Spillovers and Aggregate Effects of Health Capital: Evidence from Campaigns Against Parasitic Disease in the Americas*

Hoyt Bleakley†

April 20, 2007

Abstract

The effect of population health on aggregate income remains an open question, despite the substantial literature showing positive and long-lasting effects of health on productivity at the individual and cohort levels. This study considers several eradication campaigns against parasitic disease in the Americas. Previous work indicated that cohorts exposed to these campaigns as children had higher productivity and human capital. As these more productive cohorts entered the labor force, average incomes rose as well, above and beyond that due to the changing cohort compositon of workers. This suggests, on net, the presence of positive spillovers from health capital. I also estimate the impact of average childhood exposure to the campaign on aggregate (state-level) income, where the effect is at least as large as the sum of estimated direct and spillover effects.

Keywords: Returns to health, externalities, eradication campaigns, hookworm, malaria.
JEL codes: I12, J24, 010, H43.

*Preliminary, please do not cite without permission. Results in flux, et cetera. Comments are most welcome.

†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] uchicago [dot] edu

1 Introduction

Despite a substantial body of microeconometric evidence pointing to positive and long-lasting effects of health on productivity at the individual and cohort level, there remains considerable doubt as to the effect of population health on income at the aggregate level. While many micro- or cohort-based studies capture partial-equilibrium (i.e., fixed price) effects of health, changing the health capital of large fractions of the population will probably have general-equilibrium effects: prices will move. If you add more effective labor (of the same type) to the economy, this might depress factor ratios (land/labor or capital/labor), and thus reduce wages. If you add healthier workers (of a different skill type), this increases the relative scarcity of the unhealthy and unskilled cohorts, and perhaps increases their income. On the other hand, if health capital is used for ranking and jobs are scarce, then the new, healthy cohorts might displace the old, unhealthy ones. Finally, the health capital, especially as it might support investment more of one own general human capital, might have positive external effects on average productivity in those economies.

This study considers several eradication campaigns against parasitic disease in the Americas:[1] hookworm in the United States (circa 1913), and malaria in the United States (circa 1920), Brazil, Colombia and Mexico (circa 1955). These campaigns began because of advances in health technology and increases in funding, both from outside the affected regions, and these features mitigate concerns about reverse causality. Areas with high hookworm and malaria burdens saw large drops in the diseases following the campaigns. Furthermore, for several reasons, children bore particularly large burdens from these diseases. Existing work (Bleakley, 2006 & 2007) suggests that those cohorts treated by the eradication campaign had higher income as adults than the preceding generation, and that these across-cohort changes coincided with childhood exposure to the campaigns rather than to pre-existing trends.

The goal of the present study is to consider the general-equilibrium impact when these healthier cohorts entered the labor market. First, what was the magnitude of the reduced-form spillover effects from the entrance of treated cohorts into the labor force, and by which channels (changing factor ratios, human-capital externalities, displacement, etc.)? Second, what was the long-term aggregate effect of these campaigns, which combines both the direct effect of childhood exposure to the campaigns plus any other associated multipliers? Because these questions hinge on the entrance of cohorts treated by the eradication campaigns, the central explanatory variable in the analysis is the fraction of the population treated, which I model as each area's pre-eradication disease burdens interacted with the labor force's average childhood exposure to the respective campaign. I relate the time path of average exposure to income variables across the years 1880–2000. Like all of Gaul,

[1] The campaigns are discussed in Section 2. For survey pieces on these diseases, see Hotez *et alia* (2006) for hookworm, and Wernsdorfer and McGregor (1988) and Nájera, Liese, and Hammer (1992) for malaria.

this stretch of years can be divided into three parts: an early period in which none of the labor force had any childhood exposure to the eradication campaigns, a late period in which the entire labor force had been potentially treated as children, and a middle period in which the treated cohorts gradually filled out the ranks of the prime working-age population.

The first questions is about spillovers, which I analyze using census micro data.[2] Controlling flexibly for changes in the composition of cohorts in the labor force, I estimate a positive reduced-form correlation between income and the entrance of treated cohorts into the labor force. Moreover, I show that this rise in cohort-adjusted income in areas with higher pre-campaign disease burdens coincides with average childhood campaign exposure and not with a pre-existing trend. These estimates imply reduced-form spillovers that are around one third to one half of the direct effect on treated cohorts. Next I interpret these results in a simple model with human-capital externalities and imperfect substitutability across skill types. These results are found in Section 4.

The second question is about the aggregate effect of population health, so I turn to state-level data on per-capita personal income for U.S. states. Aggregate income per capita rises appreciably as cohorts with greater childhood exposure to the eradication efforts enter the labor force. The estimates confirm the previous results as they are of similar magnitude to the estimated sum of direct and spillover effects. And, as with the results for spillovers, these rises in income coincide roughly with the time path of treated cohort's presence in the labor force, and are not sensitive to controlling for pre-existing state-specific trends or the inclusion of variety of control variables. The magnitudes of the estimates suggest that eradication of each disease increased aggregate income in highly infected by around 30%. (For some context, consider that the log-income gap between the North and South in the U.S. was about 0.75 in the early 20th century.) These results are found in Section 6.

2 Preliminaries

2.1 Eradication Campaigns

2.1.1 The Southern United States

The efforts to eradication hookworm and malaria from the Southern U.S. had their roots in critical innovations to both knowledge and spending, much of which was from outside the region.[3] For example, the discovery of the transmission mechanism for hookworm was made in the 1890's by a European doctor whose initial 'experimental evidence' consisted of accidentally infecting himself while diagnosing a patient. At that time, hookworm infection in the American South was not

[2]The data are described in Section 3 and Appendices A–C.

[3]The historical presentation in this section draws heavily on the work of Ettling (1981) for hookworm and Williams (1951) and Harrison (1978) for malaria.

even recognized as a problem. The Federal government's initial large-scale involvement in malaria control came as a result of the effort to build a canal across the Panamanian isthmus. Knowledge generated in this effort (especially regarding the control of water and the use of insecticides) was later successfully applied to the malarious regions of the South.

The hookworm-eradication campaign (c. 1910) began soon after (i) the discovery that a variety of health problems among Southerners could be attributed to the disease and (ii) the donation by John D. Rockefeller of a substantial sum to the campaign. The Rockefeller Sanitary Commission (RSC) surveyed infection rates in eleven Southern states, and found that an average of forty percent of school children suffered from hookworm. The RSC then sponsored treatment "dispensaries" and education programs throughout the region. Follow-up studies of infection rates indicate that this campaign brought about a substantial immediate reduction in hookworm disease. (Bleakley, 2007, discusses this campaign in greater detail.)

The introduction of new knowledge about malaria control combined with a surge of interest from outside the region (mostly in the form of federal spending) to finally mount an effective attack on the problem around 1920. The region saw a decline in malaria mortality of around two thirds in the 1920s, and, following a comparatively small resurgence in the Great Depression, a reduction to negligible levels by the 1940s. Although there was probably not any single intervention that lead to the decline of malaria infection in the South, I argue that a confluence of factors substantially accelerated its eradication. (See Bleakley, 2006, for more on this.)

2.1.2 Brazil

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. The main innovation was a new chemical: dichloro-dipenyl-trichloro-ethane, or DDT. Rediscovered in the 1940s, DDT was shown to be of extraordinary value as a pesticide: it rapidly killed a variety of insects and had no immediately apparent effects on mammals. The use of DDT was valuable to the Allied war effort, and after the war, 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 the 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.[4] 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.

Most Latin American countries (including Brazil) mounted malaria eradication campaigns, and saw large declines in malaria prevalence. Throughout Latin America, the campaign ultimately proved inadequate to the task, and, in many areas, malaria partially resurged two decades later. But in almost all parts of the hemisphere, malaria never returned to its levels from before the application of DDT.

2.2 Research Design

The first factor for identifying the effect of the eradications that different areas of these countries had distinct incidences of the diseases studied. In general terms, this meant that the residents of the coastal plain of the South and the Brazilian north were much more vulnerable to infection than were their compatriots in the rest of the country.[5] 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.[6] This heterogeneity lends itself to a treatment/control strategy.

The second component in the research design is that the commencement of eradication was substantially due to factors external to the affected areas. The accelerated eradication was due to critical innovations to knowledge and increases in spending from outside the region. This contrasts with explanations that might have potentially troublesome endogeneity problems, such as, for example, positive income shocks in the endemic 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.

The introduction of this treatment (broadly defined) combines with the cross-area differences in pretreatment infection rates to form the identification strategy. By comparing the evolution of outcomes (by both cohort and time) across areas with distinct infection rates, we can assess the contribution of the eradication campaigns to the observed changes.

Moreover, the timing of the eradication campaign should induce variation in childhood malaria infection that has a marked pattern across year-of-birth cohorts. And childhood exposure to these

[4]DDT had residual action as a pesticide, meaning that its effect was sufficiently persistent after spraying that only 1-3 applications per year were mandated.

[5]In the case of hookworm, soil type and temperature were key determinants of whether the hookworm larvae survived and went on to infect human hosts. On the other hand, areas that had more slow-moving water were the preferred "nursery" for mosquitoes, the "vector" that transmitted malaria. Malaria infection was especially acute in the Mississippi and Amazon Deltas.

[6]Bleakley (2006 and 2007) presents evidence on this point. Moreover, this assertion is almost mechanically true since these diseases were more or less completely eradicated by mid-century.

diseases is thought to have particularly deleterious impacts. Indeed, Bleakley (2006 and 2007) shows that shift in the disease-income relationship coincides with childhood exposure to the eradication efforts. This can be seen graphically in this Panels A and B of Figure 1. This figure compares changes in income by cohort across areas with distinct intensities in order to assess the contribution of the eradication campaign to the observed changes.[7] 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 hookworm (Panel A) and malaria (Panel B). Each year-of-birth cohort's point estimate is marked with a dot. As is seen in the figure, a simple model of childhood exposure to the campaign (the dashed line) provides an adequate fit to the data, suggesting that those cohorts' having escaped a childhood of infection by hookworm and malaria paid off in adulthood.

But a problem with this analysis is that it might be inappropriate to assume that the earlier cohorts were not indirectly affected by these interventions. Specifically, there might be general equilibrium effects of introducing all these new, treated cohorts into the labor market. Therefore, the research design of the present study will be to relax this fixed-price assumption and examine income by state and state-of-birth that is related to the entrance of treated cohorts into the working-age population. Panel C of Figure 1 shows the time path of average childhood exposure to these campaigns.

There are two additional advantage of examining these campaigns. First, the interventions considered in the present study were of such a scale that they brought about large declines in hookworm and malaria infection in entire areas. In the context of economic development, it is precisely such a large and persistent reduction in disease burden that we would wish to consider. Second, enough time has passed since their inception that we can assess long-term consequences.

2.3 Related Literature

An important literature considers the short- and medium-run effects on income of hookworm, malaria, and related health conditions. Thomas *et al.* (2003) show that anemia has depressing effects on contemporaneous adult productivity. Numerous studies address the economic loss following malarial fevers: *inter alia*, Conly (1975) presents such an analysis in Paraguay, while Bonilla

[7] For each year of birth, OLS regression coefficients are estimated on the resulting cross section of states/municipios of birth. Consider a simple regression model of an average outcome, Y_{jk}, for a cohort with state of birth j and year of birth k:

$$Y_{jk} = \beta_k^H \; H_j^{pre} + \beta_k^M \; M_j^{pre} + \delta_k + X_j \; \Gamma_k + \nu_{jk} \tag{1}$$

in which M_j^{pre} 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. This specification allows one to examine how the relationships between income and pre-eradication malaria and hookworm ($\hat{\beta}_k^M$ and $\hat{\beta}_k^H$, respectively) differ across cohorts. (Note that the coefficients on the control variables are similarly flexible by year of birth.) See Bleakley (2006) for more details.

Castro, Kuratomi, Rodríguez, and Rodríguez (1991) consider a village 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. Shapiro (1918) examines the productivity increases among laborers in Costa Rica following deworming treatments.

A number of studies consider instead the long-term, cohort-level impact of hookworm and malaria on income and human capital. As mentioned above, Bleakley (2006 and 2007) uses a retrospective/cohort design to conclude that childhood exposure to the hookworm-eradication campaign increased adult income. Using a database of Union Army veterans, Hong (2007) finds a negative 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. Using interannual weather variation, Barreca (2007) shows that malaria exposure during infancy reduced years of schooling in the early 20th century U.S. South, although his does not find significant evidence of effects on adult income. Bleakley (2003, 2006, 2007) finds that eradicating hookworm and malaria increased literacy and regular school attendance in the episodes studied, but obtains mixed evidence on completed years of schooling. Bleakley and Lange (2006) consider the hookworm-related increase in returns to schooling in a quantity-quality model, and examine the fertility behavior of households in response to hookworm eradication.

There is a related body of studies with cohort-level designs that examines the impact of *in utero* and infant health on adult outcomes. The 'grandaddy' of this literature is perhaps the famous natural experiment of the Dutch hunger winter. (See Stein *et al.* 1975 for a review.) More recently, Almond (2006) shows that *in utero* exposure to the influenza pandemic had a variety of negative effects on later-life outcomes. Maccini and Yang (2006) analyze long-term effects of infant exposure to rainfall shocks in Indonesia. Using local resources in childhood as an instrument for adult height, Ribero and Nuñez (2000) consider the effect of health endowments on income in Colombia. Furthermore, a number of studies use fixed-effects estimators based on twins to measure the long-run effect of birth weight (Behrman and Rosenzweig (2004); Black, Devereux, and Salvanes (2006); Royer (2006)).

A smaller literature treats the impact of hookworm and malaria on aggregate income. Brinkley (1994, 1997) examines the role hookworm played in agricultural productivity in the U.S. He finds a negative conditional correlation between hookworm infection and agricultural income per capita, although he does not specifically use the RSC intervention to identify this relationship. Hong (2007) finds a large and statistically significant relationship between malaria and wealth accumulation across U.S. counties *circa* 1850. Utzinger, Tozan, Doumani, and Singer (2001) argue that the control of malaria transmission was a key factor in the development of Zambian copper mining. Furthermore, numerous authors have argued that the control of malaria in the Panamian isthmus

was crucial for the successful completion of the canal there. Sachs (2001) uses cross-country data to measure the correlation between malaria and economic development. Bloom, Canning, and Sevilla (2004) find substantial effects of broader measures of health on income across countries, and and indeed many studies in the economic-growth literature report robust effects of life expectancy (as a control variable) on GDP. More recently, Acemoglu and Johnson (2006) consider the effect of life expectancy on national income using fixed-effect and instrumental-variables estimators. They find estimates that are much smaller relative to the earlier literature.

Finally, a related literature attempts to measure any externalities associated with education in particular labor markets. Moretti (2004) reviews these studies.

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 census data from the U.S. and Brazil.

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 pre-eradication disease intensity in an individual's state of birth to conduct the analysis, which is therefore an intention-to-treat design.

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 25 and 55, inclusive, for the census years 1880-1990, which includes cohorts with years of birth ranging from 1825 to 1965. 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.

The data on native-born males from the Brazilian IPUMS-coded censuses from 1960 to 2000 are similarly pooled, resulting in birth cohorts from 1905 to 1975. These censuses contain questions

on literacy, years of education, and income (both total and earned).

I combine microdata from various censuses to construct panels of average outcomes by cohort. Cohorts are defined by census year (of observation), year of birth and state of birth. To construct these panels, I pool the micro-level census data. I then take average incomes for each cell defined by year, year of birth and state of birth. (This procedure is described in detail in Appendix A.)

Disease data are drawn from a variety of sources. U.S. data on malaria are reported from by the Census (1894), Maxcy (1923), and later in the *Vital Statistics* (Census, 1933). Hookworm infection by state is reported by Kofoid and Tucker (1921). For Brazil, data on malaria ecology are derived from Gallup, Mellinger, and Sachs (1999a). The ecology data were matched with states using a geographic information system (GIS). Appendix B contains further details.

The aggregate income data for the United States are state per-capita personal income, adjusted for price differences across areas (but not over time). The 1880-1980 data are drawn from Mitchener and McLean (2003) and extended to 2000 using data from the BEA.

4 Spillovers

4.1 Reduced-form effects

Reduced-form estimates suggest that the entrance of treated cohorts into the labor force led to increases average income, above and beyond the rise due simply to changing cohort composition. The coefficients measure the combination of external effects that are both pecuniary (working through the relative supplies of labor types) and nonpecuniary (spillovers from human capital, e.g.). The basic regression specification is as follows.

$$y_{tjk} = \beta_H \left(H_j^{pre} \times \bar{E}_t^H \right) + \beta_M \left(M_j^{pre} \times \bar{E}_t^M \right) + \delta_{jk} + \delta_{tr} + \sum_t \beta_t w_j^{1909} + \varepsilon_{tjk} \tag{2}$$

in which j is state of birth, t is census year, k is year of birth, and r is census region. The variables H_j and M_j represent the pre-campaign hookworm and malaria intensities, respectively, in area of birth j, and average childhood (temporal) exposure, in year t, to the respective campaigns among the prime working-aged population are $\bar{E}_t^H$ and $\bar{E}_t^M$. Effects due to changing cohort composition are absorbed by the δ_{jk}, while region-specific time effects load onto the δ_{tr}. Finally, w_j^{1909} is a measure of unskilled wages in 1909 from Lebergott (1964), so $\sum_t \beta_t w_j^{1909}$ allows for generic processes of post-1909 mean reversion across areas. The error term (ε_{tjk}) is assumed to be correlated within state of birth.

Controlling for cohort effects, I estimate that average incomes roses with the entrance of treated cohorts into the labor force. These results are found in Panel A of Table 1, which contains OLS estimates of equation 2 and variants thereof. Panel A, Column 1 displays estimates from the basic

specification for the United States. The regressions allow for other trends that might be spurious correlated with pre-eradication hookworm or malaria. The first strategy, adopted in Columns 2–4, is to include a term $tX_j'\Gamma$, in which t is a trend and X_j is a vector of other state-of-birth controls.[8] (The first-order effects of X_j are absorbed by the cohort fixed effects, δ_{jk}.) Results are similar if these state-level controls are instead interacted with average exposure to either campaign. Column 3 also allows for the interaction of the pre-eradication diseases rates; i.e., the regression also includes these terms: $\gamma_H \left(H_j^{pre} \times t\right) + \gamma_M \left(M_j^{pre} \times t\right)$. Finally, the regression summarized in Column 4 allows for state-specific linear trends ($\sum_j \beta_j t$ is added to the specification). For each specification, there is statistically significant evidence that the average incomes rose with the entrance to the labor market of the cohort exposed to the anti-hookworm and anti-malaria efforts in the U.S. South.

The time pattern of the income-disease relationship corresponds approximately to the entrance of treated cohorts to the working-age population. To make this comparison, the specification of equation 2 is augmented to estimate year-specific interactions of hookworm and malaria, and these estimated coefficients for the United States are plotted in Figure 2. Results from the basic and extended specification are plotted using long- and short-dashed lines, respectively. (The additional controls for the extended specification are also fully interacted with year dummies to put them on an equal footing with the disease variables.) Shifts in the coefficients are evident in the year-specific estimates for hookworm and malaria, and these are seen to coincide with the labor force's average level of childhood exposure to each respective disease (shown as the solid line in each graph). Prior to the entrance of these treated cohorts, the hookworm coefficients are stable, while there is some evidence of a downward trend in the malaria point estimates. Once this pre-trend is corrected for, the malaria coefficients bear a striking similarity to the average exposure variable. I then take the time series of coefficients displayed in Figure 2 and regress them on a trend and average exposure. These results are found in Panels B and C of Table 1. Estimates for hookworm match those found with the one-step estimates above (in Panel A), and the results for malaria are also similar once trends are accounted for. I find similar results for Brazil, although the quality of the evidence is limited by the smaller range of years for which micro data are available. This is seen in Figure 3, which plots time interactions of malaria, and the rightmost columns of Table 1.

These results imply spillover effects of non-trivial magnitudes. These numbers are shown in curly brackets throughout Table 1. To get a better sense of the reduced-form differences across areas, the parameter estimates are rescaled in each case by the difference in disease rates between the 90th and 10th percentile areas. The numbers for the U.S. are further rescaled by the mean occupational income score from cohorts born in the South between 1875 and 1895. These renormalizations show

[8] These additional controls consist of child mortality in 1890 and 1935, the number of doctors per capita in 1898, the 1910 adult literacy rate, 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.

faster income growth across the range of pre-eradication disease rates to the tune of four and eleven percent in the U.S. and one and four percent in Brazil.

4.2 Decomposition

[TO COME. Sketch: simple CES production function with different labor types (experience, education) as inputs. Allow for both imperfect substitutability among workers and externalities from the average stock of health capital, as proxied by average childhood exposure to the campaigns.]

5 Additional channels

[TO COME. Eventually work in data on migration, capital stock, females, blacks, labor-force participation.]

6 Aggregate Income

Examining state personal income per capita allows us to measure the total (direct + indirect) effect of population health on income. Consider basic regression equation for the state aggregate data:

$$\ln y_{tj} = \tilde{\beta}_H \left(H_j^{pre} \times \bar{E}_t^H \right) + \tilde{\beta}_M \left(M_j^{pre} \times \bar{E}_t^M \right) + \tilde{\delta}_j + \tilde{\delta}_{tr} + \sum_t \tilde{\beta}_t w_j^{1909} + \nu_{tj} \tag{3}$$

which is quite similar to equation 2, except for the lack of decomposition by cohort (k). This difference changes the interpretation of the average-exposure coefficients ($\tilde{\beta}_H$ and $\tilde{\beta}_M$). In particular, because there are no longer any δ_{jk} to absorb the direct effects of cohort composition, the $\tilde{\beta}$ measure both the direct effect of childhood exposure to the campaign and any indirect effects through spillovers (for example in the labor market). As above, the interactions of 1909 wages with year dummies ($\sum_t \beta_t w_j^{1909}$) flexibly control for possible convergence across states, and region-specific time effects are absorbed by the δ_{tr}. State dummies $\tilde{\delta}_j$ control for time-invariant effects by state of birth.

Aggregate income per capita rises appreciably as cohorts with greater childhood exposure to the eradication efforts enter the labor force. These results are seen in Table 2, where the first column of Panel A contain estimates of equation 3. Interactions of pre-campaign disease levels with average temporal exposure to the campaigns are both positive and statistically significant. As above, I also augment the specification to allow for pre-campaign disease levels to interact with year (in the second column) and for state-specific linear trends (in the third column). In the next three columns, I include the expanded set of controls from above, interacted with year in the regression. And in the final three columns of Panel A, I instead use Mitchener and McLean's (2003) set of "fundamental

rather than proximate" determinants[9] of income as the augmented control set (interacted with trend as before). Results are similar (i.e. within a standard error) to the baseline.

These rises in income coincide roughly with the time path of treated cohort's presence in the labor force, as seen in Figure 4. The solid line displays, once again, the average level of childhood temporal exposure to eradication for each disease. I add year-specific interactions of pre-eradication hookworm and malaria to the specification of equation 3, and these estimated coefficients for the United States are plotted as diamonds. (As above, the additional controls enter the specification also fully interacted with with year dummies for parity with the treatment of the disease variables.) The timing of the shifts in the point estimates corresponds with the ramping up of the fraction of adults potentially treated during childhood. I test this formally and display the results in Panels B and C of Table 2. The time series of disease coefficients is indeed related to average childhood exposure in a statistically significant way, even when controlling for a linear time trend. These estimates are also similar to those found using the one-step estimator in Panel A.

These estimates suggest large aggregate effects from the eradication of hookworm and malaria in the U.S. Renormalized parameter estimates are shown throughout Table 2 in curly brackets. The estimates are rescaled in each case by the difference in disease rates between the 90th and 10th percentile areas, and furthermore by the mean occupational income score from cohorts born in the South between 1875 and 1895. According to these estimates, over the course of the half century following the campaigns, eradication increased per-capita personal income in the most affected states by 23–38% in the case of malaria and 25–33% in the case of hookworm. These numbers amount to approximately 3–4 times the size of the measured time effect (from Section 4) and about twice the estimated effect at the cohort level (from Bleakley, 2006 and 2007). The aggregate effects would appear to be larger than the cohort effects, and perhaps even larger than the direct (cohort) effect plus the spillover onto other cohorts born in the same state.

7 Conclusion

[TO COME]

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.

[9]These include variables from their baseline specification: percentage of workforce in mining, 1880; percentage of population in slavery, 1860; a dummy for access to ocean or Great Lakes; the average number of cooling degree days; and dummies for French, Spanish, or Dutch colonial origins.

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.

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

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

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.

Behrman, Jere R. and Mark R. Rosenzweig (2004). "Returns to Birthweight." *Review of Economics and Statistics.* **86** (2), 586–601.

Bleakley, C. Hoyt (2007). "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics.* 02 **122** (1), 73–117.

Bleakley, Hoyt (2006). "Malaria in the Americas: A Retrospective Analysis of Childhood Exposure." Documento CEDE 2006-35, Universidad de los Andes, Bogotá September.

——— and Fabian Lange (2006). "Chronic Disease Burden and the Interaction of Education, Fertility and Growth." Working Paper 121, Bureau for Research and Economic Analysis of Development, Cambridge, Mass. May.

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.

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

——— (1997). "The Decline in Southern Agricultural Output, 1860-1880." *Journal of Economic History.* March **57** (1), 116–38.

Conley, Timothy G., Fredrick Flyer, and Grace R. Tsiang (2003). "Spillovers from Local Market Human Capital and the Spatial Distribution of Productivity in Malaysia." *Advances in Economic Analysis & Policy.* **3** (1). article 5.

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

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.

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).

Hotez, Peter J., Donald A. P. Bundy, Kathleen Beegle, Simon Brooker, Lesley Drake, Nilanthi de Silva, Antonio Montresor, Dirk Engels, Matthew Jukes, Lester Chitsulo, Jeffrey Chow, Ramanan Laxminarayan, Catherine M. Michaud, Jeff Bethony, Rodrigo Correa-Oliveira, Xiao Shu-Hua, Alan Fenwick, , and Lorenzo Savioli (2006). "Helminth Infections: Soil-Transmitted Helminth Infections and Schistosomiasis." in D. T. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, A. Mills, and P. Musgrove, eds., *Disease Control Priorities in Developing Countries*, 2 ed., Washington: World Bank, April chapter 24, pp. 467–482. Online at http://files.dcp2.org/pdf/DCP/DCP25.pdf.

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.

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.

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.

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.

Mitchener, Kris James and Ian W. McLean (1999). "U.S. Regional Growth and Convergence, 1880-1980." *Journal of Economic History.* December **59** (4), 1016–42.

——— and ——— (2003). "The Productivity of US States since 1880." *Journal of Economic Growth.* **8**, 73–114.

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.

Moretti, Enrico (2004). "Human Capital Externalities in Cities." in J. V. Henderson and c. T. Jacques-Fran eds., *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: North-Holland, chapter 51, pp. 2243–2291. Online at http://www.sciencedirect.com.proxy.uchicago.edu/science/handbooks/15740080.

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.

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

Philipson, Tomas J. (2000). "External Treatment Effects and Program Implementation Bias." Working Paper T0250, National Bureau of Economic Research, Cambridge, Mass. January.

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.

Royer, Heather (2006). "Separated at Girth: Estimating the Long-Run and Intergenerational Effects of Birthweight using Twins." Unpublished manuscript, Case Western Reserve University, Cleveland December. Online at http://wsomfaculty.case.edu/hRoyer/twins.pdf.

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.

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.

Stein, Zena, Mervyn Susser, Gerhart Saenger, and Francis Marolla (1975). *Famine and Human Development: The Dutch Hunger Winter of 1944-1945*, London: Oxford University Press.

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.

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.

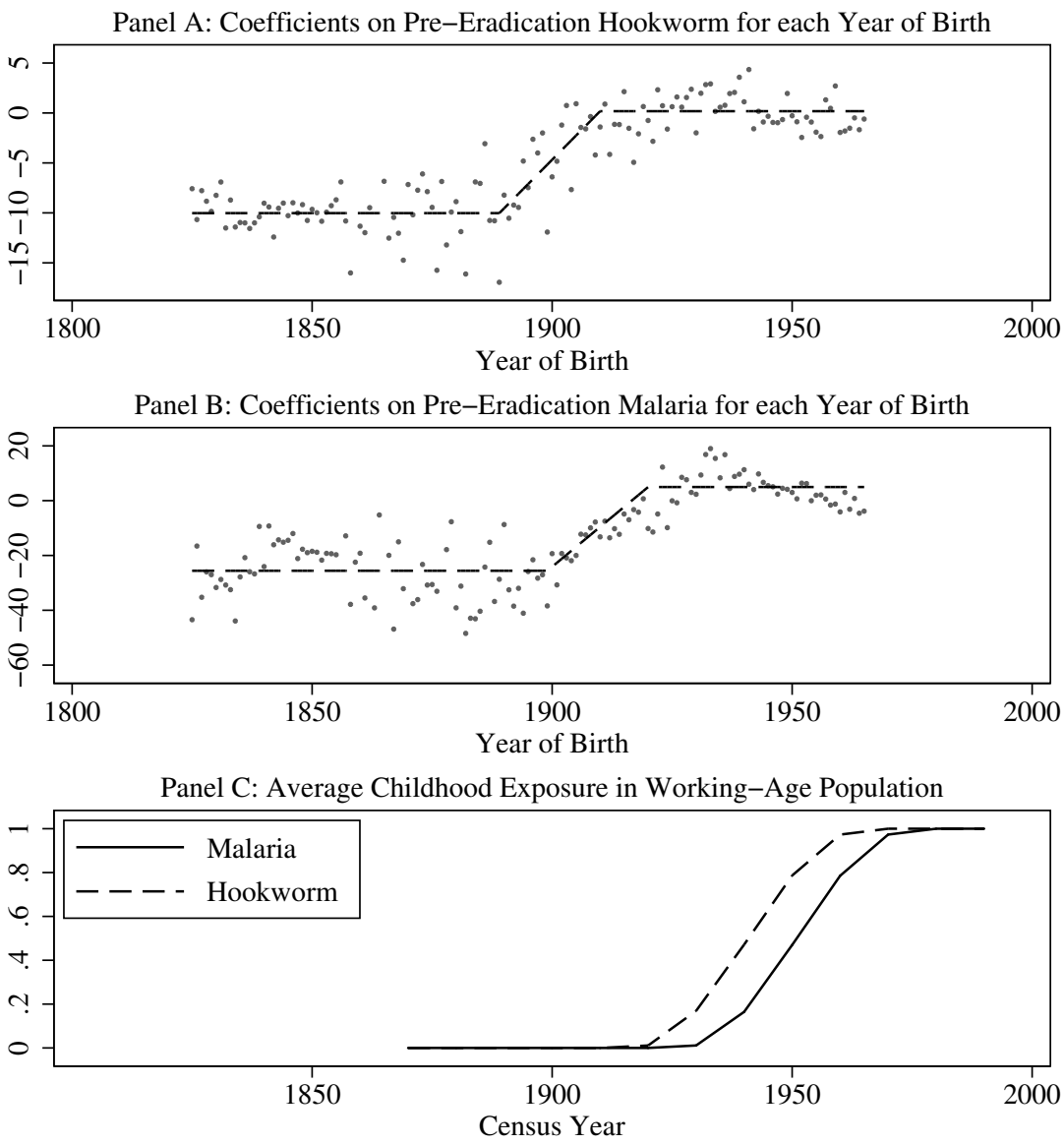
Weisbrod, Burton A., Ralph L. Andreano, Robert E. Baldwin, Erwin H. Epstein, and Allen C. Kelley (1973). *Disease and Economic Development: The Impact of Parasitic Diseases in Saint Lucia*, Madison, Wisconsin: The University of Wisconsin Press.

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.

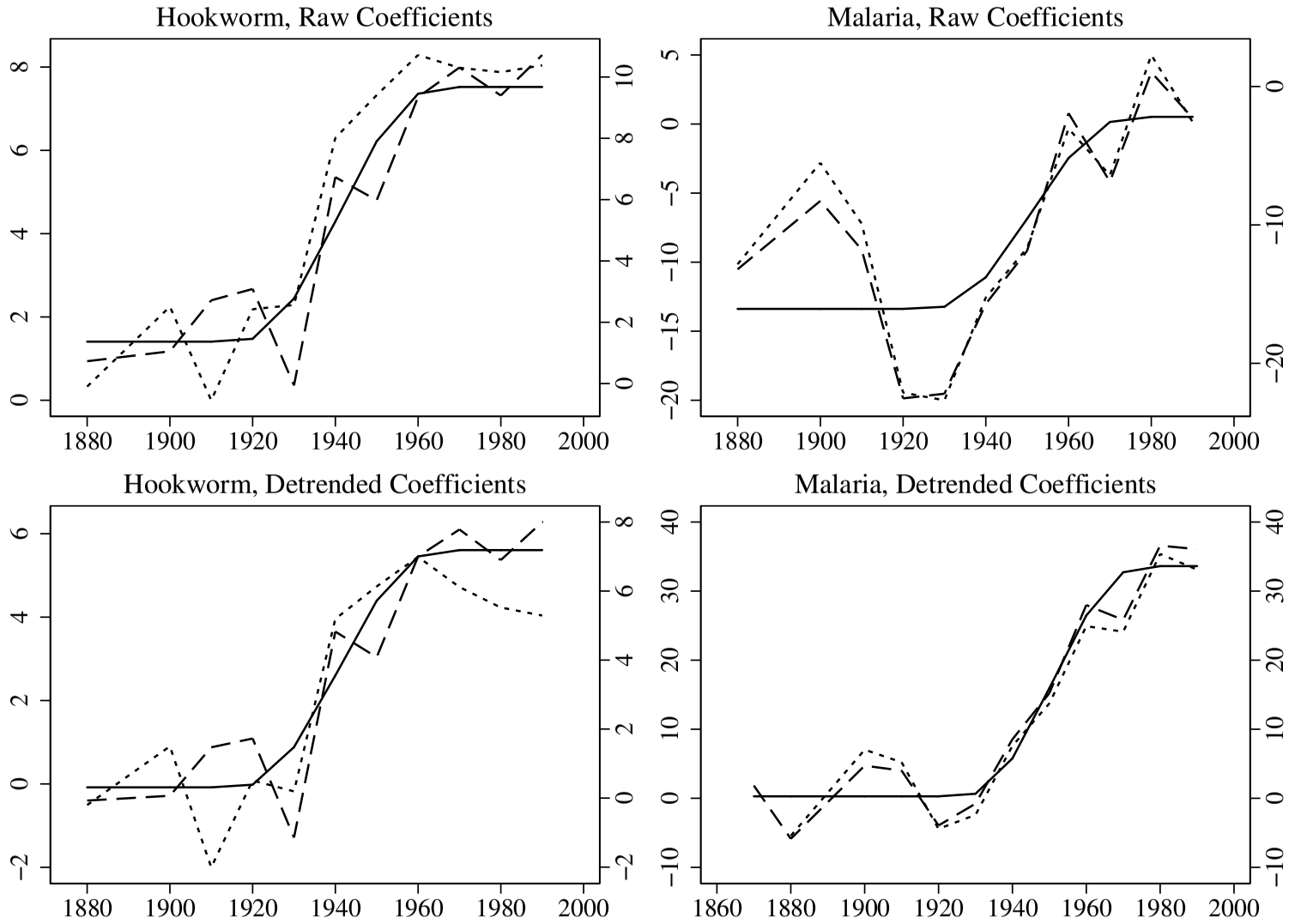
Williams, Ralph Chester (1951). *The United States Public Health Service, 1798-1950*, Bethesda, Md.: Commissioned Officers Association of the United States Public Health Service.

Figure 1: Cohort-Specific Effects and Childhood Exposure; States in the U.S.



Notes: These graphics summarize regressions of income proxies on pre-eradication hookworm-infection and malaria-mortality rates. The y axis for Panels A and B plots the estimated cohort-specific coefficients on the state-level disease measure. 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 respective eradication activities. For each year-of-birth cohort, OLS regressions coefficients are estimated on a cross section of states of birth. The state-of-birth average outcome is regressed onto malaria, Lebergott's (1964) measure of 1909 wage levels, a dummy for the Southern region, and the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variable and the disease measure. Panel C plot the average childhood (temporal) exposure to the anti-hookworm and anti-malaria campaigns among the prime working-age population (aged 25–55, inclusive) against the Census year.

Figure 2: Time-Specific Effects and Average Childhood Exposure; U.S. States

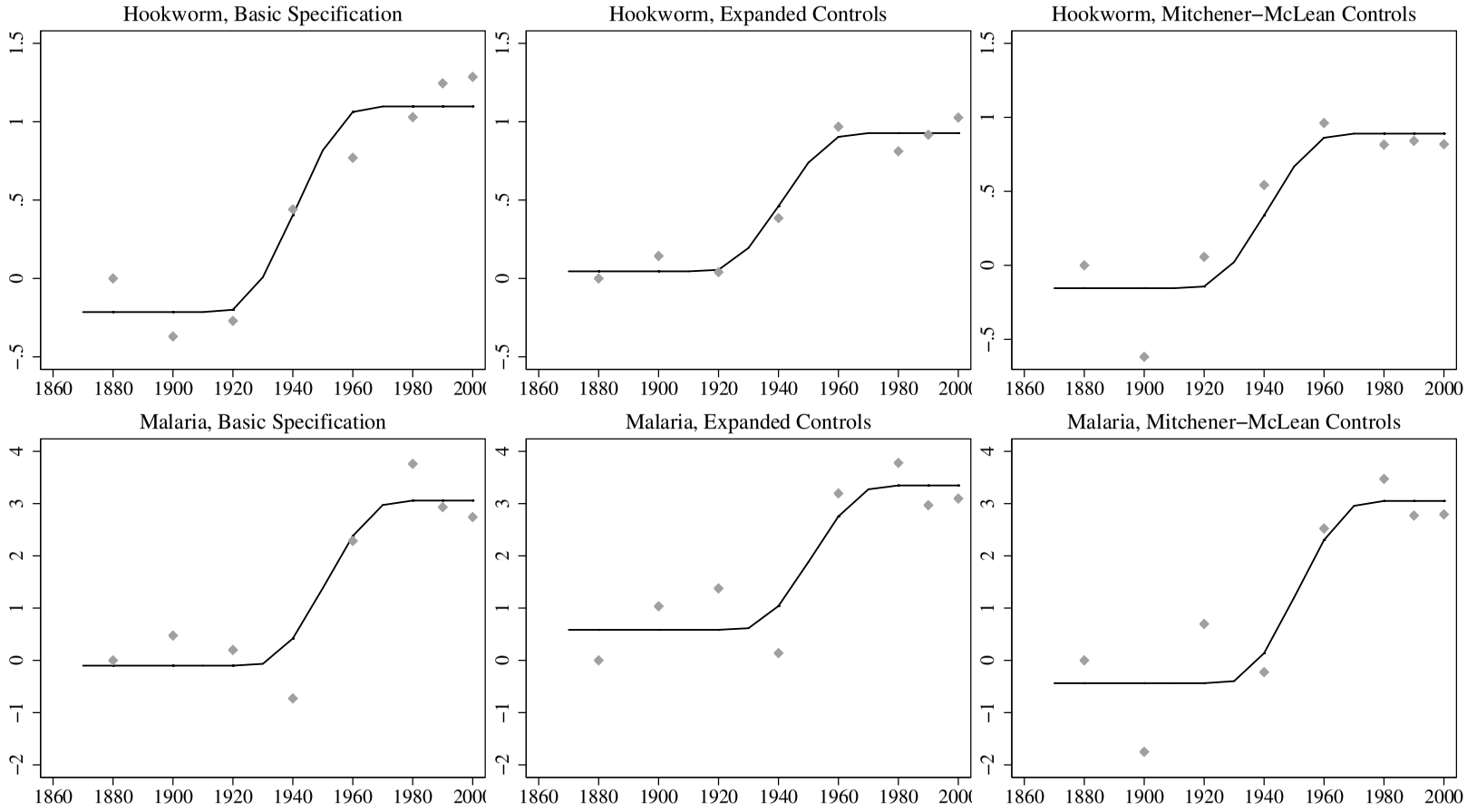


Notes: These graphics summarize regressions of income proxies on pre-eradication hookworm-infection and malaria-mortality rates. The y axis for each graphic plots the estimated year-specific coefficients on the state-level disease measure (the estimate for 1880 is normalized to zero). The x axis is census year. All underlying regressions contain fixed effects for year of birth $\times$ state of birth as well as year-specific effects of both disease measures, region dummies, and Lebergott's (1964) measure of unskilled wages in 1909. The long-dashed line, plotted against the left-hand axis, represents the point estimates from this basic specification. The short-dashed line, scaled with the right-hand axis, plots estimates from regressions that also include year-specific effects of the various control variables described in Appendix C. For the bottom row, the estimated pre-1940 trends have been removed from the series of point estimates. The solid lines in each plot are average childhood (temporal) exposures to the anti-hookworm and anti-malaria campaigns among the prime working-age population (aged 25–55, inclusive), renormalized to have the same scale as the coefficients. Appendices A and B describe, respectively, the outcome variable and the disease measures.

Figure 3: Time-Specific Effects and Average Childhood Exposure; Brazilian States

Notes: These graphics summarize regressions of log income on malaria ecology. The y axis for each graphic plots the estimated year-specific coefficients on the state-level disease measure (the estimate for 1880 is normalized to zero). The x axis is year. All underlying regressions contain fixed effects for year of birth $\times$ state of birth as well as year-specific effects of malaria, region dummies, and the log of electricity consumption *circa* 1950. The long-dashed line, plotted against the left-hand axis, represents the point estimates from this basic specification. The short-dashed line, scaled with the right-hand axis, plots estimates from regressions that also include year-specific effects of the various control variables described in Appendix C. The solid lines are average childhood (temporal) exposures to the anti-hookworm and anti-malaria campaigns among the prime working-age population (aged 25–55, inclusive), renormalized to have the same scale as the coefficients. Appendices A and B describe, respectively, the outcome variables and the disease measure.

Figure 4: Aggregate Income and Average Childhood Exposure; U.S. States



Notes: These graphics summarize regressions of income proxies on pre-eradication hookworm-infection and malaria-mortality rates. The y axis for each graphic plots the estimated year-specific coefficients on the state-level disease measure. The x axis is year. All underlying regressions contain fixed effects for state as well as year-specific effects for the following variables: both disease measures, region dummies, and Lebergott's (1964) measure of unskilled wages in 1909. The specification reported in the second column of plots alsos include year-specific effects of the various control variables described in Appendix C, and the regression for the third column includes year-specific effects of the baseline controls from Mitchener and McLean (2003). The solid lines are average childhood (temporal) exposures to the anti-hookworm and anti-malaria campaigns among the prime working-age population (aged 25–55, inclusive), renormalized to have the same scale as the coefficients. Appendices D and B describe, respectively, the outcome variable and the disease measures.

Table 1: Time-Specific Effects and Average Childhood Exposure; U.S. and Brazilian States

Independent Variables:	United States (12 Census Years)				Brazil (5 Census Years)			
Additional Controls	Basic	Expanded	Expanded	Expanded	Basic	Expanded	Expanded	Expanded
State-Specific Trends				X				X
	Panel A: One-Step Estimates from the Full Data Set							
Average Childhood Exposure to Anti-Malaria Campaign	19.825 ***	19.736 ***	24.965 ***	27.625 ***	0.046	0.084 ***	0.163	0.148
	(6.236)	(4.687)	(9.035)	(7.855)	(0.040)	(0.027)	(0.227)	(0.208)
	{0.063}	{0.063}	{0.080}	{0.088}	{0.010}	{0.019}	{0.037}	{0.033}
Average Childhood Exposure to Anti-Hookworm Campaign	5.382 *	8.980 **	8.783 *	8.640 *				
	(3.155)	(4.075)	(5.391)	(5.031)				
	{0.066}	{0.110}	{0.108}	{0.106}				
Pre-Campaign Malaria x Trend			-0.116				-0.022	
			(0.150)				(0.066)	
Pre-Campaign Hookworm x Trend			0.004					
			(0.069)					
	Panel B: Regressions on the Time Series of Malaria Coefficients							
Average Childhood Exposure to Anti-Malaria Campaign	11.669 **	10.523 **	31.681 ***		0.053 *	0.086 **	0.114 *	
	(4.112)	(4.681)	(7.220)		(0.024)	(0.017)	(0.056)	
	{0.037}	{0.034}	{0.101}	n.a.	{0.012}	{0.019}	{0.026}	
Trend x 10			-2.751 ***				-0.007	
			(0.834)				(0.115)	
	Panel C: Regressions on the Time Series of Hookworm Coefficients							
Average Childhood Exposure to Anti-Hookworm Campaign	6.382 ***	9.770 ***	8.127 ***					
	(0.730)	(0.812)	(2.066)					
	{0.078}	{0.120}	{0.100}	n.a.		n.a.		
Trend x 10			0.215					
			(0.248)					

Notes: This table reports estimates of equation 2 using OLS. The units of observation are cohorts (area and year of birth) × year for U.S. and Brazilian states, as indicated in the column headings. The dependent variable is the occupational income score. Robust standard errors (clustered on birthplace in Panel A) are displayed in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of additional terms suppressed. For the United States, 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 1870–1990 census microdata from the IPUMS and NAPP databases. For Brazil, the universe for the base sample consists of the native-born population between the ages of 25 and 55 in the 1960–2000 census microdata from the IPUMS. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity and, for the U.S., normalized by the average value of the relevant income proxy for white males born in the South between 1875 and 1895. The specification for the basic results includes the disease variables and year-specific effects of both region and a control for mean reversion. Expanded control sets, which enter into the specification interacted with trends, are described in in the text and Appendix C. Appendices A and B describe, respectively, the outcome variable and disease measures. The data for Panels B and C are the estimates displayed in Figures 2 and 3.

Table 2: Aggregate Income and Average Childhood Exposure; U.S. States

	Specification/Additional Controls:								
	Basic			Expanded			Mitchener-McLean		
Independent Variables:	*Panel A: One-Step Estimates from the Full Data Set*								
Average Childhood Exposure to Anti-Malaria Campaign	3.289 ***	5.110 **	5.046 *	3.530 ***	5.204 *	5.204 *	4.093 ***	5.414 **	5.046 *
	(1.214)	(2.499)	(2.759)	(1.344)	(2.678)	(2.876)	(1.443)	(2.490)	(2.759)
	{0.229}	{0.355}	{0.351}	{0.246}	{0.362}	{0.362}	{0.285}	{0.377}	{0.351}
Average Childhood Exposure to Anti-Malaria Campaign	1.237 ***	1.068 **	1.083 **	0.946 **	1.070 **	1.070 **	1.087 **	0.996 **	1.083 **
	(0.436)	(0.442)	(0.487)	(0.446)	(0.465)	(0.500)	(0.450)	(0.448)	(0.487)
	{0.330}	{0.285}	{0.289}	{0.253}	{0.286}	{0.286}	{0.290}	{0.266}	{0.289}
Pre-Campaign Malaria x Trend		0.002			-0.003			0.001	
		(0.009)			(0.009)			(0.010)	
Pre-Campaign Hookworm x Trend		-0.024			-0.026			-0.020	
		(0.037)			(0.037)			(0.041)	
State-Specific trends			X			X			X
	Panel B: Regressions on the Time Series of Malaria Coefficients								
Average Childhood Exposure to Anti-Malaria Campaign	3.160 ***	4.884 ***		2.763 ***	3.101		3.488 ***	3.709 **	
	(0.434)	(1.587)		(0.478)	(1.990)		(0.640)	(1.740)	
	{0.220}	{0.340}		{0.192}	{0.216}		{0.243}	{0.258}	
Trend / 10		-0.021			-0.004			-0.003	
		(0.016)			(0.022)			(0.017)	
	Panel C: Regressions on the Time Series of Hookworm Coefficients								
Average Childhood Exposure to Anti-Hookworm Campaign	1.312 ***	1.045 *		0.882 ***	0.818 ***		1.045 ***	1.180 **	
	(0.153)	(0.544)		(0.065)	(0.186)		(0.206)	(0.487)	
	{0.350}	{0.279}		{0.235}	{0.218}		{0.279}	{0.315}	
Trend / 10		0.003			0.001			-0.002	
		(0.007)			(0.002)			(0.004)	

Notes: This table reports estimates of equation 3 using OLS. The units of observation are U.S. states × year. The dependent variable is log per-capita personal income, adjusted for cross-area price differences. Robust standard errors (clustered on state in Panel A) are displayed in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of additional terms suppressed. The sample includes income data from 1880–2000, and is drawn from Mitchener and McLean (2003), except for the year-2000 data, which are drawn from the BEA web site. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity and, for the U.S., normalized by the average value of the relevant income proxy for white males born in the South between 1875 and 1895. The specification for the basic results includes the disease variables and year-specific effects of both region and a control for mean reversion. Expanded control sets, which enter into the specification interacted with trends, are described in in the text and Appendix C. Appendices D and B describe, respectively, the outcome variable and disease measures. The data for Panels B and C are the estimates displayed in Figure 4

A Construction of the Cohort-Level Data

The micro data for the RC analysis are drawn primarily from the IPUMS data for the United States and Brazil. For each country, these data are used to construct panels of income by census year, year of birth and area of birth. The cohort-level outcomes are constructed as follows.

1. The microdata are first pooled together.
2. I then define cells for each combination of year of birth, area of birth and census year. Because these averages are constructed with differing degrees of precision, I also compute the square root of the cell sizes to use as weights when estimating equation 2.

These average income by cohort/year form the dependent variable used in the present study.

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 [25,60] 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:

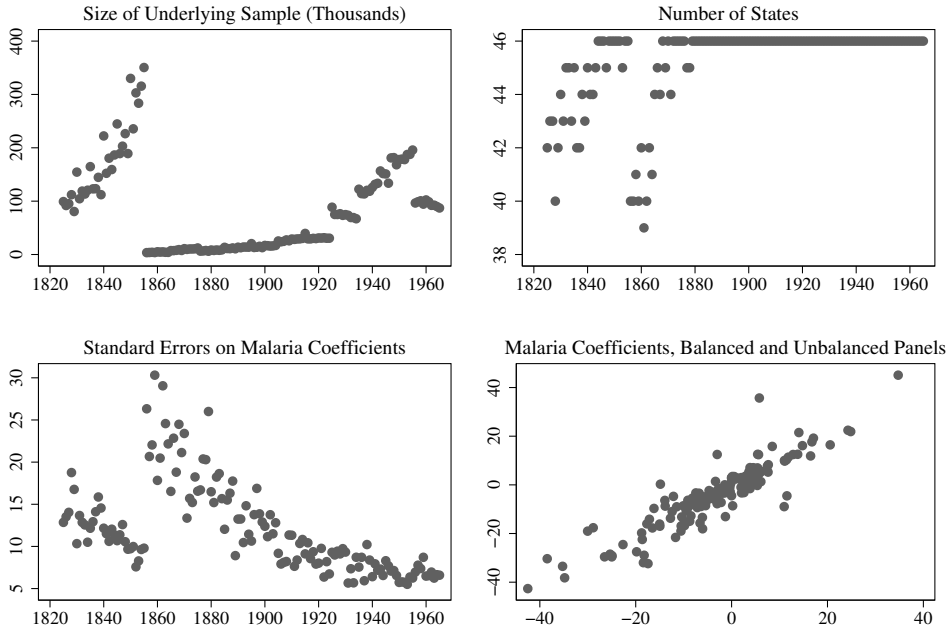
- **Occupational income score.** The occupational income score is an indicator of income by disaggregated occupational categories. It was calibrated using data from the 1950 Census, and is the average by occupation of all reported labor earnings. See Ruggles and Sobek (1997) for further details.
- **Duncan 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 51 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 full-controls specification for the occupational income score).

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 1, Panel B.

A.2 Details for the Brazilian Sample

The underlying sample used for Brazil 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–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.

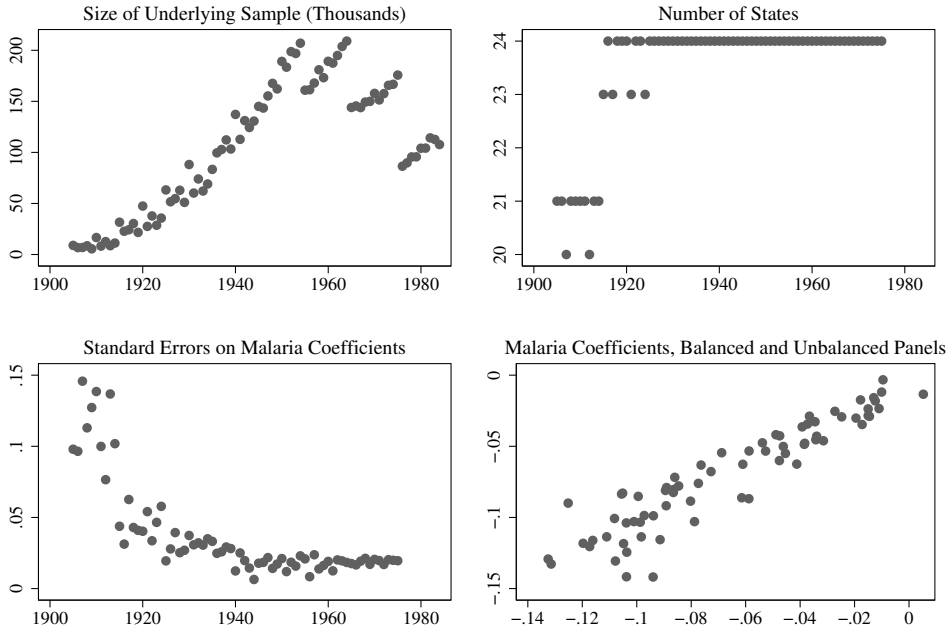
The original micro-level variables is total income, which records the total personal income from all sources in the prior month. In the empirical work above, this variable is treated in natural logs. This variable is intervalled in the 1960 census, and their midpoints are used in translating the data into income.

B Sources and Construction of the Malaria Data

Sources are indicated in parentheses at the end of each item.

- **United States, malaria.** Malaria mortality expressed a fraction of total mortality. This was measured in the 1890 Census as 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 enumerators. (Bureau of the Census, 1894.)
- **United States, hookworm.** Hookworm infection is computed from examinations of army recruits. (Kofoid and Tucker, 1921)

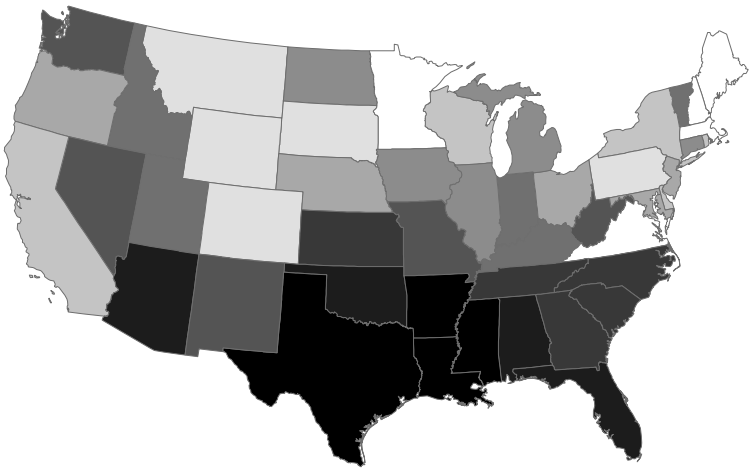
Appendix Figure A – 2: Sample Statistics for the Brazilian Sample



These graphs report additional summary statistics by year of birth for the $\hat{\beta}_t$ reported in Bleakley (2006).

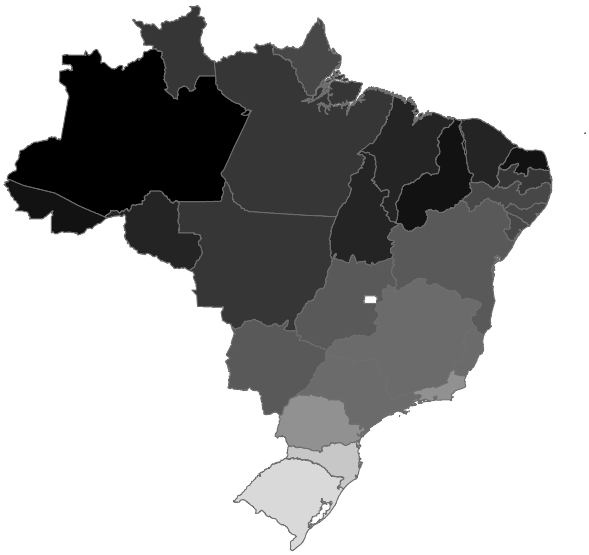
- **Brazil.** An index of malaria ecology, computed using information on climate and local vectorial capacity. The construction of these data are described in Gallup, Mellinger, and Sachs (1999a). The source data were provided as raster data in one-degree grids. A GIS program was used to extract average malaria ecology by state. (Andrew Mellinger, private communication, and author's calculations.)

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: 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 Gallup, Mellinger, and Sachs (1999a). Darker colors indicate climatic and geographic conditions more conducive to the transmission of malaria.

C Control Variables

Control variables for the United States:

- **Average wage, 1909.** I input the average monthly earnings (with board) for farm laborers by state in 1909. Various other wage measures are summarized by the same source, but are generally not available for a complete set of states. (Lebergott, 1964, Table A–24.)
- **Region of birth.** These dummy variables correspond to the Census definition of regions: Northeast, South, Midwest, and West.
- **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.)
- **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. (Census, 1894.)
- **Fertility rate, 1890.** The estimated birth rate (from Part I, page 482). (Census, 1894.)
- **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.)

D Aggregate Income Data

[TO COME]