Chronic Disease Burden and the Interaction of Education, Fertility and Growth

Hoyt Bleakley and Fabian Lange
BREAD Working Paper No. 121
May 2006

© Copyright 2006 Hoyt Bleakley and Fabian Lange

BREAD

Working Paper

Bureau for Research and Economic Analysis of Development
Chronic Disease Burden and the Interaction of Education, Fertility and Growth∗

Hoyt Bleakley†  Fabian Lange‡

April 3, 2006

Abstract

This study provides new evidence on the appropriate model of the economic and demographic transition. The episode analyzed is the eradication of hookworm disease in the American South (c. 1910). In previous work (Bleakley, 2006), it was shown that the eradication of hookworm disease led to a significant increase in school attendance and literacy. The present study shows that this increase in human capital investment was accompanied by a fertility decrease that was both economically and statistically significant. The data supports models of the demographic transition emphasizing intergenerational altruism: variables affecting childrens' economic prospects affect parental fertility decisions. The direction of intergenerational altruism can not be discerned. Parental altruism towards children and also children's altruism directed towards their parents can explain the data. The findings can not be explained by changing cultural norms, diffusion of birth control, increasing female wages, falling child mortality, or variation in female wages. We find that the eradication of hookworm had a large impact on fertility. A decline in the hookworm infection rate from 40 to 20% is associated with a decline in fertility that amounts to 40% of the entire fertility decline observed in the American South between 1910 and 1920. The relative change in fertility and schooling caused by hookworm eradication is approximately equal to aggregate comovements during the period considered.

∗We owe thanks to Eli Berman, Gary Becker, Oded Galor, Moshe Hazan, David Meltzer, Casey Mulligan, Kevin Murphy, John Strauss, Robert Topel, and seminar participants at Brown University, University of Chicago, the University of Connecticut, the Ente Luigi Enaudi (Rome), the LSE/UCL Development Seminar, the 2004 conference on Economics and Human Biology, and the 2004 Econometric Society meetings for useful comments. The authors are, of course, responsible for any errors.

†University of Chicago, GSB, bleakley @ gsb @ uchicago @ dot @ edu

‡Yale University, fabian @ dot @ lange @ at @ yale @ dot @ edu
1 Introduction

For virtually all countries, the onset of sustained per-capita income growth coincides with a rapid decline in fertility rates. This stylized fact has motivated the emergence of a diverse theoretical and empirical literature linking economic development and fertility. A significant number of researchers, most notably the participants of the European Fertility Project\(^1\) (hereafter, EFP), emphasize the spread of both new moral and cultural norms as well as birth-control technologies. Others stress the importance of factors affecting the direct costs of children, such as female wages, child mortality, and urbanization. Still others examine the link between parental investment into education and the fertility decision. These forward-looking models of fertility rely heavily on intergenerational altruism to explain the fertility transition.

The existing empirical literature on the demographic transition relies heavily on cross-country panel-data on fertility, education and other socioeconomic and cultural variables. It has proven difficult to distinguish between alternative theories of the fertility transition using such data. The observed correlations between variables measuring economic development and fertility are consistent with a variety of different models. Thus, while the empirical literature has uncovered a number of interesting patterns in the time-series, it has provided little basis for discriminating among the different theories.

We present new empirical evidence on the mechanisms driving the fertility transition. Our work differs from the vast majority of the empirical literature on the fertility transition in that it exploits quasi-experimental variation.\(^2\) We find support for models that link the fertility decision to the economic prospects of children. Our results therefore support the literature emphasizing economic incentives, and more specifically models connecting children’s economic welfare to parental fertility decisions through intergenerational altruism. We can however not distinguish the direction of intergenerational altruism: both

---

\(^1\)A comprehensive account of the European Fertility Project can be found in Coale and Watkins (1986).

\(^2\)A noteworthy exception is Paul Schultz (1985), who uses agricultural-price shocks as exogenous variation in the value of female labor in 19th-century Sweden. He shows that higher relative wages for women lead to lower fertility. In recent work using a similar design, Nancy Qian (2004) shows an effect of relative female wages on the sex mix in post-Mao China, which she argues is best understood as working through increased maternal bargaining power in the household. Another exception is due to Rosenzweig and Wolpin (1980), who consider the occurrence of a twin birth as exogenous variation in fertility.
models of child-to-parent altruism and models of parent-to-child altruism are consistent with data.

The policy intervention we examine (and describe in detail below) effectively increased the return to human capital, or, in the language of the theoretical literature, reduced the price of child “quality”. We know of no other study in the fertility-transition literature that examines similar variation.

The analysis is based on an important episode in the economic history of the Southern United States. At the beginning of the 20th century, the American South was transitioning from a high fertility, low education society to a society characterized by high levels of education and low fertility rates. We consider the impact on fertility of an improvement in health that raised the return to schooling during this period. Specifically, we examine the consequences of a particular policy intervention—the eradication of hookworm disease—in the American South during 1910–1914.3 Before the campaign, this intestinal parasite was common among children (but rare among adults) throughout the region. Hookworm infection caused anemia and listlessness, but was rarely lethal. These two facts allow us to discard effects on fertility through parental wages or child mortality.

The sudden and external origins of the eradication campaign combine with cross-area differences in pre-treatment infection rates to form our identification strategy. In spite of infecting approximately forty percent of children in the South, hookworm was not recognized as a public health problem until the turn of the century. At the beginning of the 20th century, scientific progress led to its discovery in the South. Within a short period of time, a large-scale eradication program was mounted through the efforts of the Rockefeller Sanitary Commission. This intervention originated in developments outside the region (through the innovations to knowledge and funding). The substantial regional variation in hookworm infection rates allows us to define treatment and comparison areas for the analysis.

Forward-looking models of the demographic transition predict that an increase in the returns to schooling will reduce contemporaneous fertility rates. We find robust evidence

---

3 Using a research design similar to that of the present study, Bleakley (2006) shows that hookworm eradication increased school attendance and ultimately the labor-market return to schooling.
in favor of this prediction: after hookworm eradication, fertility decreased markedly in areas that had previously suffered from high rates of hookworm infection. The abrupt onset of the hookworm eradication campaign is clearly reflected in the changing fertility rates of the “treatment” areas relative to the “control” areas during the same time-period. In contrast, we find essentially no difference in fertility trends both before the hookworm eradication campaign and after its completion.

We consider and reject several alternative hypotheses for this result. The finding is robust to controlling for a variety of additional factors, including crop mixes and production technologies, racial differences, child mortality, variation in moral norms correlated with religious denominations and pre-existing differences in fertility and human capital. We also consider the non-fertility outcomes of adults as a falsification exercise, and find no evidence that hookworm eradication was spuriously proxying for migration, income, or sectoral shocks. Furthermore, the fact that changes in schooling and fertility were abrupt and coincident with the eradication campaign is useful to our argument, especially in comparison with slowly moving variables like cultural norms.

Our results suggest that forward-looking models of the fertility transition based on intergenerational altruism can potentially explain the entire reduction in fertility observed during the demographic transition. The relative response of fertility and education to the eradication of hookworm is as large as the relative changes in fertility and education observed in the aggregate time-series. This suggests that forward-looking behavior (though not specifically hookworm eradication) can explain most if not all of the decline in fertility during the demographic transition.

We proceed as follows. Section 2 reviews the existing literature on the fertility transition. Section 3 describes hookworm disease, the historical processes that led to its eradication in the South, and the research design used in the present study. Section 4 presents the empirical results and addresses a number of alternative explanations for our findings. Section 5 shows that the quantity-quality trade-off is potentially important in explaining this region’s the fertility decline during its human-capital take-off. Section 6 concludes.
2 The Literature on the Demographic Transition

The literature on the fertility transition is large, disparate, ever evolving, and thus not easily summarized. We do however need a systematic classification of the literature in order to provide context for our empirical analysis. Our classification, which we summarize in Figure 2, centers on Gösta Carlsson’s (1966) basic distinction between Innovation and Adjustment based explanatory approaches for the fertility transition.\(^4\)

According to the Innovation hypothesis (Fig.2, node 1) cultural constraints and the absence of effective birth-control technologies prevent the effective regulation of fertility in the pre-transition stage, leading to high marital fertility rates. The introduction and diffusion of contraceptive practices and abortion and changing moral and cultural norms towards fertility control then lead to a sustained decline in fertility rates.

The Innovation hypothesis has received substantial support from the authors of the European Fertility Project (Coale and Watkins, eds. 1986) who studied the decline in fertility within each of several hundred European provinces. The onset of the fertility transition occurred within a very short period of time across most of the European continent, independent of local socioeconomic conditions. No threshold levels of important socioeconomic variables such as childhood-mortality, urbanization or education could be discerned that triggered the decline in fertility. Furthermore: “contiguous provinces that shared a cultural as well as geographical location had similar levels of nuptiality and fertility and similar patterns of decline”\(^5\). And cultural variables seem important for explaining the decline in fertility. In the words of Barbara Anderson: “...nonsocioeconomic variables, such as religion, language, ethnicity and region, explain much of the variability in marital fertility decline, even after conventional socioeconomic variables have been taken into account.”\(^6\) These findings support the view that changing cultural attitudes

---

\(^4\)Easterlin (1976) suggests that reconciling the Innovation and Adjustment approaches is necessary for explaining fluctuations in fertility across sub-populations and time. A number of contributions (recently by Brown and Guinnane (2002)) provide evidence for the Easterlin synthesis. We are sympathetic to a need to draw on both the Innovation and Adjustment approach to modeling fertility. However, we believe that it is useful to maintain clear distinctions between competing explanatory approaches in order to derive distinct predictions which can then be tested. Testing is the purpose of the empirical analysis described below and we therefore do not consider the Easterlin synthesis further.

\(^5\)Coale and Watkins (1986), page 448.

\(^6\)Coale and Watkins (1986), page 293.
towards fertility control caused the fertility transition in Europe.

This explanatory approach contrast with the Adjustment approach (Fig. 2, node 2) emphasizing changing economic incentives for child-bearing. A variety of alternative explanations within the Adjustment approach can be discerned.

A large body of work (Fig. 2, node 2.1) emphasizes variables that affect the contemporaneous costs of bearing children. Becker (1960) provides the starting point for an analysis of fertility on the basis of a straightforward consumer choice problem. This allows expressing the desired number of children in a family as a function of the price of children and parental incomes. Becker and later Jacob Mincer (1963) specifically consider female wages as important determinants of the price of children. Fertility rates decline as female wages and thus the opportunity costs of rearing children increase. Child mortality rates and the degree of urbanization are other variables that are often thought to directly affect the costs of bearing children.7

The Adjustment literature also includes a large literature emphasizing altruistic links between the generations (Fig. 2, node 2.2). If generations are linked by altruistic concerns, then the economic conditions that children face during their adulthood will become important for the fertility decision of parents. Two branches of this literature can be distinguished by the direction of altruism between the generations. The first (Figure 2, node 2.2.2) emphasizes altruism from parents towards children. Pivotal contributions to this literature are Becker, Murphy and Tamura (1990) and Galor and Weil (2000). These authors emphasize that parents have the ability to invest into the education of their children. Markets are incomplete and parents can therefore not recover the returns of investments into their children.8 This means that increasing investments into their offspring raises the price of children. This interaction between the quantity and quality of children in the budget constraint delivers the fertility transition as a response to increasing returns to investments into children.

7Doepke (2003) provides a recent review and critique of the literature linking child mortality and fertility.
8Becker and Barro (1988) emphasize the role of wages and wealth on fertility if markets are complete and can borrow against their childrens incomes. Models of fertility with complete markets are generally not prominent in the literature on the demographic transition.
Altruism from children to parents (Fig. 2, node 2.2.1) can also deliver a fertility transition. Ehrlich and Lui (1991) and more recently Boldrin and Jones (2002) hypothesize that children represent a form of old-age savings in many societies. Altruism from children towards their parents and also social norms motivates them to provide parental consumption when they are old. The need to provide for old age through children fades as countries develop forms of communal savings such as social security and financial savings mechanisms. This induces the fertility transition.

3 Background and Research Design

The explanatory approaches in Section 2 have all been constructed to fit a basic set of facts observed in the development process. This makes it difficult to discern between these approaches using time-series data. Different sets of indicators of economic development display high degrees of correlation. Societies tend to simultaneously develop financial markets, display growth in female wages, reduce child mortality, raise average education and raise adult longevity. At the same time fertility rates decline. All these variables have to be viewed as endogenous and it is not clear how to use time-series data from different countries to convincingly discern between competing explanatory approaches for the fertility transition.

It is therefore necessary to identify variation in economic variables that can plausibly be viewed as exogenous in order to examine whether the response in fertility rates and other relevant variables is consistent with the theoretical models. In the present study, we employ data from a policy intervention during the development of the American South: the eradication of hookworm disease.

3.1 Hookworm Disease

Hookworm is an intestinal parasite that lodges itself in the human intestine and absorbs nutrients from the victim’s bloodstream. The symptoms of hookworm infection (or *Ancylostoma duodenale* and *Necator americanus*) are lethargy and anemia. In rare cases, the anemia can become so severe as to cause death. The life cycle of the hookworm is dependent on unsanitary conditions. The
nematodes lay their eggs in the intestine, but the larvae are passed out of the digestive system in feces. Hookworm is therefore transmitted through skin contact with infected fecal matter. The larvae then burrow their way in through the skin. The lifespan of a hookworm is much shorter than that of a human, and so continuous reinfection is required to generate any sustained worm load.

There are two angles for managing hookworm: treatment and prevention. The treatment consists of simply taking a deworming medicine. Preventative measures include limiting skin contact with polluted soil (through the use of shoes, for example) and dealing with excrement in ways that minimize soil pollution in the first place (e.g., the use of sanitary latrines).

3.2 The Eradication Campaign

The Rockefeller Sanitary Commission for the Eradication of Hookworm Disease was formed in 1910 with the donation of one million dollars by John D. Rockefeller. Some years before, an American doctor (Charles W. Stiles) had recognized hookworm symptoms in Southerners. Through intermediaries, Dr. Stiles had convinced Rockefeller that taking on hookworm was a good foray into large-scale charity. The Commission began by conducting surveys of hookworm-infection rates among children across the region. The RSC surveyed over six hundred counties in the South and found hookworm infection to be over forty percent among children. After this finding, local opposition largely gave way to cooperation (Farmer, 1970). Another finding from the infection surveys was The RSC also noted the relation of hookworm prevalence to geography. Hookworm disease was more likely to be endemic in areas with sandy versus clay soils, or in areas that had higher temperature or average rainfall. These variables all affected the probability that hookworm larvae would survive in the soil long enough to infect a human host.

Soon after, the treatment campaign began. First, the RSC sent teams of health-care workers to counties to administer and dispense deworming treatments free of charge. RSC dispensaries visited a large and mostly contiguous fraction of the South and the campaign

---

9The historical presentation in this section draws heavily on the work of Ettling (1981).
treated over 400,000 individuals with deworming medication. Second, the RSC sought to educate doctors, teachers, and the general public on how to recognize the symptoms of hookworm disease so that fewer cases would go untreated. Another part of this publicity campaign included education about the importance of hygiene, especially with regard to the use of sanitary privies. In this period, oftentimes even public buildings such as schools and churches did not have such hygienic facilities.

Followup surveys conducted immediately afterwards as well as in the subsequent decade showed a substantial decline in hookworm infection. Although the stated goal of full eradication was not achieved, the hookworm-infection rate of the region did drop by more than half, and fewer extreme cases of the disease went unnoticed and untreated.

### 3.3 Identification Strategy

The first factor for identifying the effect of the hookworm-eradication campaign is that different areas of the South had distinct incidences of the disease. Hookworm larvae were better equipped to survive in areas with sandy soil and a warm climate. Broadly, this meant that the residents of the coastal plain of the South were much more vulnerable to infection than were those from the piedmont or mountain regions. Populations in areas with high (pre-existing) infection rates were in a position to benefit from the newly available treatments, whereas areas with low prevalence were not. This heterogeneity allows for a treatment-control strategy.

Second, the initiation of the campaign by the RSC was largely a function of factors external to the Southern states. The eradication campaign was made possible by critical innovations to knowledge: understanding how the disease worked and more importantly recognizing its presence. This contrasts with explanations that might have troublesome endogeneity problems, such as changes in government spending or positive income shocks in the infected areas. But even with the knowledge of the hookworm problem, there would have been formidable obstacles to taking action. The public-health infrastructure of this period was extremely limited. Rockefeller’s donation was an important precondition for attacking the problem.

Thirdly, the anti-hookworm campaign achieved considerable progress against the dis-
ease in less than a decade. This is a sudden change on historical time scales. Moreover, we examine outcomes over a fifty-year time span, which is unquestionably long relative to the five-year RSC intervention.

These factors combine to form the central variable in the present study:

\[(\text{Pre-treatment Infection Rate })_j \times (\text{Indicator for Post-Treatment })_t.\]

More compactly, call this variable \(Z_{jt} = H_{jpre} \times Post_t\), where \(j\) indexes the geographic area and \(t\) indicates the year. The variable \(H_{jpre}\) denotes the level of hookworm infection among school-aged children in area \(j\) at the time of the RSC’s initial survey, and \(Post_t\) is a dummy variable indicating whether year \(t\) is later than the active years of the RSC campaign (1910–1915).

We compare the evolution of outcomes (e.g., number of own children less five years old) across counties with distinct hookworm-infection rates, to order to assess the contribution of the eradication campaign to the observed changes. Estimating equation (1) measures the reduced-form differences by pre-eradication hookworm for some outcome \(Y_{ijt}\) for person \(i\) in area \(j\) at time \(t\).

\[Y_{ijt} = \beta Z_{jt} + \delta_t + \delta_j + X_{ijt}\Gamma + \varepsilon_{ijt}\]

in which the \(\delta_t\) are time dummies, the \(\delta_j\) are geographic fixed effects, and \(X_{ijt}\) is some vector of individual-level controls.\(^{11}\)

\(^{10}\)All of the estimates of this equation below are calculated using ordinary least squares (OLS) regressions.

\(^{11}\)The model is derived as follows. For individual \(i\), in area \(j\), in year \(t\), we start with an individual-level model with infection data and linear effects of hookworm:

\[Y_{ijt} = \alpha H_{ijt} + \delta_j + \delta_t + X_{ijt}\Gamma + \varepsilon_{ijt}\]

where \(H_{ijt}\) is a variable measuring the probability that \(i\)'s offspring will be infected with hookworm. Individual infection data at this level of detail are not available, so the hookworm infection rate \(H_{ijt}\) is replaced with its ecological (i.e., aggregate) counterpart:

\[Y_{ijt} = \hat{\alpha} H_{jt} + \delta_j + \delta_t + X_{ijt}\Gamma + \varepsilon'_{ijt}\]

(This equation can equally be run in aggregate form entirely, and, when estimated, it gives very similar results to those in the present manuscript.) For the instrument \(Z_{jt}\), the reduced form of this system is equation 1.
The treatment and control groups in this study are defined by individuals living in areas with high and low pre-period infection rates. This identification strategy relies on the maintained assumption that areas with higher infection rates in the pre-period also experienced larger declines following the eradication campaign. The data support for this assumption. Resurveys found a decrease in hookworm infection of thirty percentage points across the infected areas of the South (Jacocks, 1924). Such a dramatic drop in the region’s average infection rate, barring a drastic reversal in the pattern of hookworm 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. Figure 5 presents data on this issue.\(^{12}\) The basic assumption of this section — that areas where hookworm was highly endemic saw a greater drop in infection than areas with low infection rates — is borne out across states and across counties for which we have post-intervention data.

We argue that the results estimated in this framework favor a forward-looking, altruism-based model of fertility.\(^{13}\) Two additional factors allow us to interpret the estimates in this manner:

1. In the American South hookworm disease affected almost exclusively children, and not the adult population. This assumption is supported by the infection surveys of the period (see Smillie and Augustine (1926), and Figure 4).\(^ {14}\) This means that we can rule out direct effects of hookworm eradication on adults, and therefore analyze the fertility decision as a function of childhood factors.

2. Mortality from hookworm infection was extremely rare in the American South, and

\(^{12}\)This figure embodies the first-stage relationship. Consider the aggregate first-stage equation:

\[ H_{jt} = \gamma Z_{jt} + \delta_j + \delta_t + \eta_{jt} \]

This equation can be written in first-differenced form and evaluated in the post-RSC period:

\[ \Delta H^{\text{post}}_{jt} = \gamma H^{\text{pre}}_{jt} + \text{constant} + \nu_{jt}, \]

an equation that relates the observable variables graphed in Figure 5.

\(^{13}\)The data do not allow us to distinguish the direction of intergenerational altruism. Both altruism from parents to children and altruism from children to parents can explain the data.

\(^{14}\)This is not universally true. Surveys of other countries often find higher infection rates among adults, which may be due to differences in immunological or behavioral factors. This “peak shift” is analyzed further by Woolhouse (1998).
so we are able to consider the effects on fertility through morbidity and its effect on the return to schooling.

Below we estimate equation 1 and take care to consider plausible competing explanations for the observed patterns and provide evidence that allows us to reject each with a fair degree of confidence. Overall, we believe that the data provides substantial evidence in favor of altruism-based theories of the fertility transition.

4 Empirical Results

4.1 Data and Descriptive Statistics

This study links county-level data on hookworm infection with individual-level data on schooling and fertility. The aggregated data show a region with on average high levels of infection in 1910. The hookworm disease burden is however unevenly distributed across the region. In 1910 fertility rates throughout the South are high, but declining and school enrollment rates are increasing. Both these trends continue over the period 1910-1920.

These patterns of the data are evidenced by the summary statistics of various aggregate outcomes presented in Panel A of Table I. Since county boundaries change during the period we consider we use aggregated county groupings, the so-called “State Economic Areas” (SEAs) as the geographic unit, and so the $j$ above indexes SEAs. The hookworm-infection rates were computed by the Rockefeller Sanitary Commission for 550+ counties across the South. The RSC collected these data as a prelude to mounting a widespread treatment campaign. The data collection took place between 1910 and 1914 (at a single point in time for each county), and the summary statistics were constructed from samples of school-aged children in each county. The RSC surveys measured an average infection rate across SEA’s of 32.8%.

The RSC also reported county-specific details of their subsequent treatment campaign. For example, in this study we include data on the number of treatments issued by the RSC, as well as the number of individuals treated by the Commission’s staff. These numbers (scaled by 1910 SEA population) are also reported in Panel A. The second and
third columns display the means by subsamples that are separated based on the severity of their hookworm problem. Not surprisingly, the RSC directed more resources towards the areas with greater hookworm infection.

The micro-level data employed in the present study come from the Integrated Public Use Micro Sample (IPUMS), the output of a project to harmonize the coding of historical U.S. Census microdata (Ruggles and Sobek (1997)). The sample consists of native-born whites and blacks in the age range [8,16] in the case of school age children and in the age range [15,49] for female adults in the study of fertility. The age criteria for children serves to select children of school age who are likely not yet old enough to have migrated on their own. The generalized fertility rate reported in Table I, panel A is a measure of fertility only available for the aggregate level data. It is calculated as the ratio of children less than 5 year old to females aged 15-49.

Key individual-level variables are summarized in Panel B of Table I. The key outcomes are schooling attendance and fertility. The fertility measure for the individual level is the number of own children less than 5 years old in the household. The first group of columns show the results from the 1910 Census only, whereas the rightmost two columns contain summary statistics from the 1900-1950. The school enrollment variable refers to enrollment within a certain time span, typically several months, prior to the enumeration date.\textsuperscript{15}

The summary statistics already displays the features of the data that will generate the empirical results of the regression analysis discussed below. We take the evidence displayed in figure 1 as indicative that the American South was at the time transitioning from a high fertility, low schooling society to a low fertility, high schooling society. Under the assumption that high infection rates reduce the return to investments into children we expect the following patterns in schooling and fertility in our sample:

1. Schooling Enrollment Rates are lower and fertility rates are higher in SEAs with high infection rates during the pre-period;

2. Schooling Enrollment Rates increase and Fertility Rates decline during the sample period.

\textsuperscript{15}Table I does not present summary statistics for all variables employed in this study.
3. Following the eradication campaign, schooling enrollment rates increase faster in areas with high pre-period infection rates;

4. Following the eradication campaign, fertility rates decline faster in areas with high pre-period infection rates.

All of these predictions are in fact confirmed by the summary statistics reported in Table I. The low school enrollment rates in areas with high hookworm infection rates and the rapid increase in school enrollment in these same areas as hookworm was eradicated establish the validity of the assumption that high rates of hookworm infection in fact represented an impediment to investing into children, and that the eradication of hookworm disease raised the returns to education. In other words, the increase in schooling enrollment is consistent with our interpretation that the eradication of hookworm disease reduces the price of children’s quality. This allows us to test forward looking theories of the demographic and economic transitions that link investments into child ‘quality’ and fertility decisions through intergenerational altruism. These theories predict that the decline in the price of child ‘quality’ which raised investments into children should have simultaneously lowered fertility rates. It is exactly this prediction that we examine in much detail in this study.

4.2 Regression Estimates

4.2.1 Human Capital

In this subsection, we estimate the effect of the hookworm eradication on schooling enrollment using equation (1) and the 1910 and 1920 census. As described above, the variable of interest, $Z_{jt}$ is the interaction of pre-period hookworm infection, $H_j^{pre}$, with a dummy, $Post_t$, indicating whether the year comes after the RSC. Controls are as specified in the table. The results reported here are identical to those presented by Bleakley (2006), who
describes in greater detail the effect of hookworm eradication on schooling and literacy.\textsuperscript{16} In the present study, we concentrate on the fertility results and therefore only briefly consider the schooling as an outcome.

We find that living in areas with high levels of hookworm infection in 1910 is associated with a substantial additional increase in human-capital investments during 1910-1920. These empirical results are presented in Table II, Panel 1. We show estimates of the coefficient on the main variable of interest, $Z_{jt} = \text{pre-period hookworm} \times \text{post}$, for three different binary dependent variables measuring human capital investments: school enrollment, regular school attendance, and finally literacy.\textsuperscript{17} Our finding is the same for all three of these measures is the same: areas with higher levels of hookworm saw greater increases in human capital following the anti-hookworm intervention.

The estimates presented above imply plausible quantitative effects of hookworm infection on schooling enrollment. Specifically, the coefficient on $Z_{jt}$ in Table II.1, column 1 implies that one standard deviation in hookworm infection rates in 1910 is associated with an additional increase of school enrollment of 2 percentage points between 1910 and 1920. This compares with an average increase in school enrollment over this time-period of 9 percentage points. We can also compare the reduced-form effect of $Z_{jt}$ (about 0.1) to the estimated decline in infection as a function of the same variable (0.44).\textsuperscript{18} Dividing the first number by the second gives us the Indirect Least Squares estimate of infection on schooling: 0.23. This indicates that an child infected with hookworm is 23\% less likely to be attending school.

The human capital investment findings are significant for our study since they allow us to determine the sign of the effect of hookworm eradication on the returns to education.\textsuperscript{19}

\textsuperscript{16}Bleakley (2006) also follows up on the labor-market experience of cohorts exposed to the campaign as children and find that the returns to schooling rose. We refer the reader with a particular interest in the human-capital results to that study.

\textsuperscript{17}We include dummy variables for age $\times$ race $\times$ census region $\times$ census year. This represents an extremely flexible functional form for the demographic effects. We have considered more parsimonious specifications to control for these variables and our main result is robust to these changes.

\textsuperscript{18}The latter number comes from the follow-up evidence shown in Figure 6.

\textsuperscript{19}It makes no particular difference for our analysis whether we interpret the disease burden as affecting the effort cost of learning or the labor-market return to human capital.
during the 1910-1920 period in areas with high pre-period infection rates indicate a positive effect of hookworm eradication on the returns to schooling. We will now turn to the analysis of the impact of hookworm eradication on fertility choices.

4.2.2 Fertility

Table II, Panel 2 reports the results of a regression analysis of fertility similar to the one reported for schooling in the previous subsection. The results are similar across 5 different measures of fertility\(^20\) and we limit our discussion to the results using the number of children younger than 5.\(^{21}\) Fertility declines substantially faster in SEAs with high rates of hookworm infection in 1910. The estimates imply that an additional standard deviation in 1910 hookworm infection rates is associated with an additional decline in the number of children younger than 5 years old per fertile woman of about 0.025 per women between 1910 and 1920. The aggregate decline in this fertility measure during this time-period amounts to about 0.05. It is possible to roughly translate these numbers into total fertility rates (TFR), a more familiar measure of fertility. The TFR in the entire sample declined from about 4 to about 3.65 between 1910 and 1920, a decline of about 0.35. This compares with a predicted decline in the TFR of about 0.5 for an SEA with an additional standard deviation in hookworm infection rates between 1910 and 1920. The direct effects of the hookworm eradication campaign on fertility are therefore substantial.

4.2.3 Timing

We next exploit the fact that we have data available not only for 1910 and 1920, but for each decennial census between 1900 and 1950. This allows us to clearly document that the excess decline in fertility and increase in school enrollment, attendance, and literacy associated with hookworm disease is indeed concentrated in the time-period 1910-1920. For this purpose we estimate pooled regressions on the 1900-1950 data where we interact

\(^{20}\)The number of children less than 5, an indicator variable for whether the woman has a child less than 5 years old, an indicator variable for whether the woman has a child less than 1 year old, the total number of children, and an indicator variable for whether the woman has any children.

\(^{21}\)Not reported are also the results from using the intensity of treatment measures provided by the RSC. These results are very similar to those reported here and are available from the authors upon request.
the hookworm infection rate in 1910 with year dummies for all years.

Figure 6 displays the coefficients on the hookworm infection rate of a regression interacting the hookworm infection rate with year effects for the census year between 1900-1950 for both schooling and fertility regressions. To present comparisons with the base year we omit the SEA fixed effects and can thus estimate the interaction between the infection rate and year dummies for all census years between 1900 and 1950. Figure 6, panel 1 shows that the decline in the relation between hookworm and fertility was indeed concentrated in the period 1910-1920. Figure 6, panel 2 demonstrates the analogous finding for human capital investments. Clearly the period of eradication is associated with a major change in the relation between school enrollment and fertility and the 1910 local hookworm infection rate. Significant in this context is that in the period preceding the eradication campaign the relation between fertility and infection rates does not decline, but instead increases (insignificantly) in size. The excess decline in fertility rates observed for those SEAs with high infection rates is therefore not the continuation of a pre-existing trend in the data. We can furthermore demonstrate this point by augmenting this specification with both SEA fixed effects and SEA specific trends for the years 1900-1950 to capture any potential pre-existing trends in fertility varying by SEA during this time-period. The interaction between the infection rate and the year dummies then capture (since we omit the 1910 interaction with infection rates) the change in fertility observed in our treatment (high infection) areas as compared with the control (low infection) areas. For this specification we find again an excess decline of fertility in areas with high hookworm infestation rates between 1910 and 1920. The coefficient estimate for the interaction between 1920 and the infection rate in this specification is 0.1142 with a standard error of 0.0432. This is significant at the 1% level. The coefficient estimates in this specification are indeed very similar to the ones shown in figure 6-2. Again we find that the decline in fertility associated with the hookworm infection rates is observed only between 1910 and 1920, whereas between 1900 and 1910 as well as between 1920 and 1950 now significant changes in fertility specific are observed in the treatment areas compared to the control group.
4.3 Sensitivity Analysis

In this subsection, we evaluate the robustness of our finding to controlling for a variety of alternative explanations. The results from this sensitivity analysis on the 1910 and 1920 sample are provided in Table 3. Each panel shows first the coefficient on the main variable of interest (hookworm interacted with post) for all 5 fertility measures. In Panel A we reproduce the results from Table 2-2 to facilitate comparison. Each subsequent panel also displays the coefficients on the interactions of different control sets interacted with the post variable, as well as the F-statistic related to these additional controls. These regressions thus allow for a variety of alternative explanations of why areas with high hookworm infection rates might have experienced disproportional declines in fertility during the period considered here.

Panel B then considers the possibility that health and sanitary conditions in 1910 might have correlated with hookworm disease infection rates and might have had an independent effect on subsequent fertility. We thus control for measures of local health and sanitary conditions collected both by the RSC and from other sources. These variables are described in more detail in the data appendix. Of particular interest are a measure child mortality which we were able to construct using the 1900 and 1910 census questions on the number of children ever born and the number of surviving children. (Given the importance some authors have placed on child mortality for explaining the fertility transition, we will discuss child mortality again below.) We also control for fertility in the pre-period directly, as well as for pre-existing fertility trends. Because the Federal government cleaned up unsanitary conditions in and around camps, we control for World War I camp population per resident population. We also include malaria mortality due to the importance this disease played in the development of the American South.

The F-statistic on these health controls clearly indicates that these controls help explain the relative fertility patterns across geographic areas during this period. But, for us the most important finding is that the coefficient on the main variables of interest is very robust to the inclusion of these variables. This finding will be similar in all the subsequent panels of Table 3: various variables tend to explain parts of the geographic variation in the
fertility decline, but the coefficient on the interaction between 1910 hookworm infection
rates with the post-period indicator is impervious to the inclusion of these controls.

Panel C controls for a variety of changes in education systems over this time-period
and also for the variation racial composition and racial conflict across locations during this
time. The education controls capture changes in the inputs (pupil-teacher ratios, school
density, value of school buildings and equipment, teacher salaries, Rosenwald classrooms,
county spending on schools), variation in the returns to literacy across areas, and changes
in literacy rates among adults and among children. The racial variables are the fraction
of individuals black in 1910 and also the number of lynchings per capita recorded during
this time. Without discussing the coefficients on each of these controls individually we
again see that the coefficient estimate on the main variable of interest is very similar to
those reported in Panel A.

Panel D is motivated by the link between soil and climate conditions and hookworm
infection. The worm favours sandy soil and warm, humid climate. Clearly it is therefore
possible that there is a spurious correlation between hookworm infection rates and the
importance and composition of agriculture during this time period. We therefore collected
data on both the importance of agriculture in the local economy (fraction of population
living in urban centers), the value and amount of land available per capita and also the
sharecropping acreage, and finally about the fraction of land used to grow cotton and
tobacco respectively. These two crops represent the major cash crops grown in the Amer-
can South during this time. Again we find that all of these variables jointly contribute
to explain fertility trends, but they do not affect the coefficient estimates on hookworm
infection in any substantial manner.

Panel E then directly addresses the Innovation Hypothesis (see Section 2) promoted
by the Princeton Project. Here we control for the share of the population that was either
Baptist, Methodist, Catholic, or was non-churchgoing during this time. (The total of all
other religions is the omitted category.) And again we find a role for these variables,
especially for the share of Catholics. It seems as if fertility rates declined less rapidly
among Catholics then among other denominations between 1910 and 1920. However,
these variables are not individual significant and therefore strong conclusions should not
be drawn based on the results reported in Panel E. What is striking however is again how robust our main finding is.

Panel F then presents the results for controlling for all of these controls simultaneously - and again we claim a 'success' for the empirical finding that hookworm eradication lead to an excess decline in fertility rates. The main result from our sensitivity analysis therefore is the striking robustness of the coefficient estimate on our main variable of interest. This certainly strengthens our confidence that hookworm eradication is indeed related to an excess decline in fertility rates among those SEAs with high pre-period hookworm infection rates.

4.3.1 Child mortality

The previous paragraphs argued that our empirical finding is indeed robust to the inclusion of a variety of different controls addressing various potential sources of spurious correlation. Next we consider in more detail the role of one particular variable, child mortality. Some authors (Sah (1991), Kalemli-Ozcan (2002)) have directly attributed the fertility transition observed during the development process to declines in child mortality that are usually observed concurrently. In contrast, the episode we consider relies on the eradication of a disease affecting the return to education. However, even though this disease is only fatal in very rare cases one could argue that the public health measures undertaken to combat hookworm disease might potentially have reduced child mortality. It is therefore natural to wonder whether the empirical findings reported here might be spurious and simply reflect a reduction in child mortality associated with the eradication of hookworm disease.

The theoretical predictions for the fertility response to a decline in child mortality and fertility depends on whether fertility is measured as the number of surviving children or as the number of births. The fertility measure used in this study, the number of own children aged 5 and younger in the household does not correspond to either concept, but instead represents a mixture of both. It excludes those children that have died before the census was taken, but includes some children that will still perish before reaching adulthood. Nevertheless, since child mortality is highest during the first months and drops
rapidly afterwards, most children counted here will survive to adulthood. This leads us to interpret the number of children less than 5 years old as a measure of surviving children.

Consider then the possibility that the eradication of hookworm disease indeed led to a drop in child mortality in areas where hookworm disease was endemic in 1910. The question is whether there is a theoretically plausible reason why this decline in child mortality might have led to a decline in the number in surviving children? To abstract from quantity-quality considerations assume that parents simply treat the number of surviving children as an argument in their utility function without introducing any quality dimension into parental considerations. With a fixed child mortality rate $m$, and a price per birth $p$, the cost of a surviving child is $p/(1 - m)$. Thus a decrease in child mortality $m$ reduces the effective price of surviving children and therefore should lead to an increase in the number of surviving children chosen by parents. This simple application of the law of demand predicts that any beneficial effect of the campaign to eradicate hookworm disease on child mortality rates will have led to an increase in the number of surviving children in places with high infection rates. The bias introduced by child mortality and implied by this model of fertility therefore works against finding an excess decline in fertility in areas with endemic hookworm disease as documented in Table II-2. Thus, thinking about child mortality within this model would indeed strengthen our empirical argument.

An alternative model of the link between child mortality and fertility is offered by Sah (1991) and Kalemli-Ozcan (2002). They model the demand for children in the presence of uncertainty. They argue that hoarding can cause the relation between child mortality and fertility to be positive. An introduction of the dynamic nature of fertility decisions will reduce the force of this argument since any initial losses can be partially offset by having more children in later periods. In a recent study Doepke (2003) concludes that parents have to be both extremely risk averse and unable to compensate for the death of older children by having additional children to generate a positive relation between net fertility and child mortality. This greatly weakens the argument that declining child mortality might have caused the fertility transition.

There are also empirical reasons that lead us to believe that the reduction in fertility in the treatment group is not due to a reduction in child mortality. In 1900 and 1910
the census asked questions about both about the number of born and surviving children. We use these to construct an individual child mortality rate for each women. In Table III, Panel B we already showed that controlling for the effect of child mortality does not affect our results in any significant manner. We do observe that in areas with high pre-period child mortality rates the decline in fertility rates over the 1910-20 period is smaller. We can not use this as evidence on the theories linking the fertility transition with child mortality rates without knowledge about the correlation in the change of child mortality rates between 1910 and 1920 with the pre-period mortality rates. But, we can rule out in this manner that hookworm disease itself correlated with high child mortality rates and that the eradication effect worked through eliminating this correlation. Table IV shows additional evidence to this effect. There is no statistically nor economically significant relation between child mortality and local hookworm infection rates in the pre-period.22

4.3.2 Variation in response to hookworm by maternal age

A further concern that requires attention is that the hookworm eradication campaign might have affected some young women directly. If school enrollment competes with child rearing, then it is possible that hookworm eradication might affect the fertility behavior of young women directly by inducing them to stay in school longer. In addition school enrollment will improve labor market opportunities following graduation and thus might affect the price of children by increasing the opportunity costs of females. If this was indeed the mechanism that lead to a decline in fertility rates, then our results can not be interpreted as supporting a forward-looking, altruism based model of the fertility transition.

We can examine this concern directly. We estimate how the pre-period infection rate affects fertility for the entire age-profile of women in the sample in 1910 and 1920. The hookworm eradication campaign will only have affected school enrollment of women aged 15 or younger in 1910. Thus if we observe that women aged 25 or older in the treatment

---

22To consider the relation between child mortality and hookworm infection rates after the eradication campaign, we also look at infant mortality from the late 1930s. Similarly, there is no statistically or economically significant relationship between hookworm infection, circa 1910, and mortality. We have not conducted a test based on differencing these data because they are not strictly comparable.
group (high infection rates) reduced their fertility relative to the control group, then this can not be explained by more additional schooling received by women in this age-group. Figure 7 shows the interaction of 1910 infection rate with the post-period indicator for the entire age-profile 15-49. Clearly the additional drop in fertility in areas with high 1910 infection rates following the eradication campaign is not limited to age-groups which are affected directly by the campaign.

There is no evidence that the decline in fertility rates might be due to the different opportunities available to young women. Instead, the excess decline in fertility occurs at all ages. Indeed, it is largest for women between the ages of 25 and 40, when fertility itself is largest.

### 4.3.3 Labor-market outcomes of adults

We now consider how hookworm eradication affected labor-market outcomes of adults. These should, given our knowledge of the disease, not have been affected by hookworm or its eradication directly. Therefore, examining outcomes related to adults can therefore serve as a falsification exercise. Furthermore this provides further evidence that hookworm eradication indeed presents an empirical test of forward-looking models of the fertility transition - that is of models that emphasize the effect of change in childrens welfare on parental fertility decisions against the alternative that the fertility transition is caused by a change in variables (e.g. female wages) affecting adults directly. The goal is to search for evidence of shocks to income or sectoral demand that affected fertility and were spuriously correlated with the hookworm measure. These results are found in Table V, where we see little evidence that hookworm eradication affected the labor market opportunities of adults during this time-period directly. We examine the effect of hookworm eradication on labor-force participation, the choice of occupation (proxied by the occupational income score, the Duncan socioeconomic index, the probability of working in agriculture, and the probability of working as a craftsman or operative), on the probability of living on a farm, and finally the probability of having been born in a different state. The latter outcome measure can serve as a composite index indicating changes in the attractiveness of particular locales for adults. It also helps rule out directly any contamination of the
data due to selective migration. We examine all of these outcomes for males and females of
different ages. In no instance do we find evidence of an impact of hookworm eradication
on adult outcomes if we pool the data within gender across ages. If we look at each
age group separately, only in the case of men aged 50-55 do we observe a significant
impact of hookworm eradication. It is not surprising that we should have found one
significant coefficient estimate given that we now have 6 demographic (gender × age
group) subsamples and 7 outcome variables. In summary: there is no compelling evidence
of a direct effect of hookworm eradication on adult outcomes.

5 The Quantitative Importance of Forward-Looking
Models of the Fertility Transition

In the previous sections we argue that the data on fertility and schooling in the American
South following the hookworm eradication campaign is consistent with an interpretation
of the fertility transition that is based on intergenerational altruism. The reduction in the
price of quality caused by the eradication of hookworm disease lead to both an increase
in human capital investments and a decline in fertility rates.

We will now attempt to use the data on hookworm eradication to quantify the im-
portance of linking fertility decision and the schooling decision in explaining the demo-
graphic transition. The hookworm eradication campaign provides us with a response in
both schooling and fertility to the eradication of hookworm. It is natural to compare this
response with the overall relative changes in schooling and fertility observed during the
development of the American South.

5.1 Method

This paper argues that the eradication of hookworm disease affected the opportunities
to invest into human capital and, through the quantity-quality trade-off, the costs of
children. Assume that the conditions affecting the human capital investment conditions
are summarized by a parameter $r$ which depends on the hookworm infection rate $\rho$ and
other factors denoted \( \theta_r \) and can thus be written as \( r(\rho, \theta_r) \). Then let fertility \((n)\) and human capital \((h)\) depend on this parameter \( r \) and other effects \( \theta_n \) and \( \theta_h \):

\[
n(r(\rho, \theta_r), \theta_n), \ h(r(\rho, \theta_r), \theta_h)
\]

The question we are attempting to answer here is how much of the decline in fertility during the demographic transition is attributable to linkages between the quantity and quality of children. To answer this question we compare the relative changes in fertility and education attributable to the hookworm eradication campaign with the relative aggregate changes in fertility and education observed during the 1910-1920 period only. Between 1910 and 1920 the number of children younger than 5 per fertile woman declined by 0.075 (from a base of 0.55 in 1910). The school enrollment rate for ages 8-16 during this time period rose by 0.085. Thus we observe a decline in fertility per percentage-point increase in enrollment of 0.88=0.075/0.085. How does this decline in fertility relative to the increase in education compare with the changes in fertility and education observed in areas with high infection rates compared to those with low infection rates?

Ideally, we would like to estimate the change in fertility caused by some variation in the hookworm-infection rate and compare this with an estimate of the change in education caused by the same variation in the hookworm infection rate. This would give us an estimate of the importance of the quantity-quality trade-off if the hookworm infection rate affects both fertility and human capital investments primarily through the opportunities of investment into human capital. This comparison would then answer the question: If we change the investment opportunities into children’s human capital such that the school enrollment rate increases by 1 percentage point, then by how much does fertility decline? The answer to this question could then be compared to the 0.88 observed for the time-period 1910-1920 for the entire sample.

Unfortunately the structure of the data prevents us from pursuing this simple strategy. A simple regression of fertility and human capital investments is fraught with omitted variable biases. This is in fact the reason why we use the double difference estimator described above. We only observe the infection rate in the pre-period and are therefore
not able to estimate the effect of the eradication campaign on the hookworm infection rate directly. This makes it impossible to calculate the partial derivatives $\frac{\partial n}{\partial \rho}$ and $\frac{\partial h}{\partial \rho}$ directly and thus we can’t obtain the ratio $\frac{\partial n}{\partial \rho} / \frac{\partial h}{\partial \rho}$ with this strategy.

To make progress we make additional linearity assumptions and an assumption on the relation between the pre-period infection rate and the subsequent decline in infection rates following the eradication campaign. In particular assume that $r(\rho, \theta_r) = \gamma_1 \rho + \gamma_2 \theta_r + \varepsilon$, that $n(r, \theta_n) = \beta_1 r + \beta_2 \theta_n + u$, that $h(r, \theta_h) = \alpha_1 r + \alpha_2 \theta_h + u$ and finally that the $\rho_{1920} = k \rho_{1910}$ where the subscript denotes the year. Assume furthermore that the omitted variables do not change systematically over time (we need that the omitted variable bias remains the same). Then we can write the relation between fertility in 1910 and the hookworm infection rate as:

$$n_{1910} = \beta_1 \gamma_1 \rho_{1910} + \beta_1 \gamma_2 \theta_r + \beta_2 \theta_n + u_{1910} + \beta_1 \varepsilon_{1910}$$

The plim of the coefficient in a regression of $n_{1910}$ on $\rho_{1910}$ will then be given by:

$$\text{plim}(b_{n,1910}) = \beta_1 \gamma_1 + OVB$$

where $OVB$ denotes the omitted variable bias. This bias depends on the covariances between $\theta_n$, $\theta_r$ and $\rho_{1910}$ and of course the parameters $\beta_1 \gamma_2$ and $\beta_2$. Consider next the regression of $n_{1920}$ on $\rho_{1910}$:

$$n_{1910} = \beta_1 \gamma_1 k \rho_{1910} + \beta_1 \gamma_2 \theta_r + \beta_2 \theta_n + u_{1920} + \beta_1 \varepsilon_{1920}$$

Here we get

$$\text{plim}(b_{n,1920}) = \beta_1 \gamma_1 k + OVB$$

We can then difference the 2 coefficients to get:

$$\text{plim}(b_{n,1910} - b_{n,1920}) = \beta_1 \gamma_1 (1 - k)$$
In similar fashion we obtain

$$\text{plim}(b_{h,1910} - b_{h,1920}) = \alpha_1 \gamma_1 (1 - k)$$

and then

$$\frac{\text{plim}(b_{n,1910} - b_{n,1920})}{\text{plim}(b_{h,1910} - b_{h,1920})} = \frac{\beta_1}{\alpha_1}$$

This ratio is identical to the ratio of partial derivatives

$$\frac{\partial n}{\partial \rho} / \frac{\partial h}{\partial \rho} = \frac{\frac{\partial n}{\partial r} \frac{\partial r}{\partial \rho}}{\frac{\partial h}{\partial r} \frac{\partial r}{\partial \rho}} = \frac{\partial n}{\partial h}$$

we attempted to find in the first place. We can thus compare this ratio with the overall decline in fertility relative to the increase in education observed as the American South developed.

### 5.2 Results

Table VI gathers the necessary data to perform these calculations. The aggregate changes in fertility and school enrollment are obtained by averaging using the IPUMS data and reproduced in the first two rows of Table VI. The regression results reported in Table II, panel 1 and 2, column 1 provide us with the post-period interaction of the hookworm infection rate in 1910 in the schooling enrollment and fertility regressions. These coefficients are replicated in the next two rows of Table VI. The ratio of these two coefficients yields the change in fertility relative to the change in human capital investments that is associated with the hookworm eradication campaign. This number is approximately -1.24. Thus, changes in fertility and school enrollment due to the hookworm eradication campaign imply that an increase in school enrollment by 1 percentage point is associated with a decline in the number of children less that 5 years per female 15-49 of about 0.0124. For the aggregate data in 1910 and 1920 we have that the ratio of the fertility change over the education change is equal to about -0.88. The estimated changes in school enrollment and fertility implicit in the hookworm experiment is therefore of an equal order
of magnitude as the relative change in aggregate school enrollment and fertility observed during this time-period. These calculations are naturally imprecise, but they do suggest that forward looking models of the demographic transition might be sufficient to explain the decline in fertility observed during the demographic transition as economies develop.

6 Conclusion

This study contains evidence on the importance of competing models of the fertility transition observed during the development of societies to high income societies. The episode analyzed is the eradication of hookworm disease in the American South (c. 1910). In previous work, it was shown that the eradication of hookworm disease led to a significant increase in school attendance and literacy. The present study shows that this increase in human capital was accompanied by a fertility decrease that was both economically and statistically significant. A decline in the hookworm infection rate from 40 to 20% is associated with a decline in fertility that amounts to 40% of the entire fertility decline observed in the American South between 1910 and 1920.

These results can be used to test theoretical models on the interaction of fertility and human capital investments in growth. It provides broad support for models of the fertility transition that link parental fertility decisions to the educational investment decisions parents face with respect to their children. It therefore strengthens the empirical support for the emerging literature linking human capital investment and fertility in models of economic growth and demographic transitions. These models argue that increases in returns to education cause the simultaneous observed increases in human capital investments and declines in fertility rates.

The data on hookworm eradication suggests that an increase in returns to education that increases school enrollment rates by one percentage point (from a base of about 0.80 for the enrollment rate of children aged 8-16) results in a decline in the 5-year fertility rates of women aged 15-49 of approximately 0.012 (from a basis of about 0.55).

The relative change in fertility and schooling caused by hookworm eradication are approximately equal to aggregate comovements during the period considered. This cor-
respondence suggests an important role for the interaction of fertility and human capital investments in growth.

References

Augustine, D. L. and W. G. Smillie (1926). ”The Relation of the Type of Soils of Alabama to the Distribution of Hookworm Disease.” American Journal of Hygiene 6 (Supplement 1), 36-62.


Data Appendix

A Sources and Definitions for the Micro Data

The micro data consist of samples drawn from the Censuses of 1900, 1910, 1920, 1930, 1940, and 1950, accessed through the IPUMS project (Ruggles and Sobek (1997)). (The 1910 data include the black and hispanic oversamples. Results are not sensitive to their exclusion.) The sample consists of native-born whites and blacks in the age range \([8,16]\) in the case of children, and in the age range \([15,55]\) in the case of adults. The age criteria for children serves to select children of school age who are likely not yet old enough to have migrated on their own. The outcome variables are defined as follows:

- **School enrollment.** This is an indicator variable for whether the child has attended school at any time during a specified interval preceding the day of the Census. The length of this interval varies across the Censuses as follows: 1900, within the past year; 1910 and 1920, since September 1st; 1940, since March 1; 1950, since February 1.

- **Regular school attendance.** This is an indicator variable that is switched on if the child is attending and not working. We consider a child to be working if the census recorded an occupation for him/her, which corresponds to an IPUMS occ1950 code that is not missing and less than or equal to 970.

- **Literacy.** This variable is an indicator for the ability to read and/or write. Census questions contained categories for being able to read but not write, *vice versa*, both or neither. We coded the first three as literate. (The first two responses were relatively rare.) These data were collected only for the 1900–20 samples. The literacy question was only asked for individuals 10 years and older.

- **Fertility.** The census data provide the following information on own children living in the household: total number, total number of those under age five, and the age of youngest child. These variables are used to define the following dependent variables:
  - The number of children less than 5.
  - An indicator variable for whether the woman has a child less than 5 years old.
  - An indicator variable for whether the woman has a child less than 1 year old.
  - The total number of children.
  - An indicator variable for whether the woman has any children.

- **Labor-force participation.** A binary variable indicating whether the individual is working. Prior to 1940, this variable is based on whether the individual’s reported occupation was classified as a “gainful” one.
• **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 also calibrated using data from the 1950 Census.

• **Works in agriculture.** Defined as the IPUMS variable “occ1950” being equal to 100, 123, or anything in the 800s.

• **Works as Craftsman or Operative.** Defined as the IPUMS variable “occ1950” being between 500 and 699 (inclusive).

## B Sources and Definitions for the Aggregate Data

There are two units of observation for the area-level data: county and state economic area (SEA). Because county boundaries change over time and because county of residence is not available in the later Censuses, we use the SEA as the aggregate unit for the sequential-cross-section analysis. The SEAs are aggregations of counties, with an average number of 8.5 counties per SEA. SEA boundaries tend to be more stable, in part because they were often defined by a state boundary or significant natural feature (river or mountain range, e.g.).

The area-level data come from a variety of sources, but principally from the RSC annual reports and the ICPSR’s study #3, the latter of which is a collection of historical Census tabulations. The following is a list of the aggregate variables with information on sources, definitions and method of aggregation. The source is indicated in parentheses at the end of each item’s text.

• **Hookworm infection rate.** The source data are at the county level and from the period 1911–1915. The infection numbers in most cases are from surveys conducted by the Rockefeller Sanitary Commission (RSC) as prelude to (or simultaneously with) dispensing treatments. In a few instances, the RSC dispensaries had already visited the county before making the survey. For this latter case, we use the examinations conducted by the dispensaries to construct the hookworm infection rate, rather than using data that comes after the administration of the RSC treatments. (The hookworm-infection rates constructed from survey and examination have a correlation coefficient greater than 0.95 for those cases in which the survey was done first.) The infection data were aggregated to the SEA level using a population-weighted average. (RSC annual reports.)
• **Individuals treated by the RSC, per capita.** The source data are at the county level and from the period 1911–1915. The RSC dispensaries tracked how many individuals received de-worming treatments. We sum these numbers to the SEA level and divide by total population. (RSC annual reports.)

• **Sanitary Index.** The RSC conducted independent surveys of the condition of sanitation infrastructure, including whether buildings had proper latrines, clean water sources, etc. Several measures of sanitation were combined by the RSC to form an index. The coverage of this indicator is incomplete. Because SEAs were reasonably homogeneous, we construct a simple average of these indices, ignoring the missing data within SEA. (RSC annual reports.)

• **Examined by RSC per capita.** The source data are at the county level and from the period 1911–1915. The RSC dispensaries tracked how many individuals were examined by the dispensary’s medical staff. We sum these numbers to the SEA and state levels and divide by total population. (RSC annual reports.)

• **County spending.** Data were input at the county level on county-government spending on education and health/sanitation for the years 1902 and 1932. (The 1922 publication in the series did not include these categories of spending, and the 1913 publication did not include earmarked transfers from the state government.) The health spending is normalized by total population, while the education expenditure is normalized by school-age population. We construct a population-weighted average for the SEAs. (County level: U.S. Bureau of the Census, 1915b and 1935. State level: annual reports of the federal Commissioner of Education.)

• **Full-time health officer.** These data were compiled at the county level, and include information on the first year each county employed a full time health officer. We coded this variable as one if such an office was created between 1910 and 1920 (inclusive). Only one county (Jefferson county in Kentucky) had created such a post before 1910, and the results above are not sensitive to its reclassification. To aggregate to the SEA level, we create a population-weighted average of the constituent counties. (Ferrell *et al.* 1932.)

• **WWI Cantonment size per capita.** We collected data on the troop numbers that were mustered and trained at the major Army cantonments of mobilization/embarkation for the First World War. Of the 32 cantonments, there were 19 camps in the South. These were “camps and cantonments that were used for mobilizing and training combat divisions,” but this excludes “miscellaneous groups” which comprised the “special camps, usually of semipermanent construction that were intended for mobilizing and training special troops, such as the Quartermaster Department Camp, Camp Joseph E. Johnston, Jacksonville, Florida.” We input the highest value given for the number of soldiers within a camp during 1918-20, aggregated to SEA level, and then normalized by total 1910 population. (Bowen, 1928.)
• **Malaria mortality.** These data were compiled at the county level and refer to the period 1919–1921. To construct SEA rates, we create a population-weighted average of the constituent counties. (Maxcy, 1923.) The state-level data are the fraction of total mortality in 1890. (Census, 1894.)

• **Fertility Rate, 1910, and Change in fertility, 1900–10.** The fertility rate for 1910 is measured from Census tabulations under the fraction of the population under six years of age, defined as $1 - \frac{(v41 + v53)}{(v20 + v21)}$ using the ICPSR variable names. For 1900, the tabulations permit calculating the fraction of the population under five for 1900, or $1 - \frac{(v22 + v37 + v39 + v41 + v43)}{(v8 + v10)}$. When computing the approximate difference, we up-weight the 1900 number by $5/4$. To construct SEA rates, we sum the components over the constituent counties and apply the above formulae. State data is available through a separate tabulation, to which we apply the same formulae. (ICPSR Study #3.)

• **Child Mortality, 1900–10.** Defined from IPUMS samples using the methodology of Preston and Haines (1984), and aggregated to SEA level.

• **Fraction black, 1910.** These data come from the 1910 Census. Defined as the fraction of the areas males who are black, out of the total population of blacks and whites. Specifically this is defined as $\frac{(v24 + v25)}{(v24 + v25 + v22 + v23)}$, using the variable codes of the ICPSR study. To construct SEA rates, we sum the components over the constituent counties and apply the above formulae. (ICPSR Study #3.)

• **Rosenwald schools per capita.** This measures the number of classrooms per capita built by the Julius Rosenwald Fund as of 1930. The denominator normalizes the number of classrooms by the population of blacks aged 5–19 in 1930. To construct SEA rates, we create a black-population-weighted average of the constituent counties. (Johnson *et al.*, 1941.)

• **Lynchings per capita, 1900–30.** The base data is the number of lynchings per 100,000 population by county in the years 1900-30. The denominator is the county population in 1930. To construct SEA rates, we create a population-weighted average of the constituent counties. (Johnson *et al.*, 1941.)

• **Population urban.** From Census tabulations measuring the population residing in metro areas. For 1910, the urban population is contained in variable $v9$ in the ICPSR data, which we scale by the total population as defined above. The 1900 fraction urban is also defined in the 1910 data as $v13/(v13 + v14)$. We construct the change in urbanization using the difference between the two variables. To construct SEA rates, we sum the components over the constituent counties and apply the above formulae. State data is available through a separate tabulation, to which we apply the same formulae. (ICPSR Study #3.)

• **Crop acreage per capita.** The base data measures the total farmed acreage at the county level, regardless of tenancy. This is constructed with the formula $(v155 + v164 + v175)$, using the ICPSR variable names, and scaled by total population. To
construct SEA rates, we sum the components over the constituent counties and apply the above formula. (ICPSR Study #3.)

- **Sharecropped areas per capita.** The base data is a county-level measure of total acreage sharecropped ($v164$ using the ICPSR variable scheme). We scale this by total population. To construct SEA rates, we sum the components over the constituent counties and apply the above formula. (ICPSR Study #3.)

- **Cotton acreage per capita.** The base data is cotton acreage in 1910 by county. We normalize this number by the county population, as defined above. To construct SEA rates, we create a population-weighted average of the constituent counties. (Census, 1915.)

- **Tobacco acreage per capita.** The base data is tobacco acreage in 1910 by county. We normalize this number by the county population, as defined above. To construct SEA rates, we create a population-weighted average of the constituent counties. (Census, 1915.)

- **Farm value per capita.** The base data is a county-level measure of the value of farm land and buildings, regardless of tenancy. This is defined as ($v177 + v166 + v157$) using the ICPSR variable scheme). We scale this by total population, as defined above. To construct SEA rates, we sum the components over the constituent counties and apply the above formula. (ICPSR Study #3.)

- **School Term Length.** Average length of school term, in weeks. Kentucky county data are imputed from cross-tabulated data on number of schools by months. The imputation is calibrated using Alabama data, which contain a continuous measure and a cross-tabulation. (County level: Reports of state departments of education, various years, and author’s calculations. State level: annual reports of the federal Commissioner of Education.)

- **Average Monthly Salaries for Teachers.** (County level: Reports of state departments of education, various years, and author’s calculations. State level: annual reports of the federal Commissioner of Education.)

- **School density.** Number of schoolhouse operating in the county, divided by land area in square miles. (Reports of state departments of education, various years, and author’s calculations.)

- **Change in Number of Teachers per School.** (Reports of state departments of education, various years, and author’s calculations.)

- **Pupil/Teacher Ratio.** Average attendance divided by number of teachers. (County level: Reports of state departments of education, various years. State level: annual reports of the federal Commissioner of Education.)

- **Value of School Plant and Equipment.** (Reports of state departments of education, various years, and author’s calculations.)
• **Returns to Literacy for Adults.** Measured from a regression of the occupational income score on literacy status, by SEA, for the 1910 and 1920 census samples of adults. (Authors’ calculations using the 1910 and 1920 IPUMS data.)

• **Baptists.** Defined as membership of Baptist Churches (1906) per total population (measured as the average of 1900 and 1910 total population). Baptist church members are reported by variables 10 through 21. (ICPSR Study #2896.)

• **Methodists.** Defined as membership of Methodist Churches (1906) per total population (measured as the average of 1900 and 1910 total population). Methodists churches are described by variables 63 through 71 plus variable 89.(ICPSR Study #2896.)

• **Catholic.** Defined as membership of Catholic Churches (1906) per total population (measured as the average of 1900 and 1910 total population). (ICPSR Study #2896.)

• **Non-Church Members.** Defined as non-church-pertaining population (1906) as a fraction of total (measured as the average of 1900 and 1910 total population). (ICPSR Study #2896.)

**References for Data Appendix**


Figure 1, Panel 1: Fertility in the American South

Figure 1, Panel 2: School Enrollment at Age 14 in the American South

The fertility measure depicted represents the number of children less than 5 years old per woman aged 15-49. School enrollment refers to children aged 14. The data stems from the census samples made available through the Public Use Microdata Series. States included in the sample are Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, South and North Carolina, Tennessee, Texas, and Virginia.
Figure 2: The Literature on the Demographic Transition

Explanatory Approaches for Demographic Transition

1. Innovation/Diffusion
2. Adjustment

2.1 Adjustment to Contemporaneous Conditions
2.2 Forward-Looking Adjustment

2.2.1 Altruism: Children ➔ Parents
2.2.2 Altruism: Parents ➔ Children
Figure 3: Geographic Distribution of Hookworm in the American South 1911-1915

Notes: This graph documents the hookworm infection rates found across the American South. The four shades of grey correspond to the quartiles of the hookworm distribution the SEA corresponds to, where the darker shades are SEAs with the highest infection rates.
Figure 4: Average Worm Counts by Age in Alabama and Brazil around 1920

This figure shows the age-profile of hookworm infection rates reported by Smilie and Augustine (1926) for Alabama (red) and Brazil (grey). It demonstrates that the age-profile of hookworm infestation rates can vary substantially across societies. Crucial for the empirical argument in this paper is that hookworm infection rates in the American South are concentrated among school-aged children. The decline in hookworm infestation rates after age 14 in the American South is probably linked to the age at which children started wearing shoes in the American South.
Figure 4: Highly Infected Areas Saw Greater Declines in Hookworm

Panel A: States

Panel B: Counties in Alabama

Notes: The y axis displays the decrease in hookworm infection post-intervention, as measured by follow-up surveys. The x axis is the pre-treatment hookworm infection rate, as measured by the Rockefeller Sanitary Commission. Panel A displays data at the state level, as reported by Jacocks (1924). Panel B contains data from counties in Alabama, as reported by Havens and Castles (1930). Both resurveys are from the early 1920s.
Figure 6, Panel 1: Impact of 1910 Hookworm Infection Rate on Fertility.

The solid line depicts the estimated impact of the 1910 Hookworm Infection Rate on the Fertility Rate. The broken lines represent 95% confidence intervals.

Figure 6, Panel 2: Impact of 1910 Hookworm Infection Rate on Schooling.

The solid line depicts the estimated impact of the 1910 Hookworm Infection Rate on the School Enrollment Rate. The broken lines represent 95% confidence intervals.
Figure 7: The Fertility Effect of Hookworm Eradication by Age

Note: Shown is the difference in the interaction between 1910 Hookworm infection rate and age between 1920 and 1910. The solid line is a lowess plot of these coefficients.
### Table I Summary Statistics

#### Panel A: Aggregate Data (by State Economic Area, 1910)

<table>
<thead>
<tr>
<th>Whole Sample</th>
<th>By Hookworm Infection</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&gt; 40%</td>
<td>&lt;40%</td>
</tr>
<tr>
<td>Hookworm-Infection Rate</td>
<td>0.328</td>
<td>0.560</td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>Treatments Issued by the RSC, Per Capita</td>
<td>0.059</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Least Once by the RSC, Per Capita¹</td>
<td>0.029</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>School Enrollment, 1910¹</td>
<td>0.721</td>
<td>0.711</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Δ Schooling Enrollment, 1910-20</td>
<td>0.089</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Generalized 5-year Fertility Rate</td>
<td>0.558</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.127)</td>
</tr>
<tr>
<td>Δ Generalized Fertility Rate, 1910-1920</td>
<td>-0.060</td>
<td>-0.080</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>115</td>
<td>48</td>
</tr>
</tbody>
</table>

#### Panel B: Micro Data

<table>
<thead>
<tr>
<th>Sample</th>
<th>1910 Census</th>
<th>1900-1950 Census</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children</td>
<td>Female Adults</td>
</tr>
<tr>
<td>Age</td>
<td>11.9</td>
<td>28.1</td>
</tr>
<tr>
<td></td>
<td>(2.6)</td>
<td>(9.47)</td>
</tr>
<tr>
<td>Children less than 5</td>
<td>0.556</td>
<td>0.507</td>
</tr>
<tr>
<td></td>
<td>(0.862)</td>
<td>(0.828)</td>
</tr>
<tr>
<td>School Attendance</td>
<td>0.712</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>(0.453)</td>
<td>(0.397)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>17,194</td>
<td>24,488</td>
</tr>
</tbody>
</table>

**Notes:** Displayed are means and (in brackets) standard deviations. The sample of children includes native born blacks and whites age 8-16 in IPUMS from RSC surveyed units in the indicated years. The adult sample includes native born black and white females aged 15-49 in IPUMS from RSC surveyed units in the indicated years. School enrollment is calculated using all native born black and whites in the RSC surveyed units between the ages 8 and 16. The Generalized Fertility Rate is the average number of children less than 5 years per female aged 15-49.

¹) The treatment intensity variables are only available for 114 (rather than 115) SEAs.
### Table II. Panel 1 The Effect of Hookworm Eradication on Human Capital Investments

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>School enrollment</th>
<th>Regular School attendance</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>%-Infected in 1910* Post Period</td>
<td>0.0827*** (0.0220)</td>
<td>0.1565*** (0.0254)</td>
<td>0.0500*** (0.0200)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>64,676</td>
<td>64,676</td>
<td>49,476</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1207</td>
<td>0.1967</td>
<td>0.1212</td>
</tr>
</tbody>
</table>

Notes: Reported are the estimated coefficients on the interaction of the % infected by 1910 S.E.A. with the post-period indicator (year=1920) for various human capital investment measures as dependent variables. The specification corresponds to equation (1) described in the text. The reported specifications include S.E.A. fixed effects as well as fully interacted sets of Age, Race, Religion, and Post-period indicators. Simple asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of all native-born children in the IPUMS between the ages of 8 and 16 in the RSC-surveyed geographic units for 1910 and 1920.

### Table II. Panel 2 The Effect of Hookworm Eradication on Fertility

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Children less than five years old</th>
<th>At least one child less than five years old</th>
<th>At least one child less than one year old</th>
<th>Total number of children</th>
<th>At least one child</th>
</tr>
</thead>
<tbody>
<tr>
<td>%-Infected in 1910* Post Period</td>
<td>-0.1023*** (0.0239)</td>
<td>-0.0529*** (0.0127)</td>
<td>-0.0330*** (0.0110)</td>
<td>-0.1554*** (0.0542)</td>
<td>-0.0381*** (0.0135)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>77,688</td>
<td>77,688</td>
<td>77,688</td>
<td>77,688</td>
<td>77,688</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1594</td>
<td>0.169</td>
<td>0.0501</td>
<td>0.3157</td>
<td>0.3171</td>
</tr>
</tbody>
</table>

Notes: Reported are the estimated coefficients on the interaction of the % infected by 1910 S.E.A. with the post-period indicator (year=1920) for various fertility measures as dependent variables. The specification corresponds to equation (1) described in the text. The reported specifications include S.E.A. fixed effects as well as fully interacted sets of Age, Race, Religion, and Post-period indicators. Simple asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of females in the IPUMS between the ages of 15 and 49 in the RSC-surveyed geographic units for 1910 and 1920.
<table>
<thead>
<tr>
<th>Dependent Variables:</th>
<th>Children less than five years old</th>
<th>At least one child less than five years old</th>
<th>At least one child less than one year old</th>
<th>Total number of children</th>
<th>At least one child</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Baseline Results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment Hookworm</td>
<td>-0.1023***</td>
<td>-0.0529***</td>
<td>-0.0330***</td>
<td>-0.1554***</td>
<td>-0.0381***</td>
</tr>
<tr>
<td></td>
<td>(0.0239)</td>
<td>(0.0127)</td>
<td>(0.0110)</td>
<td>(0.0542)</td>
<td>(0.0135)</td>
</tr>
<tr>
<td><strong>Panel B: Health and Health Policy Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment Hookworm</td>
<td>-0.0691**</td>
<td>-0.0531***</td>
<td>-0.0127</td>
<td>-0.1415***</td>
<td>-0.0385**</td>
</tr>
<tr>
<td></td>
<td>(0.0276)</td>
<td>(0.0154)</td>
<td>(0.0136)</td>
<td>(0.0577)</td>
<td>(0.0179)</td>
</tr>
<tr>
<td>Examined by the RSC, per capita</td>
<td>0.2051**</td>
<td>0.1583***</td>
<td>-0.0059</td>
<td>0.5200***</td>
<td>0.0663</td>
</tr>
<tr>
<td></td>
<td>(0.0971)</td>
<td>(0.0577)</td>
<td>(0.0376)</td>
<td>(0.2142)</td>
<td>(0.0469)</td>
</tr>
<tr>
<td>RSC Sanitary Index</td>
<td>-0.0009</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0021</td>
<td>-0.0010**</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0004)</td>
<td>(0.0003)</td>
<td>(0.0019)</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Change in County Spending on Health and Sanitation, 1902-1932</td>
<td>-0.0525</td>
<td>-0.0230</td>
<td>-0.0091</td>
<td>-0.2260**</td>
<td>-0.0909***</td>
</tr>
<tr>
<td></td>
<td>(0.0514)</td>
<td>(0.0289)</td>
<td>(0.0247)</td>
<td>(0.1019)</td>
<td>(0.0318)</td>
</tr>
<tr>
<td>Full-time Health Officer, per capita</td>
<td>-0.0378</td>
<td>0.0245</td>
<td>-0.0297</td>
<td>-0.0166</td>
<td>-0.0083</td>
</tr>
<tr>
<td></td>
<td>(0.0345)</td>
<td>(0.0172)</td>
<td>(0.0145)</td>
<td>(0.0846)</td>
<td>(0.0211)</td>
</tr>
<tr>
<td>Malaria mortality, c. 1917</td>
<td>0.0026</td>
<td>0.0013</td>
<td>0.0022*</td>
<td>0.0052</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>(0.0032)</td>
<td>(0.0017)</td>
<td>(0.0011)</td>
<td>(0.0073)</td>
<td>(0.0016)</td>
</tr>
<tr>
<td>WWI Camp Size, per capita, c. 1918</td>
<td>-0.1184</td>
<td>-0.0909***</td>
<td>0.0273</td>
<td>-0.4529***</td>
<td>-0.0514</td>
</tr>
<tr>
<td></td>
<td>(0.0780)</td>
<td>(0.0336)</td>
<td>(0.0245)</td>
<td>(0.1558)</td>
<td>(0.0445)</td>
</tr>
<tr>
<td>Fertility Rate, 1910</td>
<td>-0.8368**</td>
<td>-0.1472</td>
<td>-0.3220*</td>
<td>-0.4441</td>
<td>-0.0301</td>
</tr>
<tr>
<td></td>
<td>(0.3751)</td>
<td>(0.1980)</td>
<td>(0.1663)</td>
<td>(0.8202)</td>
<td>(0.1973)</td>
</tr>
<tr>
<td>Decline in Fertility, 1900-10</td>
<td>0.2132</td>
<td>0.5973</td>
<td>-0.1168</td>
<td>1.1785</td>
<td>0.7843*</td>
</tr>
<tr>
<td></td>
<td>(0.8890)</td>
<td>(0.4355)</td>
<td>(0.3733)</td>
<td>(1.8200)</td>
<td>(0.4316)</td>
</tr>
<tr>
<td>Child Mortality ca. 1910</td>
<td>0.6053***</td>
<td>0.2892***</td>
<td>0.2702***</td>
<td>1.6950***</td>
<td>0.2623**</td>
</tr>
<tr>
<td></td>
<td>(0.1604)</td>
<td>(0.0905)</td>
<td>(0.0623)</td>
<td>(0.3690)</td>
<td>(0.1087)</td>
</tr>
<tr>
<td>F-stat on controls:</td>
<td>4.58***</td>
<td>7.50***</td>
<td>3.23***</td>
<td>6.43***</td>
<td>7.56***</td>
</tr>
<tr>
<td></td>
<td>(9,187)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Panel C: Education and Race Controls

<table>
<thead>
<tr>
<th></th>
<th>1910-20</th>
<th>1910-20</th>
<th>1910-20</th>
<th>1910-20</th>
<th>1910-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-treatment Hookworm</td>
<td>-0.0990***</td>
<td>-0.0481***</td>
<td>-0.0302***</td>
<td>-0.1186**</td>
<td>-0.0401***</td>
</tr>
<tr>
<td></td>
<td>(0.0254)</td>
<td>(0.0139)</td>
<td>(0.0099)</td>
<td>(0.0565)</td>
<td>(0.0136)</td>
</tr>
<tr>
<td>Change in School Term, c. 1910-20</td>
<td>0.0007</td>
<td>0.069</td>
<td>-0.0014</td>
<td>0.0119</td>
<td>0.0062</td>
</tr>
<tr>
<td></td>
<td>(0.0078)</td>
<td>(0.0042)</td>
<td>(0.0031)</td>
<td>(0.0216)</td>
<td>(0.0051)</td>
</tr>
<tr>
<td>Change in Average Monthly</td>
<td>-0.0610***</td>
<td>-0.0297***</td>
<td>-0.0254***</td>
<td>-0.0321</td>
<td>-0.0158</td>
</tr>
<tr>
<td>Teacher Salary, c. 1910-20</td>
<td>(0.0167)</td>
<td>(0.0104)</td>
<td>(0.0064)</td>
<td>(0.0481)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>Change in Teachers per School, c. 1910-20</td>
<td>0.0243</td>
<td>0.0154</td>
<td>0.0224**</td>
<td>-0.0331</td>
<td>-0.0104</td>
</tr>
<tr>
<td></td>
<td>(0.0209)</td>
<td>(0.0134)</td>
<td>(0.0096)</td>
<td>(0.0549)</td>
<td>(0.0139)</td>
</tr>
<tr>
<td>Change in Number of Schools per Square Mile, c. 1910-20</td>
<td>-0.0072</td>
<td>-0.0079</td>
<td>-0.0001</td>
<td>-0.0947**</td>
<td>-0.0284*</td>
</tr>
<tr>
<td></td>
<td>(0.0212)</td>
<td>(0.0112)</td>
<td>(0.0084)</td>
<td>(0.0423)</td>
<td>(0.0146)</td>
</tr>
<tr>
<td>Change in Pupil/Teacher Ratio, c. 1910-20</td>
<td>-0.0586***</td>
<td>-0.0193*</td>
<td>-0.0118</td>
<td>-0.1759***</td>
<td>-0.0165</td>
</tr>
<tr>
<td></td>
<td>(0.0189)</td>
<td>(0.0111)</td>
<td>(0.0078)</td>
<td>(0.0460)</td>
<td>(0.0130)</td>
</tr>
<tr>
<td>Change in Value of School Plant and Equipment, c. 1910-20</td>
<td>0.0006</td>
<td>0.0000</td>
<td>-0.0019</td>
<td>-0.0147</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0080)</td>
<td>(0.0045)</td>
<td>(0.0029)</td>
<td>(0.0193)</td>
<td>(0.0043)</td>
</tr>
<tr>
<td>Change in Returns to Literacy for Adults, c. 1910-20</td>
<td>-0.0021</td>
<td>-0.0031*</td>
<td>-0.0013</td>
<td>-0.0242***</td>
<td>-0.0048**</td>
</tr>
<tr>
<td></td>
<td>(0.0033)</td>
<td>(0.0017)</td>
<td>(0.0013)</td>
<td>(0.0072)</td>
<td>(0.0020)</td>
</tr>
<tr>
<td>Change in County Educational Spending, per child, 1902-32</td>
<td>0.0028</td>
<td>-0.0008</td>
<td>0.0015</td>
<td>-0.0068</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(0.0037)</td>
<td>(0.0021)</td>
<td>(0.0013)</td>
<td>(0.0092)</td>
<td>(0.0025)</td>
</tr>
<tr>
<td>Child Literacy Rate, 1910</td>
<td>0.0377</td>
<td>-0.1143</td>
<td>0.1985**</td>
<td>-0.2754</td>
<td>-0.2988**</td>
</tr>
<tr>
<td></td>
<td>(0.2342)</td>
<td>(0.1146)</td>
<td>(0.0954)</td>
<td>(0.5282)</td>
<td>(0.1485)</td>
</tr>
<tr>
<td>Adult Literacy Rate, 1910</td>
<td>-0.1973</td>
<td>-0.0294</td>
<td>-0.2584***</td>
<td>-0.2754</td>
<td>0.2559**</td>
</tr>
<tr>
<td></td>
<td>(0.1931)</td>
<td>(0.0993)</td>
<td>(0.0781)</td>
<td>(0.5282)</td>
<td>(0.1118)</td>
</tr>
<tr>
<td>Rosenwald-funded Classrooms, per capita, 1930</td>
<td>-0.0033***</td>
<td>-0.0012*</td>
<td>-0.0010**</td>
<td>-0.0053</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0012)</td>
<td>(0.0007)</td>
<td>(0.0004)</td>
<td>(0.0033)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Lynchings per capita, 1900-30</td>
<td>-0.0002</td>
<td>0.0001</td>
<td>0.0003</td>
<td>-0.0033</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0012)</td>
<td>(0.0008)</td>
<td>(0.0005)</td>
<td>(0.0030)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>Fraction Black, 1910</td>
<td>-0.0304</td>
<td>-0.0232</td>
<td>-0.0190</td>
<td>0.0786</td>
<td>-0.0370</td>
</tr>
<tr>
<td></td>
<td>(0.0455)</td>
<td>(0.0280)</td>
<td>(0.0165)</td>
<td>(0.1213)</td>
<td>(0.0292)</td>
</tr>
<tr>
<td>F-stat on controls:</td>
<td>4.85***</td>
<td>3.06***</td>
<td>4.91***</td>
<td>4.66***</td>
<td>2.12**</td>
</tr>
<tr>
<td></td>
<td>(13,187)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variables</td>
<td>Pre-treatment Hookworm</td>
<td>Fraction in Urban Area, 1910</td>
<td>Change in Fraction Urban, 1900-10</td>
<td>Farm Acreage, per capita, 1910</td>
<td>Farm value per capita, 1910</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------------------------</td>
<td>------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Pre-treatment Hookworm</td>
<td>-0.1020***</td>
<td>-0.0487***</td>
<td>-0.0317**</td>
<td>-0.2380***</td>
<td>-0.0494**</td>
</tr>
<tr>
<td></td>
<td>(0.0300)</td>
<td>(0.0181)</td>
<td>(0.0138)</td>
<td>(0.0670)</td>
<td>(0.0190)</td>
</tr>
<tr>
<td>Fraction in Urban Area, 1910</td>
<td>-0.0021</td>
<td>-0.0028</td>
<td>-0.0080</td>
<td>-0.2597***</td>
<td>-0.0060</td>
</tr>
<tr>
<td></td>
<td>(0.0373)</td>
<td>(0.0217)</td>
<td>(0.0170)</td>
<td>(0.0894)</td>
<td>(0.0260)</td>
</tr>
<tr>
<td>Change in Fraction Urban, 1900-10</td>
<td>-0.1015</td>
<td>-0.0139</td>
<td>-0.0665**</td>
<td>-0.3101*</td>
<td>-0.0066</td>
</tr>
<tr>
<td></td>
<td>(0.0684)</td>
<td>(0.0378)</td>
<td>(0.0264)</td>
<td>(0.1600)</td>
<td>(0.0040)</td>
</tr>
<tr>
<td>Farm Acreage, per capita, 1910</td>
<td>-0.0019</td>
<td>0.0072</td>
<td>-0.0067*</td>
<td>-0.0327</td>
<td>0.0066</td>
</tr>
<tr>
<td></td>
<td>(0.0092)</td>
<td>(0.0046)</td>
<td>(0.0034)</td>
<td>(0.0207)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td>Farm value per capita, 1910</td>
<td>-0.0191</td>
<td>0.0082</td>
<td>-0.0117**</td>
<td>-0.0795***</td>
<td>-0.0126</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0071)</td>
<td>(0.0054)</td>
<td>(0.0299)</td>
<td>(0.0077)</td>
</tr>
<tr>
<td>Sharecapped Acreage, per capita, 1910</td>
<td>-0.0148</td>
<td>-0.0157***</td>
<td>-0.0071*</td>
<td>0.0061</td>
<td>0.0074</td>
</tr>
<tr>
<td></td>
<td>(0.0101)</td>
<td>(0.0055)</td>
<td>(0.0041)</td>
<td>(0.0231)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>Cotton Acreage per capita, 1910</td>
<td>-0.0084</td>
<td>-0.0469*</td>
<td>-0.0140</td>
<td>0.0654</td>
<td>-0.0385</td>
</tr>
<tr>
<td></td>
<td>(0.0569)</td>
<td>(0.0280)</td>
<td>(0.0216)</td>
<td>(0.1220)</td>
<td>(0.0318)</td>
</tr>
<tr>
<td>Tobacco Acreage, 1910</td>
<td>-0.0082</td>
<td>0.0645</td>
<td>-0.0268</td>
<td>0.8030</td>
<td>0.1281</td>
</tr>
<tr>
<td></td>
<td>(0.2543)</td>
<td>(0.1206)</td>
<td>(0.0716)</td>
<td>(0.7174)</td>
<td>(0.1327)</td>
</tr>
<tr>
<td>Catholics</td>
<td>0.1885</td>
<td>0.1290</td>
<td>0.0760*</td>
<td>0.8610</td>
<td>0.0938</td>
</tr>
<tr>
<td></td>
<td>(0.1807)</td>
<td>(0.0893)</td>
<td>(0.0442)</td>
<td>(0.5273)</td>
<td>(0.1027)</td>
</tr>
<tr>
<td>Not Church Members</td>
<td>0.0889</td>
<td>0.0977</td>
<td>0.0410</td>
<td>0.8257</td>
<td>0.0984</td>
</tr>
<tr>
<td></td>
<td>(0.1808)</td>
<td>(0.0915)</td>
<td>(0.0481)</td>
<td>(0.5160)</td>
<td>(0.1052)</td>
</tr>
<tr>
<td>F-stat on controls:</td>
<td>5.26***</td>
<td>4.02***</td>
<td>4.91***</td>
<td>6.21***</td>
<td>4.65**</td>
</tr>
</tbody>
</table>

Panel D: Agricultural Controls

Panel E: Religion Controls
### Panel F: Include Above Controls Simultaneously

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-treatment Hookworm</td>
<td>-0.1008***</td>
<td>-0.0456**</td>
<td>-0.0090</td>
<td>-0.2501***</td>
<td>-0.0404*</td>
</tr>
<tr>
<td></td>
<td>(0.0330)</td>
<td>(0.0175)</td>
<td>(0.0134)</td>
<td>(0.0596)</td>
<td>(0.0215)</td>
</tr>
<tr>
<td>F-stat Health and Health Policy Controls</td>
<td>8.31***</td>
<td>8.35***</td>
<td>9.35***</td>
<td>15.77***</td>
<td>10.74***</td>
</tr>
<tr>
<td></td>
<td>(9,187)</td>
<td>(9,187)</td>
<td>(9,187)</td>
<td>(9,187)</td>
<td>(9,187)</td>
</tr>
<tr>
<td>F-stat Education and Race Controls</td>
<td>7.32***</td>
<td>6.90***</td>
<td>6.77***</td>
<td>5.00***</td>
<td>4.45***</td>
</tr>
<tr>
<td></td>
<td>(13,187)</td>
<td>(13,187)</td>
<td>(13,187)</td>
<td>(13,187)</td>
<td>(13,187)</td>
</tr>
<tr>
<td>F-stat Agricultural Controls</td>
<td>4.79***</td>
<td>6.88***</td>
<td>2.37**</td>
<td>5.53***</td>
<td>3.86***</td>
</tr>
<tr>
<td></td>
<td>(7,187)</td>
<td>(7,187)</td>
<td>(7,187)</td>
<td>(7,187)</td>
<td>(7,187)</td>
</tr>
<tr>
<td>F-stat Religious Controls</td>
<td>3.77***</td>
<td>4.50***</td>
<td>2.51**</td>
<td>6.09***</td>
<td>2.53**</td>
</tr>
<tr>
<td>F-stat All Controls</td>
<td>11.55***</td>
<td>9.03***</td>
<td>11.22***</td>
<td>14.29***</td>
<td>9.44***</td>
</tr>
<tr>
<td></td>
<td>(33,187)</td>
<td>(33,187)</td>
<td>(33,187)</td>
<td>(33,187)</td>
<td>(33,187)</td>
</tr>
</tbody>
</table>

Notes: Reported are the estimated coefficients on the interactions of the % infected by 1910 S.E.A as well as various controls with the post-period indicator (year=1920) for various fertility measures as dependent variables. All specifications include S.E.A. fixed effects as well as fully interacted sets of Age, Race, Religion, and Post-period indicators. F-statistics for the various control sets (not including the main variable infection*post) are provided together with their degrees of freedom. Panel G provides estimates from using all controls simultaneously. Again F-statistics on the various control-sets as well as on the entire set are provided. Description of controls are provided in the data appendix. Simple asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of females in the IPUMS between the ages of 15 and 49 in the RSC-surveyed geographic units for the indicated years.
## Table IV Hookworm Infection and Child Mortality in Pre-Period

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection*100</td>
<td>0.0141</td>
<td>0.0084</td>
<td>0.0095</td>
<td>0.0026</td>
<td>0.0073</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0106)</td>
<td>(0.0077)</td>
<td>(0.0083)</td>
<td>(0.0065)</td>
<td>(0.0079)</td>
<td>(0.0063)</td>
</tr>
<tr>
<td>Literacy Dummy</td>
<td>-0.009***</td>
<td>-0.009***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Age FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample Years</td>
<td>1910</td>
<td>1900, 1910</td>
<td>1910</td>
<td>1900, 1910</td>
<td>1910</td>
<td>1900, 1910</td>
</tr>
<tr>
<td>Sample</td>
<td>13,407</td>
<td>27,253</td>
<td>13,407</td>
<td>27,253</td>
<td>13,407</td>
<td>27,253</td>
</tr>
</tbody>
</table>

Dependent variable: Mortality rate calculated from micro-data using response to “number of children alive”, “number of children ever born” asked in 1900 and 1910. Robust standard errors in parenthesis (clustering on SEA times year). In columns (4) and (6) the race and age effects are fully interacted with year effects. Simple asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of females in the IPUMS between the ages of 15 and 49 in the RSC-surveyed geographic units for the indicated years.
<table>
<thead>
<tr>
<th>Panel</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Labor-Force Participation</td>
</tr>
<tr>
<td>B</td>
<td>Occupational Income Score</td>
</tr>
<tr>
<td>C</td>
<td>Duncan Socioeconomic Index</td>
</tr>
<tr>
<td>D</td>
<td>Lives on a Farm</td>
</tr>
<tr>
<td>E</td>
<td>Works in Agriculture</td>
</tr>
<tr>
<td>F</td>
<td>Works as Craftsman or Operative</td>
</tr>
<tr>
<td>G</td>
<td>Born in a different State</td>
</tr>
</tbody>
</table>

**Notes:** Each Panel/Column reports the results from a separate regression of the indicated variable on pre-treatment hookworm x post. The dependent variables are indicated in each Panel heading, and are derived from reported occupational and birthplace information. Robust standard errors in parenthesis (clustering on S.E.A. times post). Mean of dependent variable in square brackets. Single asterisk denote statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of all whites and blacks in the IPUMS for the indicated ages in the RSC-surveyed geographic units in 1910-20. All specifications include dummies for SEA and for age x black x Census region x year.
<table>
<thead>
<tr>
<th>Table VI How important is the Quantity-Qualty Mechanism in explaining the Demographic Transition?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>1910</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>Aggregate School Enrollment (age 8-16)</td>
</tr>
<tr>
<td>Aggregate Fertility (age 15-49)</td>
</tr>
<tr>
<td><strong>Coefficient on Infection Rate interacted with Post-Period</strong></td>
</tr>
<tr>
<td>Effect of Hookworm Eradication on Schooling</td>
</tr>
<tr>
<td>Effect of Hookworm Eradication on Fertility</td>
</tr>
</tbody>
</table>

Fertility Decline per Percentage Point Increase in Aggregate Schooling between 1910-1920: -0.88
Fertility Decline per Percentage Point Schooling Increase Implicit in Hookworm Eradication: -1.24

Notes: Reported are the aggregate school enrollment and 5-year fertility rates across the American South in 1910 and 1920 as well as the hookworm infection*(year=1920) interactions in column 1 of Table 2.1 and Table 2.2. The final two numbers are obtained by taking the ratio of the aggregate changes and the ratio of coefficient estimates respectively. See Section 5 for a discussion of the required assumptions.