls, Colombia



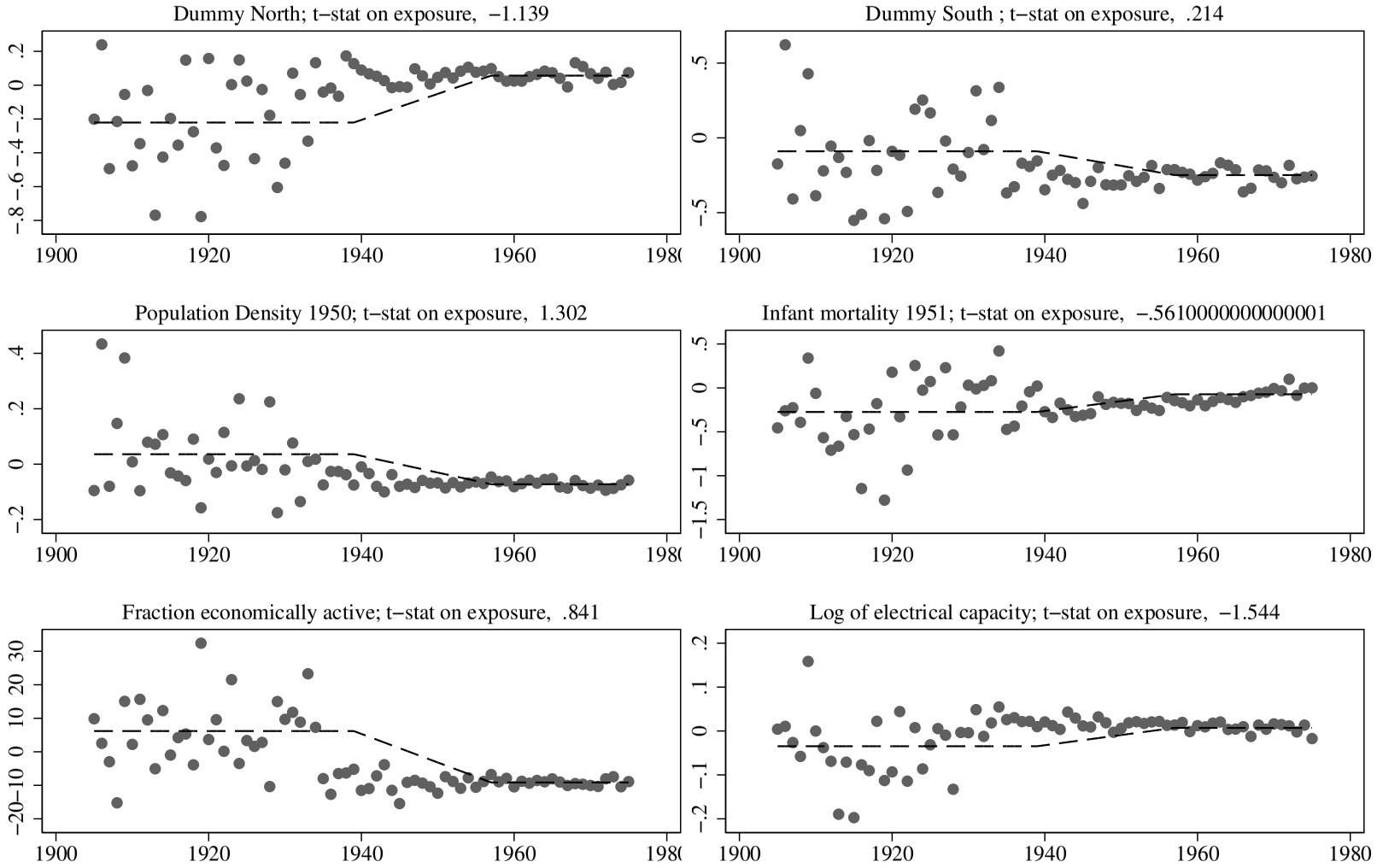
Note: Figure continues on next page.

(not for publication) Appendix Figure D – 3 (Continued): Estimates on Controls, Colombia



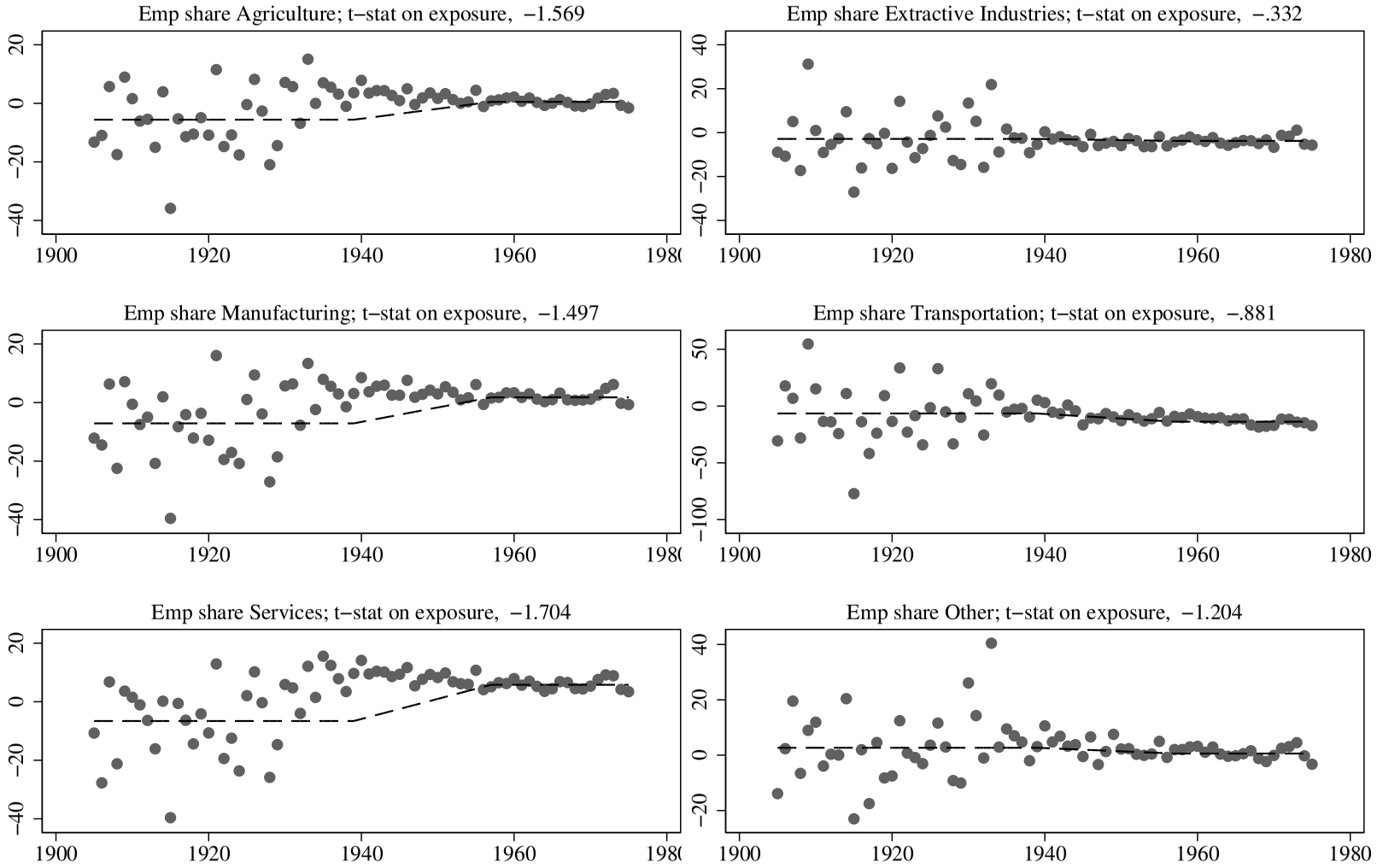
Notes: This figure plots the cohort-specific estimates on the controls variables from equation 1. Variables are defined in Appendix C.

(not for publication) Appendix Figure D – 4: Estimates on Controls, Mexico



Note: Figure continues on next page.

(not for publication) Appendix Figure D – 4 (Continued): Estimates on Controls, Mexico



Notes: This figure plots the cohort-specific estimates on the controls variables from equation 1. Variables are defined in Appendix C.

Appendix E: Panel Estimates of Childhood Exposure (not for publication)

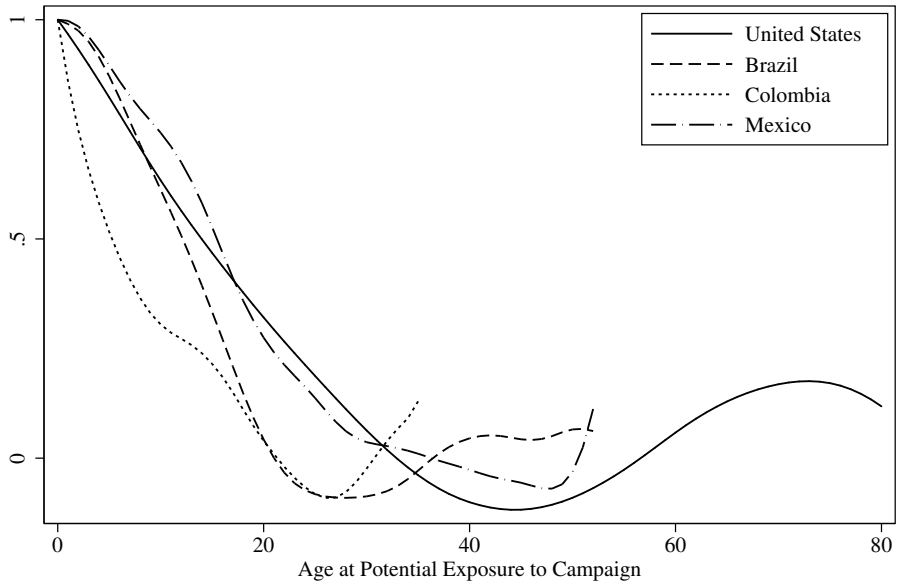
	(1)	(2)	(3)	(4)
Area of birth dummies:	No	Yes	Yes	Yes
Area of birth trends:	No	No	Linear	Quadratic
Outcome:				
	Panel A: United States			
Occupational Income Score	0.237 ***	0.118 ***	0.111 ***	0.126 ***
	(0.043)	(0.021)	(0.025)	(0.017)
Duncan's Index	0.305 ***	0.189 ***	0.121 ***	0.169 ***
	(0.063)	(0.047)	(0.026)	(0.024)
	Panel B: Brazil			
Log Total Income	0.525 ***	0.344 ***	0.338 ***	0.402 ***
	(0.090)	(0.108)	(0.088)	(0.094)
Log Earned Income	0.554 ***	0.226 ***	0.249 ***	0.261 **
	(0.196)	(0.088)	(0.082)	(0.122)
	Panel C: Colombia			
Industrial Income Score	0.092 ***	0.054 ***	0.054 ***	0.022
	(0.023)	(0.010)	(0.010)	(0.014)
	Panel D: Mexico			
Log Earned Income	0.324 **	0.155 *	0.151 ***	0.153 ***
	(0.140)	(0.085)	(0.048)	(0.047)

Notes: Each cell reports estimates, from a separate regression, of the childhood-exposure variable times pre-campaign malaria intensity using OLS. The outcome variables are as indicated in each row 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. Unlike the results in the main text, these parameters are estimated directly from cohort/panel data structured by year of birth, census year, and area of birth, the construction of which is described in Appendix A. The following equation is estimated:

$$Y_{jkt} = \tilde{\beta}\ \mathrm{M}_j \times \mathrm{Exp}_k + \delta_k + \delta_j + \delta_t + \sum_i x_j^i \times \mathrm{Exp}_k\ \gamma_i + \nu_{jkt}$$

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 is the ith state-of-birth control variable, and the δ_k, δ_t, and δ_j are fixed effects for year of birth, census year, and area of birth, respectively. The table reports estimates of $\tilde{\beta}$. 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×time average. Reporting of additional terms suppressed.

Appendix F: Smoothed Estimates by Age of Exposure to Campaign (not for publication)



Notes: This figure plots smoothed versions of the cohort-specific estimates of income on malaria from equation 1. Cohorts with full childhood exposure to the anti-malaria campaign are pooled at an age of first exposure of zero. Effect sizes are re-normalized for each country so that the estimate at full childhood exposure is 1 and that the mean for those at age of first exposure greater than 30 (25 for Colombia) is zero.