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

Hoyt Bleakley† Fabian Lange‡

January 10, 2005

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 2004), 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. 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 argue that this evidence is consistent with models of the fertility transition emphasizing economic incentives rather than changing cultural attitudes and birth control technologies. Furthermore our data supports models emphasizing intergenerational altruism. Variables affecting childrens' economic prospects affect parental fertility decisions. A consequence of this finding is that we do not require changes in the economic opportunities faced by parents directly to explain the economic and demographic transition.

*We owe thanks to Gary Becker, Amitabh Chandra, Oded Galor, David Meltzer, Casey Mulligan, Kevin Murphy, John Strauss, Robert Topel, and seminar participants at the University of Chicago, the Ente Luigi Enaudi (Rome), the 2004 conference on Economics and Human Biolgy and the 2004 Econometric Society meetings for useful comments. The authors are, of course, responsible for any errors.

†Department of Economics, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508. E-mail: bleakley[at]ucsd[dot]edu

‡Department of Economics, Yale University, 27 Hillhouse, Room 38, New Haven, CT 06511. E-mail: 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 large theoretical and empirical literature linking economic development and fertility. There is however little sign that the available empirical evidence can generate a consensus among social scientists on the appropriate model of the fertility transition. Indeed, this literature has bifurcated into two major branches.

A significant number of researchers, most notably the participants of the European Fertility Project[1] (hereafter, EFP) support theories of the fertility transition that rely on the introduction and diffusion of new birth-control technologies as well as changes in the moral and cultural attitudes towards birth-contol. A competing literature instead focuses on changing economic incentives for fertility to explain the fertility transition. This division dates back at least to Gősta Carlsson (1966) who termed the former approach the "innovation" approach and the latter the "adjustment" approach.

Within the adjustment category, we distinguish between theories that rely heavily on economic incentives that affect the parents directly and those that emphasize the economic welfare and opportunities of children. Examples of variables that affect parents directly (and contemporaneously with their fertility decisions) are, for instance, the female wage (Becker, 1960; Mincer, 1963, and many others) and child mortality (Doepke, 2003; Kalemli-Ozcan, 2002). Models that emphasizes the economic welfare and opportunities of children usually rely on intergenerational altruism (plus some incompleteness of markets for intertemporal consumption smoothing). One approach emphasizes altruism of parents towards children: because parents cannot easily borrow against their children's future labor earnings, educational investment may be reduced by borrowing constraints. This restriction affects fertility through the interaction of average investment per child with number of children in the parental budget constraints. Alternatively, some emphasize the importance of child-to-parent altruism (Ehrlich and Lui, 1991; Boldrin and Jones, 2002). In developing countries, parents rely on children for old age support since no good substitutes for saving exist. Sub-

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

stitutes for saving become available as economies develop, leading to a decline in fertility. Within these approaches, investments into children (most notably education) and fertility are joint decisions.

The existing empirical literature relies heavily on analyzing the fertility transition using cross-national panel-data on fertility, education and other economic and cultural variables. The empirical problem is that it is difficult to distinguish between alternative theories using such data. The observed correlations between variables measuring economic development and fertility are consistent with a variety of different models. Indeed, most of the theoretical models have been constructed with an eye towards matching the time series. Thus, while the empirical literature has uncovered a variety of interesting patterns in the time-series, it has provided little basis for discriminating among the different theories.

We present new evidence that allows us to discern between these hypotheses. Our work differs from the vast majority of the empirical literature on the fertility transition in that it exploits quasi-experimental variation.[2] The present study links fertility to the economic prospects of children rather than to the labor-market opportunities of the parents. These results therefore support the literature emphasizing economic incentives, and more specifically models connecting children's economic welfare to parental fertility decisions through intergenerational altruism. 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 in the middle of its demographic and economic transition from a high fertility, low education economy to a society characterized by high levels of education and low fertility rates. (Figures 1

[2]One noteworthy exception is by Paul Schultz (1985), who uses certain agricultural-price shocks as exogenous variation in the value of female (i.e., potential mothers') labor in 19th-century Sweden. He shows that higher relative wages for women lead to lower fertility, which he interprets as a response to economic incentives. On the other hand, 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 by Rosenzweig and Wolpin (1980), who consider the quasi-experiment of a twin birth as exogenous variation in fertility.

and 2 show total fertility and school enrollment rates in the American South between 1850 and 2000.) 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. 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 adult productivity or child mortality.

The sudden and external origins of the eradication campaign combine with the 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 within the region. Within a short period of time, a large-scale eradication program was mounted through the efforts of the Rockefeller Sanitary Commission. This intervention is best understood as originating in developments outside the region (through the innovations to knowledge and funding). The substantial variation across the region in hookworm infection rates allows us to define treatment and comparison areas for the analysis. Areas with high infection rates had more to gain from the newly available treatments, whereas areas with little hookworm disease did not.

Hookworm infection increased the return to schooling or, alternatively stated, decreased the net price of raising a "quality" child. Using a research design similar to that of the present study, Hoyt Bleakley (2004) shows that hookworm eradication increased school attendance and ultimately the labor-market return to schooling. We take this result as indicative of the fundamental prices faced by the household, and then consider the response of fertility decisions.

Forward-looking models of the demographic and economic transition predict that such an increase in the returns to education will reduce contemporaneous fertility rates. We find robust evidence in favor of this prediction: the reduction in hookworm infection was accompanied by a fertility decline that was both economically and statistically significant. After hookworm eradication, fertility decreased markedly in areas that had previously suffered from

high rates of hookworm infection. This is true in absolute terms as well as relative to comparison counties that had lower levels of hookworm infection. We find this effect using either a two-period double difference or a multi-period setup that allows for differential trends across areas. Furthermore, the conclusion is robust to controlling for a variety of other alternative hypotheses, including crop-specific shocks, racial differences, and pre-existing differences in fertility and human capital.

Finally, our results suggest that the link between education and fertility can potentially explain the entire reduction in fertility rates observed during the region's demographic transition. In particular, we find that the relative responsiveness of fertility and education to the eradication of hookworm campaign is almost as large as the decline in fertility relative to the increase in education observed in the aggregate data during this time-period. This suggests that forward-looking behavior (though not specifically hookworm) is capable of explaining most if not all of the decline in fertility during the transition of the American South from a low income, high fertility society to a region of high incomes and low fertility.

The plan for the rest of the study is as follows. Section 2 reviews the existing literature on the fertility transition. Section 3 describes the historical episode that provides us with the natural experiment we consider in this study and discusses the identification strategy. This Section describes both the disease, the historical processes that led to the eradication of hookworm disease and the data used in this study. We also derive the identifying equation in Section 3. Section 4 presents the empirical results and addresses a number of alternative explanations for our findings. Section 5 shows, using the results presented in Section 4, that the quantity-quality trade-off is potentially very important in explaining this region's the fertility decline during the economic 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 for our empirical analysis to inform the debate on appropriate model for explaining the fertility transition. The classification we suggest is based on the basic distinction between *Innovation*

and *Adjustment* introduced by Gősta Carlsson (1966) as main explanatory approaches for the fertility transition. We summarize our classification using the tree presented in Figure 3. The remainder of this section to discusses the literature based on this classification.

Gősta Carlsson argued that we can distinguish between an *Innovation* approach and an *Adjustment* approach to explaining the fertility transition.[3] The *Innovation* approach (Fig. 3, node 1) explains the fertilty decline as the result of the (i) introduction and diffusion of birth-control practices and (ii) changing moral attitudes governing birth control. The *Adjustment* hypothesis (Fig. 3, node 2) relies instead on individuals adjusting their fertility behavior to varying economic incentives.

According to the *Innovation* approach either the absence of effective birth-control technologies and/or moral and cultural factors prevent the effective regulation of fertility in the pre-transition stage, leading to high marital fertility rates. The transition is then identified with the introduction and diffusion of contraceptive practices and abortion as countries develop. A broad interpretation of the innovation hypothesis emphasizes changing moral and cultural norms towards fertility control.

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. Much of the work done within the auspices of the European Fertility Project (hereafter EFP) consisted of dating the onset of the fertility decline using marital fertility[4] rates. The fertility decline occurred within a very short period of time across most of the European continent, independent of the social and economic conditions of the different societies. No threshold levels of important socioeconomic variables such as childhood-mortality, urbanization or education could be discerned that

[3]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 Guinanne (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.

[4]One of the major findings of the EFP was that the fertility transition was characterized by a sustained decline in marital fertility rates. Most of the decline in fertility rates is achieved by reductions in marital fertility rates, not by lowering rates of marriage or declines in pre-marital fertility rates. Pre-transition fluctuations in fertility rates are often achieved by variation in the age of marriage.

triggered a 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 to be 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 towards fertility control caused the fertility transition in Europe.

This explanatory approach contrast with the *Adjustment* approach that emphasizes decentralized adjustments to changing economic incentives for child-bearing as causing the fertility transition. It is not surprising that an explanatory approach emphasizing economic incentives has received substantial attention among economists. Thomas Malthus (1798) was one of the first proponents of linking fertility to economic variables, even though his model did not predict a sustained fertility decline.

A large body of work (Fig. 3, node 2.1) emphasizes variables that affect the contemporanous 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 prices of children and other consumption goods as well as the income of parents. Variables affecting income and the price of children therefore play an important role in his analysis. Becker and later Jacob Mincer (1963) argue that female wages affect the opportunity costs of bearing children. Fertility rates decline as female wages and thus the economic costs of rairing children increase. Paul Schultz (1985) finds substantial support for this explanation of the fertility transition in Swedish data from the second half of the 19th century. Another variable that has received substantial attention and directly affects the costs of bearing children is child mortality.[7] Furthermore urbanization will lead to increases in the price of children, since the costs of food and clothing are usually higher in urban setting while the contribution of children to

[5]In Coale and Watkins (1986), page 448.
[6]In Coale and Watkins (1986), page 293.
[7]Doepke (2003) provides a recent review and critique of the literature linking child mortality and fertility.

household income are smaller than in rural settings. To summarize, the contemporaneous view emphasizes variables that impinge directly on the choice of parents through income and prices of children. Such variables are for instance female wages, infant mortality rates and urbanization since these affect the costs of bearing children directly.[8]

The *Adjustment* literature also includes a large literature emphasizing altruistic links between the generations coupled with incomplete markets (Fig. 3, 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. We can further distinguish two types of models by the direction of altruism between the generations. The first (Figure 3, 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. 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.

Altruism from children to parents (Fig. 3, 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 to parents, but also societal norms provide a reason why children will provide for parental consumption when these are too old to provide for themselves. Underdeveloped societies are generally characterized by an absence of financial savings instruments. As countries develop they develop forms of communal savings such as social security and usually also develop financial savings mechanisms. Both social security and financial markets serve as substitutes

[8]Becker and Barro (1988) emphasize the role of wages and wealth on fertility if markets are complete and parents are altruistic towards their children. Parents then take into consideration the economic conditions faced by their off-spring later in life. This allows for variables affecting children to enter into the choice problem of parents. The predictions that arise from this extensions differ however substantially from those when markets are incomplete. In particular, models of the fertility decline that rely on altruistic considerations have generally also imposed that some markets are incomplete. We therefore do not consider Becker and Barro as a candidate for modeling the fertility transition.

for saving through children and therefore reduce fertility.

3 Identifying the Fertility/Schooling Relationship: The Hookworm-Eradication Campaign

The empirical difficulty in distinguishing between the hypotheses of Section 2 is that they all have 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 exogeneous in order to examine whether the response in fertility rates and other relevant variables is consistent with the theoretical models. Paul Schultz (1985) provides an example of such an approach. He examined the response in fertility rates to variation in female and male wages caused by variation in world prices for agricultural products intensive in female labor. Schultz found about 1/4 of the decline in fertility observed in Sweden between 1850-1910 can be attributed to increases in female wages. His results therefore support an *Adjustment* interpretation of the fertility decline. And, his results suggest that economic variables that impinge on the contemporaneous costs of child bearing are important for understanding the fertitlity transition. A different approach is taken by Rosenzweig and Wolpin (1980). These authors rely on quasi-exogeneous variation at the individual level. They examine the impact of twin births on the average level of education and find support for the quality-quantity tradeoff in data from India. Unfortunately there are few other empirical papers that rely on quasi-exogenous variation to evaluate the competing explanations for the fertility transition.

In the present study, we employ data from a policy intervention during the development of the American South: the eradication of Hookworm disease. We argue that our results provide evidence that forward-looking, altruism based behavior contributes to the reduction in fertility observed during this time-period. This argument is based on three main assumptions regarding Hookworm disease and the policy intervention initiated by the Rockefeller Sanitary Commission:

1. The policy intervention was plausibly exogenous to the development of the South. Below, we discuss how the eradication campaign originated from outside the region and in response to scientific innovations.

2. Hookworm disease affected mainly children and not the adult population. This assumption is supported by the infection surveys of the period (see Smillie and Augustine (1925), e.g.).[9]

3. Mortality from hookworm infection was extremely rare in the American South, and so we are able to consider the effects on fertility through morbidity and its effect on the return to schooling.

A number of additional auxiliary assumptions are made and we attempt to provide supporting evidence for each of these assumptions below. Finally, we 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.

3.1 The Eradication of Hookworm Disease

In the remainder of this Section, we discuss the eradication of hookworm disease in greater detail. We start by presenting the disease and the historical circumstances that led to its eradication as well as the data used in this study. We then argue that this eradication

[9]Note that 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).

represented an exogenous event in the development of the American South and is therefore particularly well suited for our purposes. The section closes with a more detailed description of the identification strategy.

3.1.1 Hookworm Disease

Hookworm is an intestinal parasite that infects humans. The worm lodges itself in the victim's intestine and absorbs nutrients from the bloodstream. As a result, the symptoms of hookworm infection (or *uncinaria*) are lethargy and anemia. The life cycle of the hookworm is dependent on unsanitary conditions. The nematodes lay their eggs in the intestine, but these larvae are passed out of the digestive system in feces. Hookworm is therefore transmitted through skin contact with infected excrement. The larvae burrow their way in through the skin, generating what was referred to in the South as "ground itch" in the era before people understood that hookworm was a potential problem. The lifespan of a hookworm is much shorter than that 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 de-worming medicine, which has a strong purgative effect. In the early 20th century the treatment available was a combination of thymol and epsom salts. Since then, safer and more effective medicines have become available. On the other hand, preventing reinfection will also have the effect of reducing the condition since the "worm burden" goes down over time without reinfection. Preventative measures include taking measures to limit skin contact with possibly polluted soil (the use of shoes being a leading example) and dealing with excrement in ways that minimize soil pollution in the first place (for example, the use of sanitary latrines). Of these, the private benefit to the former action is clear, whereas the latter activity clearly has an external component.

Symptoms of hookworm disease are clearly recognizable in historical documents going back as far as 1550 B.C. in Egypt. Moreover, the historical record is replete with subsequent evidence (RSC, 1911). The hookworm was itself identified in 1838 in Europe, although the transmission mechanism was not fully understood until 1898. Although no direct evidence

exists for earlier periods, Brinkley (1997) makes the case that hookworm was prevalent as least as early as the American Civil War. In 1910, hookworm was endemic all across the coastal plain of the American South. Important for our purposes is that a European-educated American physician (Charles W. Stiles) recognized hookworm symptoms in the American South in the last decade of the XIXth century.

3.1.2 The Rockefeller Sanitary Commission

The Rockefeller Sanitation Commission for the Eradication of Hookworm Disease (RSC) was formed in 1910 with the donation of one million dollars by John D. Rockefeller. This was emblematic of the emerging movement of large-scale philanthropy of the day. In this period, the so called "robber barons" were aging and turning from wealth accumulation to philanthropy. Through intermediaries, Dr. Stiles had convinced Rockefeller that taking on hookworm was a good foray in large-scale charity.

One of the Commission's first undertakings was to conduct surveys of hookworm infection rates among children across the region. One reason for this was logistical: the RSC could more efficiently allocate resources if it knew where the problem was. Another reason was political. Many leaders simply did not believe Stiles' hypothesis that a tiny parasite could be a root cause for some many of the problems of the large underclass (both white and black) in the region. As a result, the RSC surveyed over 600 counties in the South and found hookworm infection to be over 40% among children. After this finding, local opposition largely gave way to coöperation (Farmer, 1970).

Another finding from the infection surveys was 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. This consisted of two components. First, the RSC sent teams of health-care workers to counties to administer and dispense de-worming treatments free of charge. Through this system, some 400,000 individuals were treated with

de-worming medication. The second aspect of the campaign was publicity. The RSC sought to educate doctors, teachers, and the general public in how to recognize the symptoms of hookworm disease so that fewer cases would go untreated. Another part of this publicity campaign was 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.

3.1.3 Identification Strategy

The strategy for identifying the effect of the the hookworm-eradication campaign is based on three factors:

1. the incidence rate of hookworm disease varied widely in the American South;
2. the initiation of the campaign by the RSC was largely a function of factors external to the Southern states[10];
3. the timing of the treatment efforts was reasonably abrupt and well defined.

The marked heterogeneity in hookworm infection across the region was one of the first facts demonstrated by the RSC surveys. Much of this variance was due to geographical differences, as discussed by Augustine and Smillie (1926). 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 the denizens of 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.

[10]The historical presentation in this section draws heavily on the work of Ettling (1981).

As to point (2), the central impetus for the eradication campaign was due to critical innovations to knowledge. This contrasts with explanations that might have potentially troublesome endogeneity problems, such as, say, changes in government spending or positive income shocks in the highly infected areas. Instead, the innovations were to medical knowledge: understanding how the disease worked and more importantly recognizing its presence.

Such innovations were not related to or somehow in anticipation of the future growth prospects of the affected areas, and therefore should not be thought of as endogenous in this context. For example, as mentioned above hookworm's transmission mechanism was discovered by a European physician who accidentally infected himself while diagnosing a patient in Egypt. At that time, hookworm infection in the American South was not even understood to be a problem. That would wait until a one Charles W. Stiles came onto the scene.

Moreover, the process by which Stiles — later popularly heralded as "the discoverer of hookworm disease" — came to his conclusions was quite indirect. A zoologist by training, he was hired by the U.S. Department of Agriculture as part of its long-established tradition of supporting agriculture through the dissemination of knowledge and technology. (Public health among *humans* was clearly outside his purview.) His initial impetus to look for helminth infection among humans was a *mistaken* belief, based on faulty research that he had studied in Europe. Believing that the intestinal nematodes that afflicted farm animals could infect human hosts as well (they cannot), he sought and found the tell-tail signs of hookworm infection prevalent in the American South in the late 1890's.

Even if others had stepped forward to publicize the extent of the hookworm problem, they would have nevertheless faced formidable obstacles in taking effective action against the disease. The public-health infrastructure of this period was extremely limited, not just in the Southern states, but throughout the country. This was well before the era of "Big Government."

It was more by chance than by design that Stiles found himself in contact with an associate of John D. Rockefeller, who himself became convinced of the importance of the hookworm problem. Rockefeller was among those whose fortunes were behind this emerging

philanthropy movement. Moreover, as mentioned above, the very availability of funds for such large-scale philanthropy was a recent development, following a string of decades that saw the accumulation of enormous fortunes by the so-called robber barons. Around the turn of the century, several of these millionaires turned their attention from wealth accumulation to wealth *de*cumulation through charitable giving.

The "sudden" introduction of this treatment (broadly defined) combines with the cross-area differences in pre-treatment infection rates to form the identification strategy employed in the present study. The variable of interest, Z_{jt}, is

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

More formally, let $Z_{jt} = H_j^{pre} \times Post_t$ where j indexes the geographic area and t indicates the year. The variable H_j^{pre} denotes the level of hookworm infection among school-aged children in area j at the time of the RSC's initial survey. Since these surveys took place prior to the treatment campaign, we label this number "pre." On the other hand, the variable $Post_t$ is a dummy variable indicating whether year t is later than the active years of the RSC campaign (1910–1915).

By comparing the evolution of outcomes (eg. schooling, fertility) across counties with distinct hookworm-infection rates, we can assess the contribution of the eradication campaign to the observed changes. This comparison can be made easily in the standard difference-in-difference framework. Estimating equation (1) measures the regression-adjusted double difference for some outcome Y_{ijt} for person i in area j at time t.

$$Y_{ijt} = \beta Z_{jt} + \alpha Post_t + \gamma H_j^{pre} + X_{ijt}\Gamma + \varepsilon_{ijt} \tag{1}$$

where Y_{ijt} is the outcome of interest and X_{ijt} is some vector of individual-level controls. Since areas with higher infection rates had "more to gain" from the intervention, we expect there to be greater drops in hookworm infection in these areas. The medical evidence is that infection depresses the energy level of the infected, and so drops in hookworm prevalence should increase investment in schooling. That is, we expect β to be positive.

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 contains substantial support for this assumption. Resurveys summarized by Jacocks (1924) indicated a decrease in hookworm infection of thirty percentage points across the infected areas of the South. Such a dramatic drop in the region's average infection rate can only plausibly have taken place if infection rates fell *more* in highly infected areas than in areas with moderate infection rates. Furthermore we have direct county level evidence on this question for Alabama from Havens and Castle (1930). Figure 4 relates the pre-period infection rates to the subsequent drop of infection rates in affected areas. The assumption that areas where hookworm was highly endemic saw a greater drop in infection than areas with low infection rates is born out across those states and counties for which we have post-intervention data.

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 high levels of infection in 1910, followed by substantial declines in fertility and increases in schooling in the period considered. Panel A of Table I contains summary statistics of various aggregate outcomes. Since county boundaries change during our sample, 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%.

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 first group of columns show the results from the 1910 Census only, whereas the rightmost two columns contain summary statistics from the Censuses of 1900, 1910, 1920, 1940, and 1950. 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 1 year old to females aged 15-49.

Key individual-level variables are summarized in Panel B of Table I. The principal controls are age, race, and gender. 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 results in the remainder of the study employ these variables. The question on schooling refers to enrollment within a certain time span, typically several months, prior to the enumeration date.

The summary statistics already displays the features of the data that will generate the empirical results of the regression analysis discussed below. Under the assumptions that (i) high infection rates reduce the return to investments into children and (ii) the American South in 1910 was undergoing the transition from the Malthusian to balanced-growth periods, we expect the following patterns in schooling and fertility in our sample:

1. Schooling Enrollment Rates are higher in counties with low infection rates during the pre-period;

2. Fertility Levels are lower in counties with low infection rates during the pre-period;

3. Schooling Enrollment Rates and Fertility Rates decline during the sample period;

4. Following the eradication campaign, schooling enrollment rates increase faster in areas with high pre-period infection rates;

5. Fertility Rates decline faster in county groups (SEAs) with high infection rates following the eradication campaign.

All of these predictions are in fact confirmed by the summary statistics reported in table I. Facts (1) and especially (4) allow us to establish the validity of the assumption that high rates of hookworm infection in fact represented an impediment to investing into children. With rational agents we can establish the effect of hookworm disease and its eradication on the returns to schooling by considering its effect on schooling enrollment decisions. To the extent that we can firmly empirically establish the link between the eradication of hookworm disease and the increase in schooling enrollment rates, we also establish 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 childrens' 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.

4.2 Regression Estimates

4.2.1 Schooling

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 (2004), who describes in greater detail the effect of hookworm eradication on schooling and literacy. In the present study, we concentrate on the fertility results and therefore only briefly consider the schooling as an outcome. We refer the reader with a particular interest in the schooling and literacy results to Bleakley (2004).

The empirical results for schooling are presented in Table II, Panel 1. Estimates of the variable of interest, pre-period hookworm $\times$ post, are shown in the first row of coefficients. The first three columns contain regressions with different sets of main effects and controls. Column 4 allows for the controls themselves to interact with Post_t. The interaction with Post_t of the age variable and the black indicator variable are both significantly different from zero, but the inclusions of these interactions do not substantially affect the coefficient on Z_{jt}: areas with higher levels of hookworm saw greater increases in schooling following the anti-hookworm intervention.

We find a substantial additional increase in schooling attendance during 1910-1920 among children living in areas that had high levels of hookworm infection in 1910. This is true in absolute terms and also relative to areas with lower levels of infection. Specifically, the coefficient on Z_{jt} in table II.1, column 1 implies that a county with a 1910 infection rate of 50% would, relative to a county with a 0% infection rate, experience an additional increase in schooling enrollment of five percent points. In 1910 the standard deviation of school enrollment across SEA's was 0.11 and that of hookworm infection rates was 0.23. These numbers imply that a one-standard-deviation increase in lagged hookworm is associated with a post-RSC increase in schooling enrollment as large as 1/4 of a standard deviation in school enrollment. Relative to the mean increase in school enrollment in the entire sample between 1910 and 1920 (see table I) the predicted increase in enrollment due to a standard deviation in hookworm infection rates is about 1/3.

The estimates presented above imply plausible quantitative effects of hookworm infection on schooling enrollment. We can 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).[11] Dividing the first

[11]The latter number comes from the follow-up evidence shown in Figure 6.

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 significance of these findings in the present context is that they allow us to determine the sign of the effect of hookworm eradication on the returns to education.[12] The sizable excess increase in education observed 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 number of children aged 5 or lower per woman declines substantially faster in SEAs with high rates of hookworm infection in 1910. The estimates imply that a county with an infection rate of 50% would see an extra decline of about 0.03-0.04 in the number of children younger than 5 per fertile woman. This compares with an estimated decline of between 0.04-0.05 for a county with no hookworm disease burden in 1910. It is possible to roughly translate these numbers into total fertility rates, a more familiar measure of fertility. Then the mean total fertility rate in the sample in 1910 was approximately 4 and between 1910 and 1920 the decline in fertility rates in an SEA with no hookworm infestation amounted to about 0.05. The decline in the total fertility rate between 1910 and 1920 in an SEA with a 1910 infection rate of 50% amounted to approximately 0.08-0.09. The direct effects of the hookworm eradication campaign on fertility are therefore substantial.

The finding that the fertility decline in areas with high hookworm infection rates significantly exceeded the average decline during the 1910-1920 period is robust to controlling for different sets of controls and also for interacting these controls with Post_t. Table III documents the relation between fertility rates and 1910 infection rate for the entire period

[12] 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.

1900-1950. Columns (1)-(4) show that the findings from table II, panel 2 also hold for this time-period when including various sets of controls for age, race and time. Columns (5) and (6) use the information contained in the RSC on the intensity of efforts to eradicate hookworm disease as independent variables. The conclusions are consistent with those reported above. Those areas that were subjected to large eradication efforts experienced greater declines in fertility.

Figure 5 displays the coefficients on the hookworm infection rate of a regression interacting the hookworm infection rate with year effects[13] (see also table III, column 2). We can see that both before and after the eradication campaign fertility was higher in areas with large infection rates in 1910. However this positive relation dropped significantly in size following the eradication campaign. 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.

4.3 Alternative Explanations

In this subsection, we evaluate a number of alternative explanations for the empirical pattern described above. The first concern relates to infant and child mortality. The empirical evidence does not support the assertion that the excess decline in fertility rates in highly infected areas is related to greater declines in child mortality in these areas. Furthermore theory suggests that any additional decline in child mortality associated with the eradication campaign should have increased fertility, absent the effects of the quantity-quality trade-off. The next concern relates to the possible presence of omitted variable bias. We include additional controls, but this does not affect the coefficient on the infection rate in a statistically measurable manner, nor does it induce changes that are economically important. We also consider the possibility that the eradication of hookworm disease might have affected the opportunities of young women directly and therefore led to the excess reduction in fertility. The

[13]The 1930 microdata is not available at the time of writing.

decline in fertility rates is, however, observed mainly at ages that are not directly affected by hookworm disease. Thus, we conclude that direct effect of the eradication of hookworm disease can not have generated the observed effects. Finally, as a falsification exercise, we estimate equation 1 for the labor-market outcomes of adults, and find no evidence of omitted income or sectoral shocks.

4.3.1 Child mortality

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 the 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 highly prevalent in 1910. Is it theoretically plausible that this will lead to an excess decline in the number of surviving children as documented in table II and III for reasons other than a trade-off between the quantity and quality of 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 the results documented in Tables II and III.

Sah (1991) and Kalemli-Ozcan (2002) 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 extremely risk averse and that they have to be unable to replace dying children to generate a positive relation between net fertility and child mortality. If parents can respond to the death of a child by increased subsequent fertility, then not even very high degrees of risk aversion will lead to this result.

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. Table IV shows that there is no statistically significant relation between child mortality and local hookworm infection rates in the pre-period. The point estimates reported in Table IV imply that one standard deviation in the hookworm infection rate corresponds to a change in the child mortality rate of between 1/4 and 1/3 of a percentage point. This compares with a mean child mortality rate in the sample of 11.42%. Thus the variation in the child mortality rate with the hookworm infection rate is statistically insignificant and economically small, even if not irrelevant in the period.[14]

[14]To 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

4.3.2 Pre-eradication differences across areas

The differential drop in fertility following eradication cannot be explained by a variety aggregate-level controls for pre-eradication differences across areas. The estimates reported above already include detailed micro-level controls for time-varying demographic effects, but we might be concerned with area-level shocks. Results from the baseline specification are shown in Panel A of Table V. As above, the regressions include detailed demographic controls and SEA-specific linear time trends in the 1900-50 sample. We also report specifications in which we add county-group-level data from 1910 on demographics, literacy, agriculture, mortality, and fertility. Paralleling the treatment of the pre-eradication infection rate, these additional variables enter the specification as an interaction with Post. None of the resulting estimates of the effect of hookworm are statistically or economically different from the baseline numbers.

4.3.3 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. To rule out this effect as causing our results, we estimate how the pre-period infection rate affects fertility for the entire age-profile of women in the sample including the years 1910 and 1920 only. 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 group (high infection rates)

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.

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 6 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.

4.3.4 Labor-market outcomes of adults

We now consider impacts on the labor-market outcomes of adults, who should not have been directly affected by the RSC, as a falsification exercise. 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 VI, where we see little evidence of omitted labor-market shocks. The first outcome studied is labor-force participation, defined as whether the individual reports an occupation. For both women and younger men, there is no statistically significant evidence of a change in this measure as a result of the reduction in hookworm infection. On the other hand, for men older than fifty years, there is a significantly significant decline in labor-force participation associated with hookworm eradication, consistent with the models of child-to-parent altruism in retirement. Panels B and C contain measures of labor earnings, imputed using the individual's reported occupation. In no instance do we find significant effects of hookworm on these income proxies. Finally, in Panels D and E, we look for evidence of sectoral shifts correlated with hookworm eradication and find none.

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 a decline in fertility rates. And, this decline in fertility rates was largest in areas where the pre-treatment infection rates were highest.

We will now attempt to use this data to quantify the importance of linking the decision on the quality and the quantity of children in explaining the demographic 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 ρ and other factors denoted θ_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 θ_n and θ_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 linkage 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 0.067 (from a base of 0.56 in 1910). The school enrollment rate for ages 8-16 during this time period rose by 8.68 percentage points. Thus fertility per percentage-point of additional enrollment declined by 0.0077=0.067/8.68. How does this decline in fertility relative to the increase in education compare with the changes in fertility and education observed in high 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 0.0077 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 ρ_{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 θ_n, θ_r and ρ_{1910} and of course the parameters $\beta_1\gamma_2$ and β_2. Consider next the regression of n_{1920} on ρ_{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

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

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 VII gathers the necessary data to perform the calculations described above. The regression results reported in table II, panel 1 and 2, column 1 provide us with the coefficient on the infection rates in the schooling and fertility rates for 1910 and 1920 respectively. The aggregate changes in fertility and school enrollment are obtained by averaging using the IPUMS data. The predicted difference in fertility rates between areas with a 1 percentage point difference in the 1910 hookworm infection rate declines by about 0.0007 between 1910 and 1920. At the same time the school enrollment rate in high infection areas relative to those in the low infection areas increases by 0.001 per percentage point difference in the hookworm infection rate in 1910. In the aggregate data we see that the fertility measure decreases by about 0.0752 between 1910 and 1920 and school enrollment increases by about 0.0854. The ratio $\frac{b_{n,1910}-b_{n,1920}}{b_{h,1910}-b_{h,1920}}$ then 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 -0.68. 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 0.0068. 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 change in aggregate school enrollment and fertility.

The calculations performed here are naturally imprecise, but they do suggest that the quantity-quality trade-off alone 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 a number of 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 quality of 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 approximately 2/3 of a percentage point (from a basis of about 0.50 children per women aged 15-49).

The relative change in fertility and schooling caused by hookworm eradication are approximately equal to aggregate comovements during the period considered. This correspondence 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.

Barro, Robert and Jong-Wha Lee. (2000). "International Data on Educational Attainment: Updates and Implications", *Working Paper no. 42,* Center for International Development at Harvard University (data-set downloadable at: http://www2.cid.harvard.edu/ciddata/)

Becker, Gary S. (1960). "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries.* Princeton, N.J.: Princeton University Press (for N.B.E.R.).

Becker, Gary S., and Robert J. Barro. (1988). "A Reformulation of the Economic Theory of Fertility" *Quarterly Journal of Economics* 53(1): 1-25.

Becker, Gary S., Murphy, Kevin M., and Robert Tamura, (1990). "Human Capital, Fertility and Economic Growth" *The Journal of Political Economy,* Vol. 98, 5.2 S12-S37, 1990.

Bleakley, C. Hoyt (2004). "Disease and Development: Evidence from Hookworm Eradication in the American South." Unpublished manuscript, UCSD. August. Available at http://www.econ.ucsd.edu/ bleakley.

Boldrin, Michele and Larry E. Jones (2002), "Mortality, Fertility, and Saving in a Malthusian Economy", *Review of Economic Dynamics* 5, 775-814.

Brinkley, G.L. (1997). "The Decline in Southern Agricultural Output, 1860-1880." *The Journal of Economic History.* March 57 (1), 116-138.

Brown, John C. and Timothy W. Guinnance (2002), "Fertility Transition in a Rural, Catholic Population: Bavaria, 1880-1910," Population Studies, 56 (2002), 35-49.

Caldwell, John C. (1978). "A Theory of Fertility: From High Plateau to Destabilization", *Population and Development Review*, Vol. 4, No. 4, 553-577.

Carlsson, Gősta (1966). "The decline of fertility: innovation or adjustment process", *Population Studies* 20(2): 149-174.

Coale, Ansley and Susan C. Watkins (eds.) (1986), *The Decline of Fertility in Europe.* Princeton: Princeton University Press.

Doepke, Matthias (2000). "Growth and Fertility in the Long Run." Unpublished manuscript available at http://www.econ.ucla.edu/doepke/research/index.html.

——— (2003). "Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Facts?" Unpublished manuscript available at http://www.econ.ucla.edu/doepke/research

Easrterlin, Richard A. (1976). "The economics and sociology of fertility: a synthesis," in Charles Tilly (ed.) *Historical Studies of Changing Fertility.* Princeton: Princeton University Press, pp. 57-134.

Ehrlich, Isaac and Francis T. Lui (1991). "Intergenerational Trade, Longevity, and Economic Growth", *The Journal of Political Economy*, Vol.99, No. 5, 1029-1059.

Ettling, J. (1981). *The Germ of Laziness; Rockefeller Philanthropy and Public Health in the New South*, Cambridge: Harvard University Press.

Farmer, Jr., H. F. (1970). "The Hookworm Eradication Program in the South, 1909-1925." PhD dissertation, University of Georgia.

Galor, Oded, and David N. Weil. (2000). "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review* 86(3):375-87.

Havens, L.C. and R. Castles (1930). "The Evaluation of the Hookworm Problem of Alabama by Counties." *Journal of Preventive Medicine.* 4(2), 109-114.

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

Jacocks, W. P. (1924). "Hookworm Infection Rates in Eleven Southern States as Revealed by Resurveys in 1920-1923." *Journal of the American Medical Association.* May **82** (20), 1601–2.

Kalmli-Ozcan, Sebnem (2002). "A Stochastic Model of Mortality, Fertility and Human Capital Investment." Forthcoming, *Journal of Development Economics.*

Malthus, Thomas R. (1798) *An essay on the Principle of Population, as it Affects the Future Improvement of Society with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers.* London, printed for J. Johns, in St. Paul's Church-Yard. available at http://www.ac.wwu.edu/~stephan/malthus/malthus.0.html

Meltzer, David (1992). "Mortality decline, the demographic transition, and economic growth", Doctoral Thesis, University of Chicago, Dept. of Economics, December 1992.

Mincer, Jacob (1963). "Market Prices, Opportunity Costs, and Income Effects." In *Measurement in Economics: Studies in Mathematical Economics and Econometrics*, by Carl F. Christ et al., Stanford, Calif: Stanford Univ. Press.

Qian, Nancy (2004). "Quantity-Quality and the One Child Policy: The Positive Effect of Family Size on School Enrollment in China." Working Paper available at http://econ-www.mit.edu/graduate/candidates/download_res.php?id=130

Rockefeller Sanitary Commission (1910-1915). *Annual Report*, New York, N.Y.: Rockefeller Foundation.

Rosenzweig, Mark R. and Kenneth I. Wolpin (1980). "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment", *Econometrica*, Vol. 48, Issue 1, 227-240.

Ruggles, S. and M. Sobek (1997). *Integrated Public Use Microdata Series: Version 2.0*, Minneapolis, Minn.: Historical Census Projects, Univ. of Minnesota. Machine readable dataset available at http://www.ipums.umn.edu.

Sah, Raaj K. (1991) "The Effects of Child Mortality Changes on Fertility Choice and Parental Welfare." *Journal of Political Economy* 99 (3): 582-606.

Schultz, Paul T. (1985). "Changing World Prices, Women's Wages, and the Fertility Transition: Sweden, 1860-1910", *The Journal of Political Economy,* Vol. 93, No. 6, 1126-1154.

Smillie, W. G., and Augustine, D. L. (1925). "Intensity of Hookworm Infestion in Alabama", *Journal of the American Medical Assocation*, Dec. 19, pp.1958–63.

Soares, Rodrigo (2002). "Mortality Reductions, Educational Attainment, and Fertility Choice", Working Paper.

U.S. Bureau of the Census (1915). "Report on the Statistics of Agriculture in the United States." in "Thirteenth Decennial Census of the United States, 1910," Vol. VI & VII, Washington: U.S. Government Printing Office.

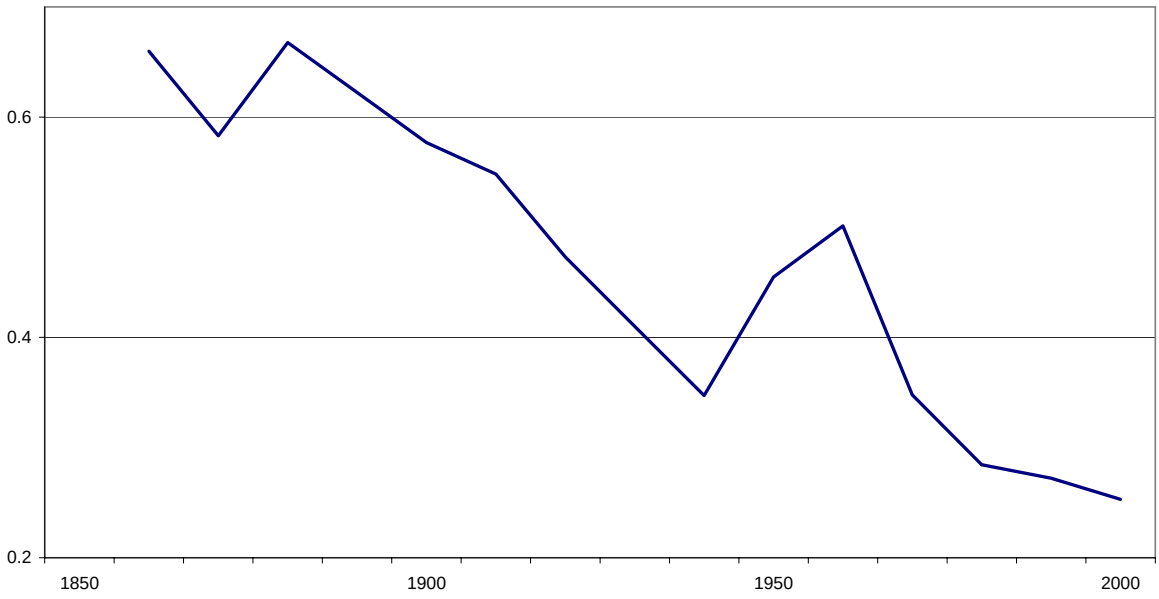
Vance, R. B. (1932). *Human Geography of the South*, Chapel Hill: University of North Carolina Press.

Willis, Robert J. (1982). "The Direction of Intergenerational Transfers and Demographic Transition: The Caldwell Hypothesis Reexamined," *Population and Development Review*, Vol. 8, Supplement: Income Distribution and the Family, 207-234.

Woolhouse, M.E.J. (1998). "Patterns in Parasite Epidemiology: The Peak Shift," *Parasitology Today*, 14:10, pp.428–34.

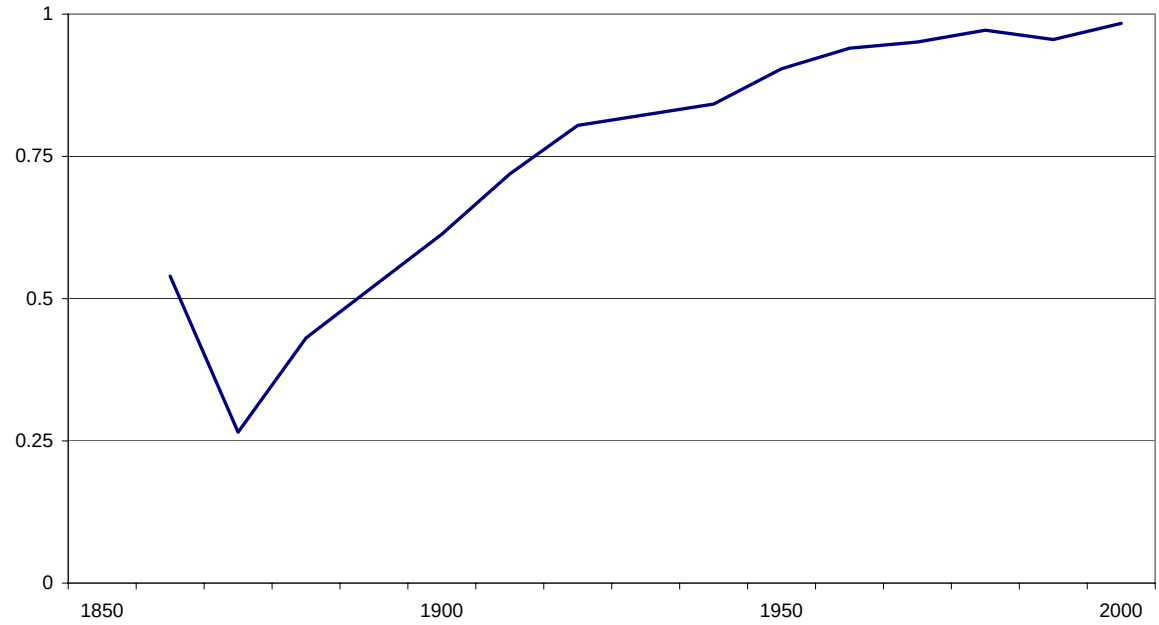
World Bank, 2002. "World Development Indicators." Machine readable dataset available at http://www.worldbank.org/data/wdi2002/

Figure 1: Fertility in the American South



The fertility measure depicted represents the number of children less than 5 years old per woman aged 15-49. The measure is estimated from census extracts provided by the IPUMS.

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



School Enrollment is estimated from census extracts provided by the IPUMS for children aged 14.

Figure 3 The Literature on the Demographic Transition

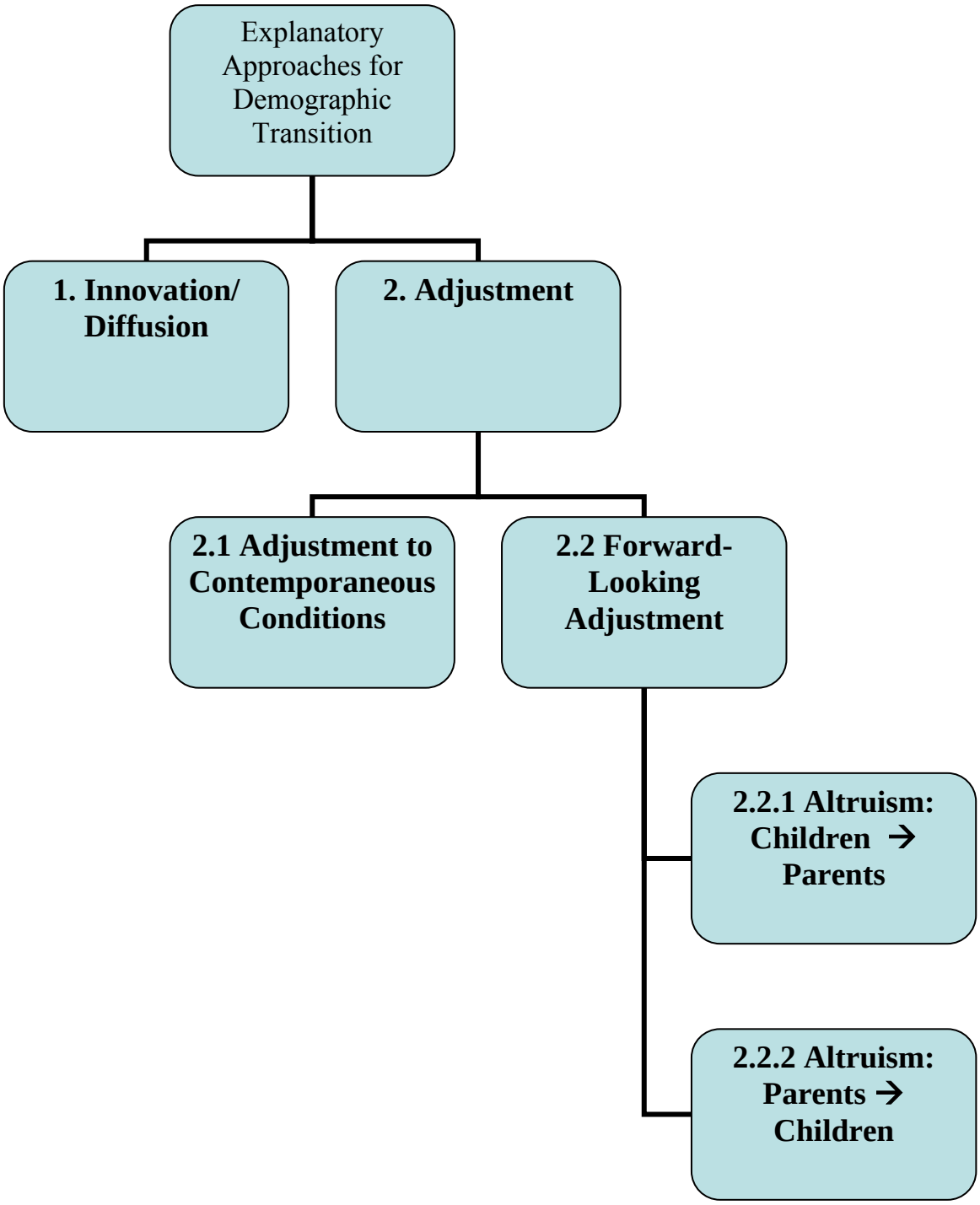
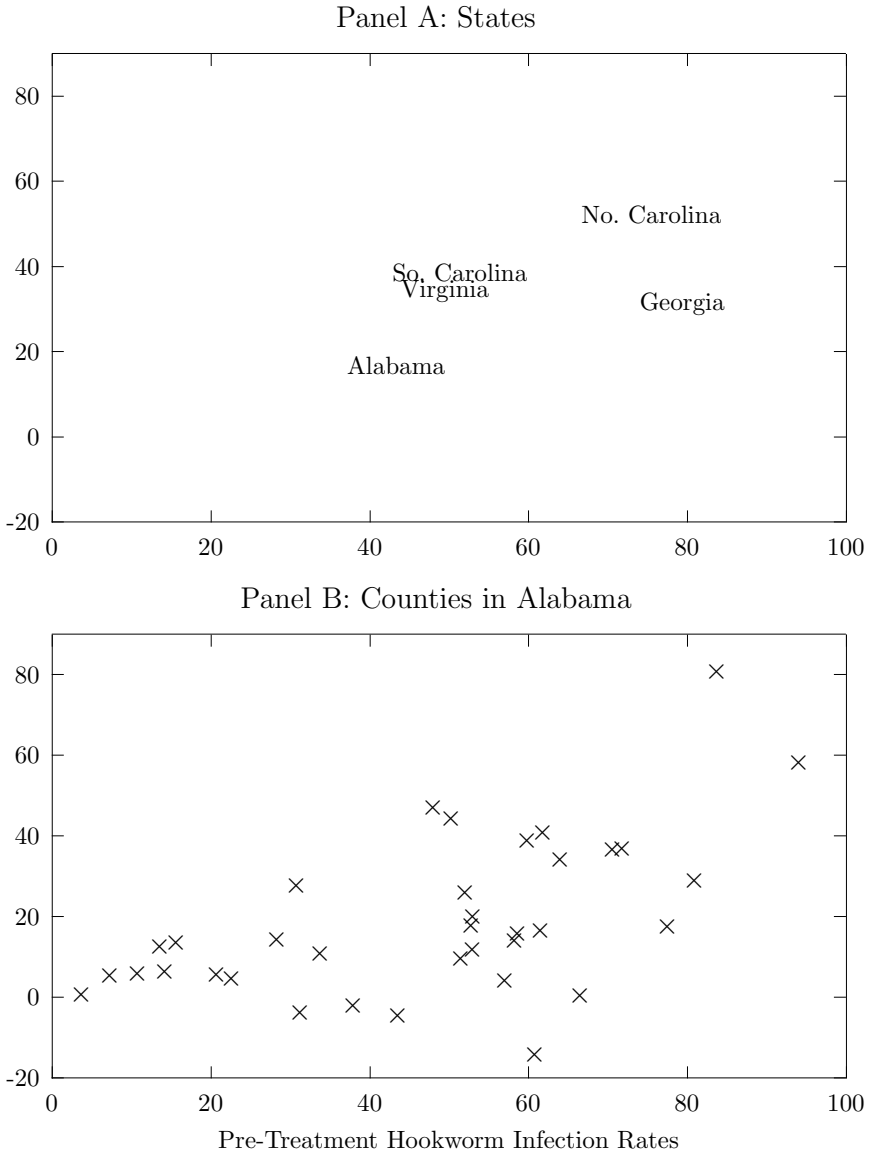
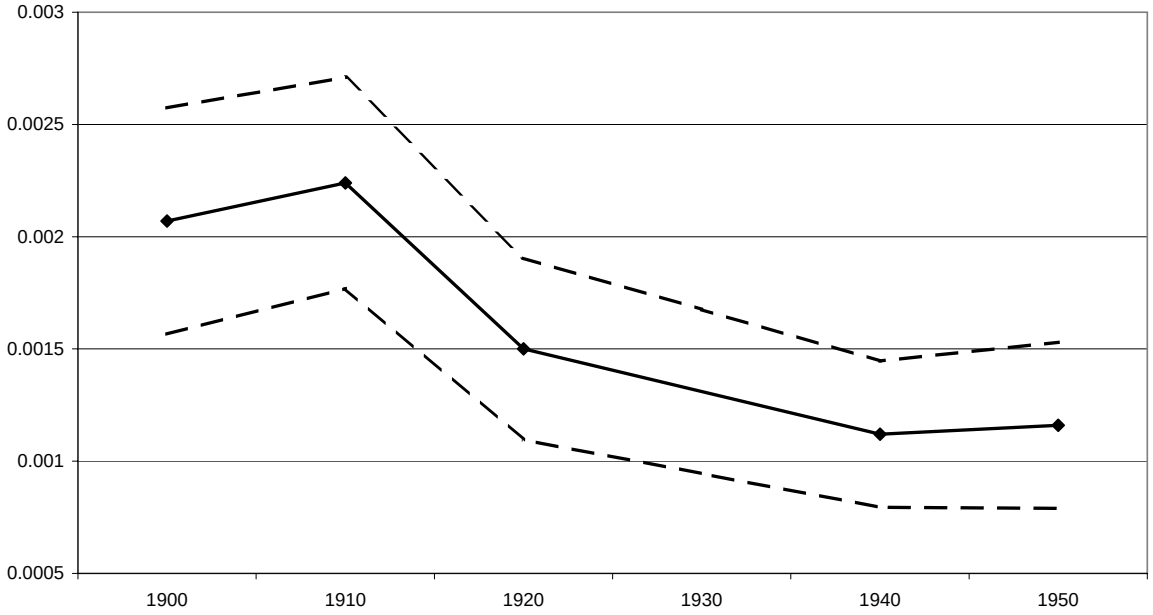


Figure 4: Highly Infected Areas Saw Greater Declines in Hookworm



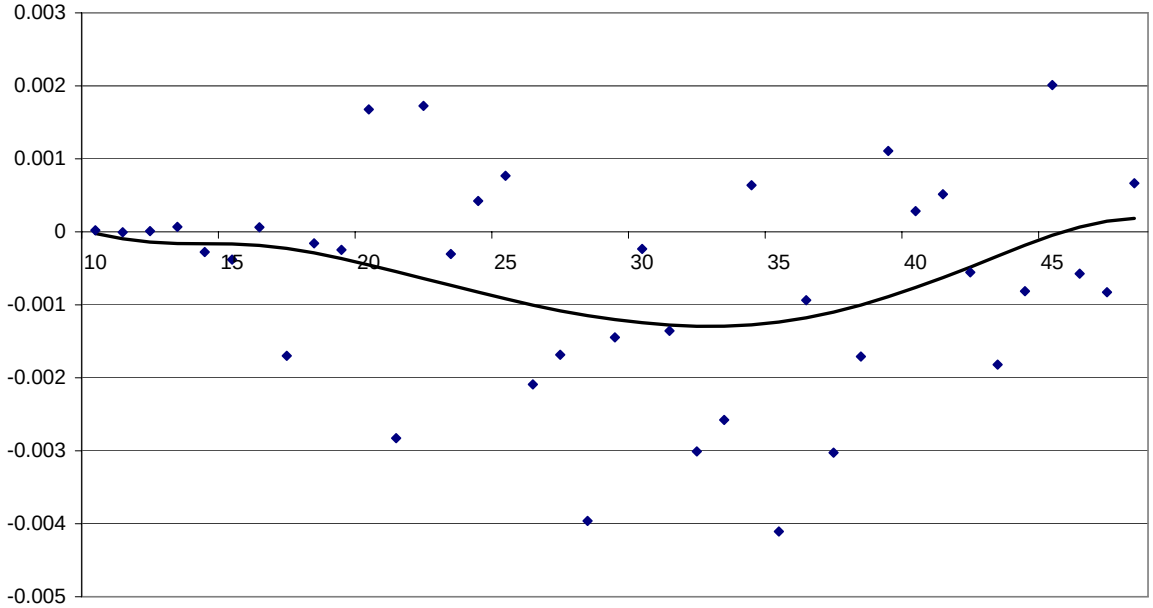
Notes: The y axis displays the decrease in hookworm infection post-intervention, as measured by follow-up surveys. The x axis is the pre-treatment hookworm infection rate, as measured by the Rockefeller Sanitary Commission. Panel A displays data at the state level, as reported by Jacocks (1924). Panel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

Figure 5: 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: 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)				
	Whole Sample	**By Hookworm Infection**		**Source**
		> 40%	<40%	
Hookworm-Infection Rate	0.320	0.55	0.16	
	(0.230)	(0.137)	(0.117)	RSC Annual Reports
Treatments Issued by the RSC, Per Capita	0.06	0.1	0.03	
	(0.067)	(0.069)	(0.052)	RSC Annual Reports
Individuals Treated at Least Once, Per Capita	0.03	0.05	0.02	
	(0.029)	(0.027)	(0.022)	RSC Annual Reports
School Attendance, 1910	0.72	0.71	0.73	IPUMS, author's
	(0.104)	(0.099)	(0.108)	calculations
Change in School Attendance, 1910-20	0.09	0.1	0.08	IPUMS, author's
	(0.080)	(0.090)	(0.072)	calculations
Generalized Fertility Rate	0.75	0.79	0.72	IPUMS, author's
	(0.162)	(0.152)	-0.17	calculations
Change in Generalized Fertility Rate	-0.060	-0.073	-0.050	IPUMS, author's
	(0.121)	(0.125)	(0.117)	calculations
Sample Size	115	48	67	

Panel B: Micro Data				
	1910 Census		**1910-1950 Census**	
Sample	**Children**	**Adult Females**	**Children**	**Female Adults**
Age	11.9	28.2	11.9	29.3
	(2.6)	(9.49)	(2.6)	(9.84)
Black	0.37	0.36	0.33	0.32
	(0.482)	(0.481)	(0.470)	(0.465)
Children less than 5		0.56		0.51
		(0.856)		(0.825)
School Attendance	0.71		0.8	
	(0.453)		(0.397)	
Sample Size	17,194	19,776	140,161	265,195

Notes: Displayed are means and (in brackets) standard deviations.

Sample Selection (Children): Native born blacks and whites age 8-16 in IPUMS from RSC surveyed units in the indicated years.

Sample Selection (Adults): All black and white females aged 15-49 in IPUMS from RSC surveyed units in the indicated years.

Variables: 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 for the aggregate level data is the ratio of children less than 1 years old to females aged 15-49.

Table II.1 The Effect of Hookworm Eradication on School Enrollment, 1910-1920

	(1)	(2)	(3)	(4)
Main Variables of Interest				
1910 Infection Rate* Post Period	0.105**	0.103***	0.104***	0.090***
	(0.047)	*(0.023)*	*(0.023)*	*(0.023)*
1910 Infection Rate	-0.06			
	(0.040)			
Post Period	0.053**	0.052***	0.049***	0.028***
	(0.020)	*(0.009)*	*(0.009)*	*(0.011)*
Controls				
Age	Yes	Yes	Yes	Yes
Age^2	No	No	Yes	Yes
Age*Post Period	No	No	No	Yes
Age Squared*Post Period	No	No	No	Yes
Black	Yes	Yes	Yes	Yes
Black*Post Period	No	No	No	Yes
Black*age	No	No	Yes	Yes
Female	Yes	Yes	Yes	Yes
Female*Post period	No	No	No	Yes
SEA Fixed Effects	No	Yes	Yes	Yes
Sample	64,676	64,676	64,676	64,676

Notes: Dependent variable: Binary Indicator for School Enrollment. Robust standard errors in parenthesis (clustering on S.E.A. times post). Single asterisk denote statistical significante at the 90% level of confidence, double 95%, triple 99%. Sample consists of all native-born white and black children in the IPUMS between the ages of 8 and 16 in the RSC-surveyed geographic units for the indicated years. Reporting of constant term suppressed.

Table II.2 The Effect of Hookworm Eradication on Fertility, 1910-1920

	(1)	(2)	(3)	(4)
Main Variables of Interest				
1910 Infection Rate* Post Period	-0.0621	-0.0668***	-0.0837***	-0.0805***
	(0.0566)	*(0.0239)*	*(0.0246)*	*(0.0245)*
1910 Infection Rate	0.2458***			
	(0.0425)			
Post Period	-0.0521**	-0.047***	-0.035***	-0.0015
	(0.0241)	*(0.012)*	*(0.011)*	*(0.0235)*
Controls				
Age	Yes	Yes	Yes	Yes
Age^2	No	No	Yes	Yes
Age*Post Period	No	No	No	Yes
Age Squared*Post Period	No	No	No	Yes
Black	Yes	Yes	Yes	Yes
Black*Post Period	No	No	No	Yes
Black*age	No	No	Yes	Yes
SEA Fixed Effects	No	Yes	Yes	Yes
Sample	73,695	73,695	73,695	73,695

Notes: Dependent variable: Number of own children less than 5 years old living in the household. Robust standard errors in parenthesis (clustering on S.E.A. times post). Single asterisk denote statistical significante at the 90% level of confidence, double 95%, triple 99%. Sample consists of all white and black females in the IPUMS between the ages of 15-49 in the RSC-surveyed geographic units for the indicated years. Reporting of constant term suppressed.

Table III The Effect of Hookworm Eradication on Fertility, 1900-1950

	(1)	(2)	(3)	(4)	(5)	(6)
Main Variables of Interest						
1910 Infection Rate* Post Period	-0.0921**		-0.1145***	-0.1311***		
	(0.0406)		*(0.0396)*	*(0.0281)*		
1910 Infection Rate	0.2154***		0.2350***			
	(0.0345)		*(0.0339)*			
Treatment per capita* Post					-0.4260***	
					(0.1053)	
Individuals treated per capita* Post						-0.8900***
						(0.2290)
Infection Rate* 1900 Effect		*0.2071****				
		(0.0504)				
Infection Rate* 1910 Effect		0.2241***				
		(0.0470)				
Infection Rate* 1920 Effect		*0.1500****				
		(0.0405)				
Infection Rate* 1940 Effect		*0.1124****				
		(0.0326)				
Infection Rate* 1950 Effect		*0.1158****				
		(0.0370)				
Controls						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Age, Race Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
(Age, Race)*Year	No	No	Yes	Yes	Yes	Yes
SEA Fixed Effects	*No*	*No*	*No*	*Yes*	Yes	Yes
Sample	*265,195*	*265,195*	*265,195*	*265,195*	*257,173*	*257,173*
R^2	*0.14*	*0.14*	*0.15*	*0.16*	*0.16*	*0.16*

Notes: Dependent variable: Number of own children less than 5 years old living in household. Robust standard errors in parentheses (clustering on S.E.A. times post). Simple asterisk denotes statistical significance at the 90% level of confindence, double 95%, triple 99%. Sample consists of black and white 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

	(1)	(2)	(3)	(4)	(5)	(6)
Infection*100	0.0141	0.0084	0.0095	0.0026	0.0073	-0.0005
	(0.0106)	*(0.0077)*	*(0.0083)*	*(0.0065)*	*(0.0079)*	*(0.0063)*
Literacy Dummy for Mother					*-0.009****	*-0.009****
					(0.002)	*(0.002)*
Race FE			Yes	Yes	Yes	Yes
Age FE			Yes	Yes	Yes	Yes
Sample Years	1910	1900, 1910	1910	1900, 1910	1910	1900, 1910
Sample Years	13,407	27,253	13,407	27,253	13,407	27,253
R^2	0	0	0.07	0.07	0.08	0.08

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 (2), (4) and (6) the race and age effects are fully interacted with year effects. Simple asterisk denotes statistical significance at the 90% level of confindence, double 95%, triple 99%. Sample consists of black and white females in the IPUMS between the ages of 15 and 49 in the RSC-surveyed geographic units for the indicated years.

Table V, Sensitivity Analyses

	1910-20	1900-50
Interact 1910 SEA Aggregates with Post:		
	Panel A: Basic Specification	
Hookworm Infection Rate	-0.076***	-.051***
	(.0159)	*(.019)*
	Panel B: Race and Literacy	
Hookworm Infection Rate	-.091***	-.063***
	(.017)	*(.019)*
Fraction Black	-0.02	-.110***
	(.030)	*(.029)*
Fraction Literate	-.114**	-.162***
	(.046)	*(.040)*
	Panel C: Agricultural Variables	
Hookworm Infection Rate	-0.074***	-.050***
	(.0164)	*(.017)*
Fraction Living on a Farm	-0.011	-.087***
	(.021)	*(.024)*
Cotton Output per Capita	-0.012***	-.020***
	(.004)	*(.004)*
Tobacco Output per Capita	-0.069**	-.030
	(.027)	*(.038)*
	Panel D: Child Mortality	
Hookworm Infection Rate	-0.077***	-.051***
	(.0155)	*(.019)*
Child Mortality, 1900-10	0.266***	-.016
	(.087)	*(.091)*
	Panel E: Pre-Eradication Fertility	
Hookworm Infection Rate	-0.032**	-.058***
	(.016)	*(.020)*
Fraction with Children < 5 years old	-0.259***	0.041
	(.053)	*(.065)*
Sample	*44,095*	*165,324*

Notes: Dependent variable: Number of own children less than 5 years old living in the household. Robust standard errors in parenthesis (clustering on S.E.A. times post). Single asterisk denote statistical significance at the 90% level of confidence, double 95%, triple 99%. Sample consists of all white and black females in the IPUMS between the ages of 25 and 45 in the RSC-surveyed geographic units for the indicated years. All specifications include dummies for SEA and for age x black x Census region x year. Regressions for the 1900-50 sample include SEA-specific linear trends.

Table VI Estimated Effects of Hookworm Eradication on Adult Occupational Outcomes

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Samples:							
Females:				Males:			
All Ages	Ages [25,35)	Ages [35,50)	Ages [50,55]	All Ages	Ages [25,35)	Ages [35,50)	Ages [50,55]
Parameter estimates:							
Panel A: Labor-Force Participation							
-0.0056	0.0207	-0.0264	-0.0382	-0.0069	0.0051	-0.0048	-0.0492 **
(0.0284)	(0.0349)	(0.0331)	(0.0501)	(0.0065)	(0.0084)	(0.0094)	(0.0234)
{48240}	{21052}	{21758}	{5430}	{48947}	{19749}	{22808}	{6390}
[0.3238]	[0.3277]	[0.3237]	[0.3091]	[0.9790]	[0.9766]	[0.9824]	[0.9760]
Panel B: Occupational Income Score							
0.0581	0.2195	-0.1159	-0.0626	-0.0186	0.9070	-0.6510	-0.1192
(0.4163)	(0.5295)	(0.5058)	(0.7358)	(0.4912)	(0.5863)	(0.6794)	(1.0823)
{48240}	{21052}	{21758}	{5430}	{48947}	{19749}	{22808}	{6390}
[3.3444]	[3.4879]	[3.3236]	[2.8716]	[18.5736]	[18.2216]	[19.1484]	[17.8854]
Panel C: Duncan Socioeconomic Index							
-0.3499	-0.2522	-0.3388	-1.0809	-0.4529	1.1373	-1.5683	-0.7018
(0.6185)	(0.8407)	(0.7316)	(1.0012)	(0.7931)	(0.8798)	(1.1703)	(1.8652)
{48240}	{21052}	{21758}	{5430}	{48947}	{19749}	{22808}	{6390}
[4.3239]	[4.5981]	[4.2919]	[3.3925]	[19.8944]	[18.9747]	[21.0538]	[19.1638]
Panel D: Works in Agriculture							
-0.0360	-0.0206	-0.0460	-0.0318	-0.0416	-0.0769	-0.0002	-0.1030
(0.0342)	(0.0417)	(0.0401)	(0.0590)	(0.0382)	(0.0489)	(0.0492)	(0.0665)
{48240}	{21052}	{21758}	{5430}	{48947}	{19749}	{22808}	{6390}
[0.1844]	[0.1830]	[0.1866]	[0.1820]	[0.6043]	[0.5783]	[0.6027]	[0.6909]
Panel E: Works as Craftsman or Operative							
-0.0022	-0.0013	-0.0037	-0.0055	0.0033	0.0508	-0.0425	-0.0138
(0.0092)	(0.0147)	(0.0117)	(0.0164)	(0.0265)	(0.0377)	(0.0339)	(0.0424)
{48240}	{21052}	{21758}	{5430}	{48947}	{19749}	{22808}	{6390}
[0.0217]	[0.0243]	[0.0216]	[0.0126]	[0.1335]	[0.1404]	[0.1389]	[0.0950]

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 information. Robust standard errors in parenthesis (clustering on S.E.A. times post). 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 VII How important is the Quantity-Quality Mechanism in explaining the Demographic Transition?

	1910	1920	Difference
Effect of Infection on Schooling[1]	-0.0570	0.0480	0.1050
Effect of Infection on Fertility[2]	0.2458	0.1737	-0.0721
Aggregate School Enrollment (age 8-16)	0.7190	0.8044	0.0854
Aggregate Fertility (age 15-49)	0.5483	0.4731	-0.0752

Fertility decline per Percentage Point Schooling implicit in Hookworm Eradication Data: **-0.69**

Fertility decline per Percentage Point increase in Schooling (aggregate 1910-1920): **-0.88**

1.) See Table II, Panel 1, column 1

2.) See Table II, Panel 2, column 1

Data Appendix

The empirical component of the study is an analysis of sequential cross sections (SCS) from different points in time. That is, we compare a particular age group in one Census to that same age group in a later Census. This is coupled with a decomposition of the changes over time by area, on the basis of each area's pre-treatment-campaign infection level. These two comparisons constitute the difference-in-difference strategy. In this appendix, we discuss the micro data employed in the SCS analysis. We later describe the construction of the aggregate data for hookworm and the additional control variables that factor into analysis.

A Sources and Definitions for the Micro Data

The micro data for the SCS component are samples drawn from the Censuses of 1900, 1910, 1920, 1940, and 1950, accessed through the IPUMS project (Ruggles and Sobek (1997)). 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 attendance.** 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.

  See http://www.ipums.umn.edu/usa/peducation/schoola.html for more detail.

- **Fertility.** Defined as number of own children less than five years of age living in the household.

- **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. From 1940 on, the question corresponds more closely to the modern BLS definition.

- **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, such as in Section 4. 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, I 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.)
- **Fertility, 1910.** The fertility rates for 1910 are measured using the IPUMS samples. We construct SEA rates, by average within our sample of native white and black females.

- **Literacy rates, 1910.** These data were compiled at the county level and come from the 1910 Census. Child literacy refers to ages 10–20 and is constructed as follows: $1 - (v50/v49)$, using the variable codes of the ICPSR study. Adult literacy refers to males of voting age, defined as $1 - (v37/v26)$. To construct SEA rates, we sum the components over the constituent counties and apply the above formulae. (ICPSR Study #3.)

- **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 $(v24 + v25)/(v24 + v25 + v22 + v23)$, 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.)

- **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.)

- **Child Mortality, 1900–10.** Defined from IPUMS samples using the Preston-Haines methodology, and aggregate to SEA level.