NBER WORKING PAPER SERIES

MOSQUITOES: THE LONG-TERM EFFECTS OF MALARIA ERADICATION IN INDIA

David Cutler
Winnie Fung
Michael Kremer
Monica Singhal

Working Paper 13539
http://www.nber.org/papers/w13539

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2007

We thank the National Malaria Eradication Programme of India for its generous assistance. We are grateful to participants of the Center for International Development Conference on Health Improvements for Economic Growth for their comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Mosquitoes: The Long-term Effects of Malaria Eradication in India
David Cutler, Winnie Fung, Michael Kremer, and Monica Singhal
NBER Working Paper No. 13539
October 2007
JEL No. H51,I18,J24

ABSTRACT

We examine the effects of malaria on educational attainment by exploiting geographic variation in malaria prevalence in India prior to a nationwide eradication program in the 1950s. Malaria eradication resulted in gains in literacy and primary school completion rates of approximately 12 percentage points. These estimates imply that the eradication of malaria can explain about half of the gains in these measures of educational attainment between the pre- and post-eradication periods in areas where malaria was prevalent. The effects are not present in urban areas, where malaria was not considered to be a problem in the pre-eradication period. The results cannot be explained by convergence across areas. We find gains for both men and women as well as for members of scheduled castes and tribes, a traditionally disadvantaged group.¸

David Cutler
Department of Economics
Harvard University
1875 Cambridge Street
Cambridge, MA 02138
and NBER
dcutler@harvard.edu

Michael Kremer
Harvard University
Department of Economics
Littauer Center M20
Cambridge, MA 02138
and NBER
mkremer@fas.harvard.edu

Winnie Fung
Department of Economics
Harvard University
Littauer Center
1805 Cambridge Street
Cambridge, MA 02138
fung@fas.harvard.edu

Monica Singhal
Harvard University
JFK School of Government
79 JFK Street
Cambridge, MA 02138
and NBER
monica_singhal@harvard.edu

1 Introduction

Malaria is one of the three infectious diseases responsible for the most deaths in human history.[1] Today, malaria is endemic in over 100 countries and affects 40% of the world's population. The World Health Organization (2001) estimates that there are 300 million malaria cases and almost one million deaths from malaria in the world each year. Faced with this huge global burden, international organizations have redoubled their efforts to combat the disease. The United Nations has made combating malaria one of its Millennium Development Goals. The Roll Back Malaria Global Partnership, formed by the WHO, United Nations Children's Fund, United Nations Development Program, and the World Bank, aims to halve the malaria burden by 2010.

Many argue that improving health, while important in and of itself, can also lead to higher economic growth and development. Gallup and Sachs (2001) use cross-country growth regressions to argue that countries with intensive malaria grew 1.3% less per year and that a 10% reduction in malaria was correlated with 0.3% higher growth. Other macroeconomic studies, such as those by Bloom and Canning (2005) and Alleyne and Cohen (2002), also conclude that improvements in health can lead to higher economic growth. On the other hand, Acemoglu and Johnson (2005) argue that the wave of international health innovations that began in the 1940s did not lead to a disproportionate increase in log per capita GDP or average levels of education in the areas with high pre-intervention disease burden.

In this paper, we examine one of the mechanisms through which health may have a causal impact on development by studying the effects of malaria eradication on human capital accumulation. We use the national malaria eradication program in India in the 1950s as a quasi-experiment and exploit geographic variation in malaria prevalence prior to the eradication campaign. We compare educational gains for cohorts born before and after the program in areas with varying pre-eradication prevalence. We find that the eradication program had a large, positive, and significant effect on

[1] The other two diseases are pneumonia and tuberculosis (Kiple, 1993 and Hoff et al., 2000).

literacy and primary school completion. The effects are smaller but still significant for middle and secondary school completion. We do not see effects of the eradication program in urban areas, where malaria was not considered to be a problem in the pre-eradication period. These findings cannot be explained by simple convergence in educational outcomes across areas. The results are robust to allowing for differential time trends across areas with high and low initial educational attainment and are also robust to a variety of specification tests. We find that the program resulted in gains for both men and women, and the gains for members of scheduled castes and tribes are similar to gains for the rest of the population.[2]

This study relates most closely to two recent papers. Bleakley (2007) studies the effect of malaria eradication campaigns on the income and education of native males in the United States, Brazil, Colombia, and Mexico. Using malaria mortality rates and an ecology index to identify pre-eradication disease prevalence, Bleakley finds that childhood exposure to malaria lowers labor productivity and leads to lower adult income. Results for years of schooling are mixed. Lucas (2005) studies married women in Paraguay, Sri Lanka, and Trinidad in the 1940s to 1960s and finds that malaria eradication leads to increases in female education and literacy rates.[3]

Two other studies use weather conditions to instrument for malaria exposure in the United States and examine the effects on long run health and economic outcomes. Hong (2007) finds that malarial risk leads to adverse long run health outcomes, lower labor force participation, and lower wealth. Barecca (2007) finds that in utero malaria exposure leads to lower educational attainment.

Our results expand on this literature in several ways. We distinguish explicitly between malaria that is endemic and malaria that is epidemic. Malaria is endemic to an area if transmission is stable and the disease is continually present. Epidemic malaria has unstable transmission and may occur with substantial variation over time.[4] The measures of malaria prevalence used by Bleakley (2007) and Lucas (2005) may capture both types of malaria. Hong (2007) and Barecca (2007) focus on variations in malaria that are driven by fluctuations in weather factors; these estimates may therefore be capturing primarily the effects of epidemic malaria. We find that the eradication

[2] Members of scheduled castes and tribes have traditionally been economically and socially disadvantaged in India.

[3] Lucas (2007) finds that malaria eradication in Sri Lanka led to an initial increase in fertility followed by lowered fertility in the second generation.

[4] We discuss the potential consequences of both types of malaria for individual health and educational outcomes in Section 2.1.

program resulted in significant gains in educational attainment in both types of areas, although the gains are slightly higher in areas where malaria was endemic prior to the program. The variation in pre-eradication prevalence is at the district level, allowing us to exploit quite localized variation in malaria exposure. We also test for heterogeneity in treatment effects, to see whether traditionally disadvantaged groups also benefited from the program.

The paper proceeds as follows. Section 2 provides an overview of the epidemiology of malaria and the mechanisms through which malaria may affect educational attainment, describes malaria in India in the pre-eradication era, and discusses the National Malaria Control Program in India and its effectiveness. Section 3 outlines our empirical strategy. Section 4 describes the data. Section 5 presents our results, and Section 6 concludes.

2 Malaria in India

2.1 Overview of Malaria Epidemiology

> "The problem of existence in very many parts of India is the problem of malaria. There is no aspect of life in this country which is not affected either directly or indirectly by this disease. It constitutes one of the most important causes of economic misfortune, engendering poverty, diminishing quantity and quality of food supply, lowering the physical and intellectual standards of the nation and hampering increased prosperity in every way." - *John Sinton, Director of the Malaria Survey of India, 1936*

Malaria is a protozoal infection transmitted to human beings by mosquitoes. The classic symptom of malaria is bouts of fever with spikes on alternating days. Headaches, malaise, fatigue, nausea, and anemia are also common. Severe forms of the disease can result in organ failure, delirium, impaired consciousness, and generalized convulsions, followed by persistent coma and death.[5] Infants and children suffer the worst from malaria and high morbidity and mortality rates. The other high risk group is pregnant women, for whom malaria often results in pre-term labor or low birth weight full-term births, as well as spontaneous abortions and still-births.

Malaria has been hypothesized to have lifelong effects on cognitive development and educational attainment through at least three channels: chronic malaria-induced anemia, time lost or wasted

[5] WHO, http://www.who.int/malaria/faq.html.

in the classroom due to illness, and low birth weight.[6] Through the first channel, epidemiological studies have associated malaria with anemia, epileptic convulsions, and growth faltering during the first three years of life (Shiff et al., 1996). These in turn can cause learning disabilities and affect children's cognitive development (Boyle et al., 1994). A randomized evaluation of treatment for malaria among school children in Kenya found resulting reductions in anemia and increases in sustained attention (Clarke et al., 2007).

Regarding the second channel, malaria can also affect school attendance directly through illness. In Kenya, primary school students have been estimated to have on average four episodes of malaria per year and to miss five school days per episode, amounting to 20 school days missed per child per year (Leighton and Foster, 1993). Malaria can explain up to 8% of all school absenteeism and up to 50% of absenteeism due to preventable medical causes (Brooker et al., 2000). A study in The Gambia shows that the use of insecticide-treated mosquito nets, which presumably reduces the threat of malaria, lowers school absenteeism due to fever (Aikins, 1995).

Lastly, malaria in pregnancy can cause low birth weight because of fetal growth retardation or premature delivery (Duffy and Desowitz, 2001). This can in turn reduce the physical, cognitive, and neurosensory development of the child, resulting in lower human capital accumulation (McCormick et al., 1992).

In areas where malaria is endemic, individuals can acquire immunity over time through years of continued exposure and multiple infections. The effects of malaria are therefore most pronounced in childhood and youth, when individuals have not acquired immunity. Immunity may also be reduced during pregnancy. In areas where malaria is epidemic, individuals may have little or no acquired immunity. In these areas, malaria can affect both children and adults and can result in severe adverse health consequences.

2.2 The Pre-Eradication Era

References to malaria can be found in Vedic writings dating to 1600 B.C. (Desowitz, 1991) and two classical books on Ayurveda describe malaria as the "king of diseases" (Rao, 1959). Efforts to control malaria date back to the early 1900s but were revolutionized in the mid-1940s with the

[6] Cerebral malaria may also result in long-term adverse effects on brain function (Holding and Snow, 2001).

advent of DDT (dichlorodiphenyl trichloroethylene).[7] DDT was effective, non-toxic to humans, and "dirt-cheap to manufacture" (Desowitz, 1991). Aggressive campaigns using DDT were launched almost simultaneously around the world, leading to the rapid eradication of malaria in Taiwan, much of the Carribean, the Balkans, parts of northern Africa, northern Australia, and large parts of the South Pacific (Davis, 1956).

DDT was first used in India by the military in 1944 and became available for civilian anti-malaria operations in 1945. During the late 1940s, a number of pilot programs and trials took place throughout the country. International organizations such as WHO, UNICEF, and the Rockefeller Foundation also sponsored demonstration projects. These pilot projects were very successful, and in 1951, the national Planning Commission endorsed the development of a comprehensive, nationwide malaria control program. In April of 1953, the National Malaria Control Program (NMCP) was launched. Funding for the program was primarily from bilateral and international sources, and program implementation was overseen by the WHO. The timing of the program is plausibly exogenous, since it was driven by the advent of DDT.

Prior to the eradication program, malaria was considered the greatest health problem facing India. Survey evidence estimates that immediately after partition in 1947, India suffered from 75 million cases of malaria (doubled during epidemic years) and 800,000 deaths directly attributable to malaria annually (Sinton, 1935, 1936; Rao, 1959). The population of India in 1947 was 344 million, implying an annual prevalence rate of 22%. The estimate of malaria deaths implies that malaria was responsible for approximately 10% of total annual deaths in the pre-eradication era.

Of the four human malaria parasites (*Plasmodium falciparum, P. vivax, P. malariae,* and *P. ovale*), two are prevalent in India: *P. vivax* and *P. falciparum*.[8] *P. falciparum* is associated with the most severe forms of malaria and accounts for most malaria fatalities. It is the primary cause of malarial infections in Africa, where 90% of malaria deaths currently occur. Data on the relative prevalence of these parasites in India during the pre-eradication era is unfortunately unavailable; data from the immediate post-eradication period suggest that approximately 30% of cases were due to *P. falciparum* (NMEP, 1996). This percentage has increased in recent years.

[7] Early experiments prior to 1910 focused on breeding control. These attempts were generally considered failures. From 1910 to 1944, measures such as drainage and the filling up of breeding places were undertaken. Some limited success was also achieved using larvicidal chemicals such as oil, Paris green, and later pyrethrum.

[8] *P. malariae* also exists, but is confined to tribal areas of the country (NMEP, 1986).

2.3 National Malaria Control Program

The National Malaria Control Program's main operational activity was to spray DDT in human dwellings and cattle sheds. Two rounds of spraying were conducted per year, usually between May and September, the peak transmission times.

A five-year plan was formed with the goal of establishing 125 Malaria Control Units by 1956, each covering a population of one million (NMEP, 1986). The program was successful in achieving this target (*Table 1*). By 1956, 134 units had been established and 112 million people were estimated to be protected. By 1958, almost 200 units had been established and 165 million people were under protection.

The program was so successful that in 1958 it was reformulated as the National Malaria Eradication Program, with the goal of completely eradicating malaria from the nation. This effort was adopted in coordination with a WHO campaign to eradicate malaria from the entire region, launched after the Eighth World Health Assembly in 1955 (WHO, 1967). By 1960-61, the entire country was brought under the program.

Figures 1.1 to 1.4 illustrate the rapid geographic expansion of coverage as districts were phased into the program. Once a district was incorporated into the program, it remained in the program in all subsequent years. The statement of the Planning Commission indicated that priority targeting of areas should be based on endemicity and food producing capacity. The timing of the phase-in for particular districts may therefore not be exogenous.

Large urban areas were relatively free from malaria prior to the eradication era (League of Nations, 1930). In fact, urban malaria was considered to be a negligible problem and malaria control efforts were left to local governments.[9] Prevalence of malaria in urban areas increased only later (although not nearly to the levels of pre-eradication rural malaria), and the government launched an Urban Malaria Scheme in 1971 to address the growing problem of urban malaria. We therefore focus our analysis on the rural sample and later use the urban sample as a placebo group.

[9] The following quote describes the treatment of urban malaria during this time: "As per the plan of operations formulated at the time of launching of the National Malaria Eradication Programme, all the roofed structures in the rural areas received indoor residual insecticidal spray except urban areas with a population of over 40,000. In such urban areas, the indoor residual insecticidal spray was confined only to the peripheral belt to a depth of 1 to 1.5 km. Antilarval measures were recommended in towns and cities. The implementation of antilarval operations was made the responsibility of the local bodies. Due to financial constraints many local bodies failed to implement the control measures. Though malaria epidemics were recorded earlier in Bombay, Delhi, Lucknow, etc., these could be immediately contained. Hence, malaria in urban areas was not considered as a major problem" (NMEP, 1996).

2.4 Program Effectiveness

While the campaign was unsuccessful in eradicating malaria from India, it did achieve tremendous reductions in malaria prevalence. The NMEP, which began tracking malaria prevalence from 1961 using blood smear data, estimates that in 1965 there were 100,000 cases of malaria, compared to 75 million cases in the pre-eradication era (*Table 2*).

Although the 1965 figures are likely to underestimate true malaria prevalence, there is no doubt that there was a dramatic reduction in malaria prevalence over this period. *Table 3* shows three measures of malaria prevalence during the control program: the child spleen rate, child parasite rate, and infant parasite rate. The child spleen rate is the percentage of examined children aged 2-9 with an enlarged spleen.[10] The child and infant parasite rates are the percentage of children aged 2-9 or infants examined in a blood survey who are found positive for a malaria parasite. All three measures show substantial declines over this period. Vital statistics data on causes of death indicate that the number of malaria deaths between 1952 and 1963 dropped by 91.2% and 98.3% in Uttar Pradesh and West Bengal, respectively (*Figure 2*). These states were chosen because they are large states with relatively complete data and their boundaries did not change much during the reorganization of Indian states in 1956.

Malaria prevalence remained low throughout the 1960s but experienced a slight resurgence in the 1970s, peaking in 1976. However, even at the peak of the resurgence, the prevalence rate was only 1.1% (*Table 2*). Reported prevalence decreased again, although not to the low levels seen in the immediate post-eradication period. This may partially be a result of increased accuracy in reported caseloads over time.

3 Empirical Strategy

Our study focuses on the effects of malaria exposure on educational attainment. The basis of our empirical strategy is a differences-in-differences design, exploiting geographic variation in the prevalence of malaria prior to the eradication program.[11] We compare outcomes at a point in time

[10] The spleen rate is a commonly used measure of exposure to malaria.

[11] Although there is variation in when districts were phased into the program during the eradication era, the timing of phase-in may be related to malaria severity and other relevant factors. In addition, measures of malaria prevalence (child spleen rate, child parasite rate, and infant parasite rate) show declines in both sprayed and unsprayed areas over this period, suggesting that even those in unsprayed areas may have benefited from the program (NMCP, 1986).

for individuals in birth cohorts born before and after the eradication era in areas with high and low pre-eradication malaria prevalence. Ideally, we would like to know the district of birth for each individual, but our outcomes data report only the district of current residence. An identifying assumption of our analysis is therefore that district of residence is a good proxy for district of birth. According to the 1991 Census of India, only 7.5% of the rural population are reported to be living in districts other than their district of birth.

Since malaria likely affects cognitive development and educational attainment mainly during infancy and childhood, we focus on the effects of malaria exposure at birth. Malaria prevalence at birth can be seen as an approximation of the individual's malaria exposure during the first few years of life, when the effect of malaria is likely the strongest. We also report results relaxing this assumption.

We run regressions of the following form, for individual i in birth cohort c in district d:

$$Outcome_{icd} = \beta_0 + \beta_1 (Post)_c * (High)_d + \gamma_d + \alpha_c + X\beta_2$$

where $Post$ indicates that the individual's birth cohort is after the eradication era and $High$ indicates whether individual i's district was a high prevalence district prior to eradication. γ and α are district and birth cohort fixed effects, and X is a vector of individual characteristics, including gender, scheduled caste, scheduled tribe, and household religion.

4 Data

4.1 Map of Pre-Eradication Endemicity

A central problem in assessing the impact of malaria is the identification of a suitable indicator for the prevalence of the disease. As Gallup and Sachs (2001) point out, there is a lack of high-quality data on malaria incidence in the most severely affected countries. In their study, they use historical maps of the geographical distribution of malarial risk to derive an index of malaria prevalence.[12]

In this paper, we use a 1948 government map that classifies areas into categories of malaria

Finally, the phase-in of the program was quite rapid, raising further difficulties in exploiting variation in timing of coverage.

[12] Other studies have used malaria mortality rates (Bleakley, 2007), malaria morbidity rates (Lucas, 2005), spleen rates (Lucas, 2005), and malarial fevers (Hong, 2007) as indicators of disease prevalence.

prevalence. The map was obtained from the Ministry of Health and Family Welfare, Government of India. The pre-eradication malaria map classifies areas into six categories of endemicity: (1) areas above 5000 feet; non malarious, (2) known healthy plain areas; spleen rate under 10%, (3) variable endemicity associated with dry tracts; potential epidemic areas, (4) known areas liable to fulminent epidemic diluvial malaria, (5) moderate to high endemicity; fulminent epidemics unknown, and (6) hyperendemicity of jungly hill tracts and terai land. This map was based on spleen rate surveys and climate factors, although the exact mechanism by which category boundaries were constructed is not known. It should be noted that the malaria endemicity variable is not a simple proxy for climatic factors such as rainfall. Rainfall and malaria prevalence have a non-monotonic relationship. Rainfall is necessary to create collections of water for mosquitoes to breed, but the breeding sites can also be flushed and destroyed by excessive rains.[13]

Using geographic information system (GIS) software, we digitize the 1948 malaria endemicity map. *Figure 3* shows the digitized map. There are 466 districts in the 1991 census. However, the National Sample Survey (NSS), used for our outcomes measures, groups some districts together, resulting in 431 NSS "districts." We follow the NSS district coding. We drop three island districts (Andaman and Nicobar Islands, and Lakshadweep) for which malaria prevalence data are unavailable. In our analysis, we group districts in categories 1 and 2 together, and define them as non-malarious regions. Districts in categories 3 and 4 are grouped together and defined as potential epidemic areas, while districts in categories 5 and 6, where malaria is endemic, are classified as malarious areas. We classify each district into one of these three groups.

As can be seen in *Figure 3*, some districts have more than one possible classification. 295 districts (68.9%) are "unambiguous," that is, they contain only one of the six categories or two categories that fall in the same group (e.g. categories 1 and 2). 53 districts (12.4%) are "majority" districts. These districts have one group that comprises a majority of the possible categorizations and are assigned that group. For example, districts with categories 3, 4, and 5 are categorized as potentially epidemic. The remaining 80 districts (18.7%) have no majority and are "discretionary." In our main specifications, we assign these districts to the highest possible prevalence category. For example, districts with categories 4 and 5 are classified as malarious. Under this classification, 24

[13] From National Center for Infectious Diseases, Division of Parasitic Diseases, http://www.cdc.gov/malaria/distribution_epi/epidemiology.htm.

districts (5.6%) are coded as non-malarious, 114 (26.6%) are potential epidemic, and 290 (67.8%) are malarious. Misclassification of districts should, in general, bias our results downward. We also perform extensive robustness analysis to test whether the method of district classification affects the empirical findings.

4.2 Outcomes

We use data on educational attainment from the 43rd round of the Indian National Sample Survey (NSS), conducted in 1987. The NSS is an all-India representative household consumer expenditure survey set up by the Government of India in 1950. It includes a parallel employment and unemployment survey every five years. The NSS has four "thick" rounds that have the largest samples: namely, 1983, 1987, 1993, and 1999. We use the 43rd round (1987) because it is the earliest thick round that contains district identifiers. Choosing an early round mitigates possible mortality bias, and using the district identifiers allows us to examine outcomes at a very local level. The NSS reports district of current residence but not district of birth.

Table 4 provides summary statistics for the entire survey sample in the 43rd round. Almost 300,000 households were sampled, giving slightly over one million individual observations. 78% of sample individuals live in rural areas, 41% are female, and the average age in the sample is 28.

For our main specifications, we focus on the rural sample. We restrict the sample to individuals aged 15 and over for literacy and primary school completion outcomes and to individuals aged 20 and over for higher level educational outcomes. We exclude those born during the eradication era (1953 to 1961). The ten birth cohorts born from 1962 to 1972 are defined to be post-eradication cohorts and the forty birth cohorts born from 1912 to 1952 are defined to be pre-eradication cohorts. We demonstrate in the robustness analysis that the results are not sensitive to the choice of age cutoff or the choice of initial birth cohort for the pre-eradication group.

Schedule 10 (employment schedule) of the NSS gives information on education. In the NSS, individuals report the highest level of education they have attained. *Table 5* provides summary statistics for the pre- and post-eradication cohorts. The variable "literate" is a dummy variable that equals one if the individual can read and write; "primary" is a dummy that equals one if the individual has finished at least primary school; etc. For the pre-eradication cohorts, the average literacy rate and completion rates for all levels of schooling are highest in non-malarious

districts, consistent with malaria prevalence depressing educational attainment. The differences in educational outcomes between the non-malarious districts and the potentially epidemic and malarious districts are generally much smaller for the post-eradication cohorts. In some cases, the point estimates for educational attainment variables for these cohorts are even higher for the districts that were malarious in the pre-eradication period relative to those that were non-malarious.

5 Results

5.1 Baseline Results

Table 6 shows the results of our baseline specification for a number of educational outcomes. The coefficients of interest are the coefficients on post*malarious and post*potential epidemic, which capture the effect of being born during the post-eradication period in a district that was malarious or potentially epidemic in the pre-eradication period. The OLS estimates for literacy rate and primary, middle, and secondary schooling are positive and statistically significant, reflecting the benefits of malaria eradication on educational attainment. The benefits decrease in magnitude for the higher levels of education completion. Effects on college education are insignificant. This is perhaps unsurprising, given that the nationwide average level of college completion is less than 2%, even in the post-eradication era.

Column 1 shows that the gain in literacy rate between those born in post-eradication versus pre-eradication cohorts is 12.5 percentage points higher for those born in a malarious region compared to those born in a non-malarious region. Those born in potential epidemic areas also benefit from an 11.7 percentage point relative increase in literacy rate. *Column 2* shows that malaria eradication leads to a 12.9 percentage point increase in the probability of completing at least primary school for those born in malarious regions and to an 11.7 percentage point increase for those born in potential epidemic areas. The benefits decrease in magnitude for the higher levels of education completion (middle and secondary school) but are still significant at the 10% level. The gain in all educational attainment variables is higher for those in malarious areas as compared to potential epidemic areas, as we would expect since malarious areas experienced larger declines in malaria prevalence.

These increases are quite substantial relative to a nationwide average literacy rate of 32% and a primary school completion rate of 20% for the pre-eradication cohorts. From *Table 5*, we see

that the raw difference in literacy rate between the pre- and post-eradication cohorts in malarious regions is 0.241.[14] This means malaria eradication accounts for (0.125 / 0.241) = 51.9% of the increase in literacy rate in malarious regions. Similarly, malaria eradication accounts for (0.117/ 0.215) = 54.4% of the increase in literacy rate in potential epidemic areas. Similar calculations demonstrate that eradication was responsible for 49.0% of the increase in primary school completion in malarious areas and 48.9% of the increase in potential epidemic areas.

The NSS education data only include levels of education completion. We calculate the effect of malaria on years of schooling by approximating years of schooling for each level of education completion. We define less than primary as zero years, primary as five years, middle as eight years, secondary as 12 years, and college as 16 years (Duraisamy, 2002). We find that malaria eradication increases years of schooling by 0.89 years for those born in malarious areas and 0.78 years for those born in potential epidemic areas, relative to those born in non-malarious areas (*Column 6*). Both effects are significant at the 5% level.

There are several things to note when interpreting the above results. First, cohort attrition may lead to a downward bias in the estimated effect of malaria eradication. In the pre-eradication era, there were 75 million cases of malaria per year and 800,000 deaths. Those who survived to older ages were presumably the stronger cohort members. Weaker cohort members who would have died in the pre-eradication period were able to survive post-eradication. Hence the educational attainment of individuals in the eradication period is averaged over both strong and weak cohort members, and the estimated gain in educational attainment for those who would have survived is therefore likely to be downward biased.

Second, the estimated effect of malaria eradication includes both direct and indirect effects. An example of a direct effect is the reduction in school absence for an individual who now does not suffer from malaria. An example of an indirect effect is the reduction in school absence for an individual who now does not need to take care of younger siblings who suffer from malaria.

[14] As described in the previous section, the pre-eradication cohorts are defined as 1912-1952, and the post-eradication cohorts are defined as 1962-1972.

5.2 Robustness Tests

5.2.1 Empirical Specification and Sample Selection

Our robustness analysis focuses on the literacy and primary school completion outcomes, since these effects are the strongest in the baseline specification. *Column 2* of *Table 7* presents the results with district and year of birth dummies but excluding individual control variables. Excluding the individual controls causes the point estimates to decrease slightly, but the estimated effects are still quite close to the baseline effects, and all are statistically significant at or above the 5% level. The results are also robust, and in fact somewhat stronger, when we exclude the state of Jammu and Kashmir, which experienced conflict and two wars during this time period (unreported).

Next, we examine effects for urban areas. As discussed, malaria was not nearly as serious a problem in urban India as in rural India prior to the eradication era. *Column 3* of *Table 7* shows the results of our baseline specification for the urban sample. The point estimates suggest that individuals in endemic or potentially epidemic districts had improvements in literacy and primary school completion of approximately one to five percentage points relative to those in non-malarious districts, effects that are much smaller than those in the rural sample. None of the estimates for the urban sample are statistically significant. These results suggest that the gains observed in the rural sample are not due to overall economic improvements in malarious areas that are unrelated to the eradication campaign, since such improvements would have most likely affected both rural and urban areas.

We also examine the effects on the eradication era cohorts. Since districts were phased into the program over this period, some individuals in these cohorts would have been under the program during their birth and childhood years and others would not. We would therefore expect a smaller effect of the program on the eradication era cohorts. From *Column 4*, we see that eradication era cohorts in malarious areas experienced a 7.3 percentage point gain in literacy and a 10.9 percentage point gain in primary school completion on average. Similarly, eradication era cohorts in potential epidemic areas experienced a 5.8 percentage point gain in literacy and an 8.9 percentage point gain in primary school completion. All of these effects are strongly significant.

We next adjust the choice of pre-eradication cohorts. We limit the pre-eradication cohorts to the ten birth cohorts born between 1942 and 1952 (*Column 5*). The estimates are slightly

smaller than the baseline estimates but still significant. We see that those born in malarious regions experienced an approximately 9.5 percentage point increase in literacy and primary school completion as compared to those born in non-malarious districts. Those born in potential epidemic areas experienced about 9 percentage point increase in educational attainment. The estimates are higher when we limit the pre-eradication cohorts to the twenty or thirty birth cohorts prior to 1953 but still lower than the baseline estimates.[15]

One explanation for this pattern is that not all of the pre-1953 cohorts are truly "pre-eradication." Our baseline specification assumes that only individuals born after the eradication program is introduced benefit, but some of the pre-1953 cohorts may also have benefited from the program during their later childhood years. Bleakley (2007), for example, assumes that the effect of malaria exposure is uniform across youth, defined as birth to 18 years of age, and finds evidence supporting this hypothesis. If individuals benefit from a reduction in malaria exposure during their childhood years, our baseline estimates will be biased downward. This bias will be more severe when the pre-eradication cohorts are limited to the 1942-1952 cohorts.

We explore this possibility further by adjusting our pre-eradication window to allow for effects of the program at different years of life. If cohorts born immediately prior to 1952 benefited from the program, the exclusion of these cohorts should cause the estimated effects to increase. *Column 6* excludes the five birth cohorts prior to 1953. The estimated effects (with the exception of literacy in potential epidemic areas) are slightly higher than the baseline estimates. The effects increase slightly further when we exclude the ten cohorts prior to 1953 (*Column 7*). These results provide some suggestive evidence, consistent with Bleakley (2007), that the benefits of reducing malaria exposure may accrue throughout the childhood years.

Finally, we test whether our results are robust to the way we have classified districts into malaria prevalence groups based on the 1948 endemicity map. *Column 2* of *Table 8* restricts the sample to "unambiguous" districts, as defined in the data section. The effects in this restricted sample are almost identical to the effects in the full sample; the literacy gain in malarious areas is 12.6 percentage points and the gain in potentially epidemic areas is 10.2 percentage points. The gain in primary school completion is 13.5 percentage points for malarious regions and 10.4 percentage

[15] The results are also robust to restricting the sample to individuals aged 20 and above. Results available from authors on request.

points for potentially epidemic areas. All of these effects are strongly significant. *Column 3* restricts the sample to "unambiguous" and "majority" districts. *Columns 4* and *5* show the results classifying districts into the least and most malarious categories possible, respectively. The results are similar to the baseline results and are significant at or above the 5% level in all cases.

5.2.2 District Convergence

A concern with the identification strategy is whether districts with high malaria prevalence followed the same trends over time relative to areas with low malaria prevalence. In particular, we might be concerned that educational attainment in potential epidemic and malarious areas would have converged toward non-malarious regions even in the absence of the eradication program. We test for convergence in two ways (*Table 9*).

First, we control for pre-eradication educational attainment at the district level interacted with the post dummy. This specification allows districts with low and high initial educational attainment to experience differential gains across cohorts. We construct two measures of initial educational attainment: average literacy and average primary school completion at the district level averaged over the pre-eradication (1912-1952) birth cohorts. Including an interaction of the post dummy with initial literacy does not affect our estimates of the effects of the eradication program (*Column 2*). Those in malarious areas experience a relative increase of approximately 13 percentage points in literacy and primary school completion, and those in potential epidemic areas experience relative increases of approximately 11 percentage points. The coefficient on post*initial literacy is negative, as we would expect if there is convergence in educational attainment over time. Controlling for post*initial primary completion gives almost identical estimates of the effects of malaria eradication, and the coefficient on post*initial primary is also negative, as expected (*Column 3*). The results are very similar if initial literacy and primary school completion are included together (*Column 4*) and are also very similar if we instead use measures of the educational attainment of the first pre-eradication cohort (1912) or the middle pre-eradication cohort (1932) interacted with the post dummy (unreported).[16]

Second, we include post*region dummies in addition to district and year of birth fixed effects. A "region" is a sub-state geographic area defined by the NSS. There are 77 NSS regions and

[16] Results available from authors on request.

the median region contains 10 districts. This strategy allows us to identify the effects of malaria eradication using only *within-region* district variation in pre-eradication prevalence.[17] Our results are robust to the inclusion of post*region controls. *Column 5* shows that those born in potential epidemic areas or malarious areas experienced a 10 percentage point increase in literacy and 13 to 15 percentage point increase in primary school completion relative to those born in non-malarious regions. All effects are significant at or above the 5% level.[18] These results provide further evidence that our findings are not driven by mere convergence across geographic areas.

5.3 Heterogeneity in Treatment Effects

We next test whether the effects of malaria eradication on educational outcomes vary by gender or caste (*Table 10*). We find similar gains for males and females in malarious areas: males experience a 10.4 percentage point increase in literacy rate while females experience a 14.1 percentage point increase. The percentage point gains in primary school completion are 11.4 for males and 13.6 for females. Females experienced somewhat higher gains than males in potential epidemic areas. We also examine the effects for individuals that are members of a scheduled caste (SC) or scheduled tribe (ST) versus those that are not. Scheduled castes and scheduled tribes are groups that have been historically disadvantaged in India. We find that these groups experienced slightly larger effects for literacy and slightly smaller effects for primary school completion than those experienced by the non-scheduled caste and tribe. Overall, these results suggest that gains from the eradication program were broad and resulted in increased educational attainment across groups.

6 Conclusion

This paper examines the effects of a large scale eradication program that drastically reduced malaria in India over a short time period. Exploiting the heterogeneity in indigenous malaria rates and the exogenous implementation of the eradication program, we find that malaria eradication leads to substantial increases in educational attainment. The program resulted in gains in literacy and primary school completion of approximately 12 percentage points. These estimates suggest that

[17] 44 out of 77 regions (57%) have inter-regional variation in malaria prevalence.

[18] The point estimates are still positive when post*state controls rather than post*region controls are included, although the results are not statistically significant. We cannot reject that the coefficients are the same as in the baseline.

the reduction of malaria can explain about half of the increases in these measures between pre- and post-eradication cohorts. Effects for higher levels of schooling are smaller but still sizeable and significant. In all cases, estimated gains are slightly higher in areas where malaria was endemic prior to the program as compared with areas where malaria was potentially epidemic. The gains from the program were broad-based and accrued even to traditionally disadvantaged groups.

Our results are robust to the precise definition of pre- and post-eradication cohorts as well as to adjustments in the categorization of districts. We find no effect of the program in urban areas, a natural placebo group since urban malaria was not a problem in the pre-eradication era. The results cannot be explained by convergence across areas. The results are also unaffected when we allow districts to experience differential gains based on their initial levels of educational attainment.

Overall, we find strong support for the belief that improvements in health and in the disease environment can have a causal effect on human capital accumulation. If the rate of return to schooling is 10% a year, our estimates imply that malaria reduces income by 7-10% through the channel of educational attainment alone. This may be an important causal channel between health and economic growth.

References

[1] Acemoglu, Daron and Simon Johnson (2005). "Disease and Development: The Effect of Life Expectancy on Economic Growth," NBER Working Paper Number 12269.

[2] Aikins MKS (1995). "Cost-effectiveness Analysis of Insecticide-impregnated Mosquito Nets (Bednets) Used as a Malaria Control Measure: A Study from the Gambia," PhD thesis, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, University of London.

[3] Alleyne, George and Daniel Cohen (2002). "The Report of Working Group I of the Commission on Macroeconomics and Health," WHO Commission on Macroeconomics and Health, April.

[4] Barreca, Alan (2007). "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria," UC Davis mimeo.

[5] Bleakley, Hoyt (2007). "Malaria in the Americas: A Retrospective Analysis of Childhood Exposure," University of Chicago mimeo.

[6] Bloom, David and David Canning (2005). "Health and Economic Growth: Reconciling the Micro and Macro Evidence," CCDRL working paper.

[7] Boyle, C.A., P. Decoufle, M. Yeargin-Allsopp (1994). "Prevalence and Health Impact of Developmental Disabilities in U.S. Children," *Paediatrics*, 93(3): 399-403.

[8] Brooker, S., H. Guyatt, J. Omumbo, R. Shretta, L. Drake and J. Ouman (2000). "Situation Analysis of Malaria in School-aged Children in Kenya - What Can Be Done?" *Parisitology Today*, 16(5): 183-186.

[9] Clarke, Siân, Matthew Jukes, Joseph Njagi, Lincoln Khasakhala, Bonnie Cundill, Julius Otido, Christopher Crudder, Benson Estambale and Simon Brooker (2007). "Intermittent preventive treatment in schoolchildren - a new strategy to reduce malaria-related anaemia and improve learning: results of a randomised controlled trial," mimeo, London School of Hygiene and Tropical Medicine.

[10] Davis, Kingsley (1956). "The Amazing Decline of Mortality in Underdeveloped Areas," *American Economic Review*, 46(2): 305-318.

[11] Desowitz, Robert (1991). *The Malaria Capers: More Tales of Parasites and People, Research and Reality,* New York: W.W. Norton.

[12] Duffy, P.E. and R.S. Desowitz (2001). "Pregnancy Malaria Throughout History: Dangerous Labors," in *Malaria in Pregnancy: Deadly Parasite, Susceptible Host*, ed. by P.E. Duffy, and M. Fried, London: Taylor and Francis.

[13] Duraisamy, P. (2002), "Changes in Returns to Education in India, 1983-94: by Gender, Age-cohort and Location," *Economics of Education Review*, 21(6): 609-622.

[14] Gallup, John Luke and Jeffrey Sachs (2001). "The Economic Burden of Malaria," *American Journal of Tropical Medicine and Hygiene*, 64: 85-96.

[15] Hoff, Brent and Carter Smith III (2000). *Mapping Epidemics: A Historical Atlas of Disease*, Franklin Watts, New York.

[16] Holding, P.A. and R.W. Snow (2001). "Impact of Plasmodium Falciparum Malaria on Performance and Learning: Review of the Evidence," *American Journal of Tropical Medicine and Hygiene*, 64 supplement: 68-75.

[17] Hong, Sok Chul (2007). "A Longitudinal Analysis of the Burden of Malaria on Health and Economic Productivity: The American Case," University of Chicago mimeo.

[18] Kiple, Kenneth (1993). *The Cambridge World History of Human Disease*, Cambridge University Press.

[19] League of Nations (1930). "Report of the Malaria Commission on its Study Tour in India," Geneva, August.

[20] Leighton, Charlotte and Rebecca Foster (1993). "Economic Impacts of Malaria in Kenya and Nigeria," Major Applied Research Paper Number 6, Office of Health, Health Services Division, Bureau of Research and Development Agency for International Development.

[21] Lucas, Adrienne (2005). "Economic Effects of Malaria Eradication: Evidence from the Malarial Periphery," Brown University mimeo.

[22] Lucas, Adrienne (2007). "The Impact of Disease Eradication on Fertility and Education," Wellesley College mimeo.

[23] McCormick, M.C., J. Brooks-Gunn, K. Workman-Daniels, J. Turner, and G.J. Peckman (1992) "The Health and Development Status of Very Low-Birth-Weight Children at School Age," *Journal of the American Medical Association*, 267:16.

[24] NMEP (1986). *Malaria and its Control in India, Vol. 1.* National Malaria Eradication Program, India.

[25] NMEP (1996). *Epidemiology and Control of Malaria in India.* National Malaria Eradication Program, India.

[26] Rao, B.A. (1959). *Malaria in India*, New Delhi: Central Health Education Bureau.

[27] Shiff, C., W. Checkley, P. Winch, J. Minijas, and P. Lubega (1996). "Changes in Weight Gain and Anaemia Attributable to Malaria in Tanzanian Children Living under Holoendemic Conditions," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 90: 262-265.

[28] Sinton, J. A. (1935). "What Malaria Costs India, Nationally, Socially and Economically." *Records of the Malaria Survey of India* 5: 223–64 and 413–89.

[29] Sinton, J.A . (1936). "What Malaria Costs India, Nationally, Socially and Economically." *Records of the Malaria Survey of India* 6: 91–169.

[30] World Health Organization (1967). *The World Health Organization: Twenty Years in South-East Asia, 1948-1967*, New Delhi: World Health Organization, Regional Office for South-East Asia.

[31] World Health Organization (2001). *Macroeconomics and Health: Investing in Health for Economic Development.*

FIGURE 1.1: NMCP 1953-54

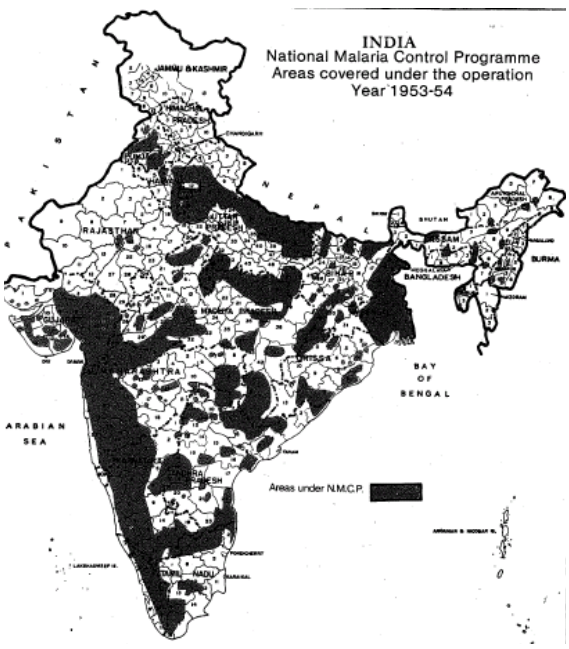

FIGURE 1.2: NMCP 1954-55

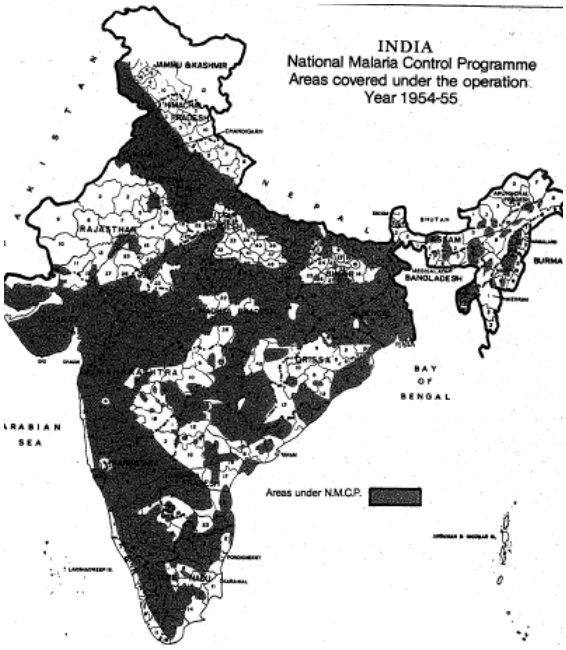

FIGURE 1.3: NMCP 1956-57

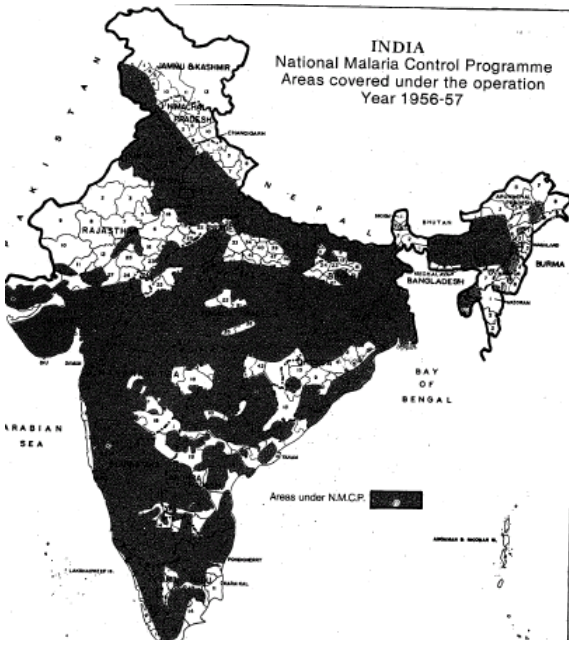

FIGURE 1.4: NMEP 1959-61

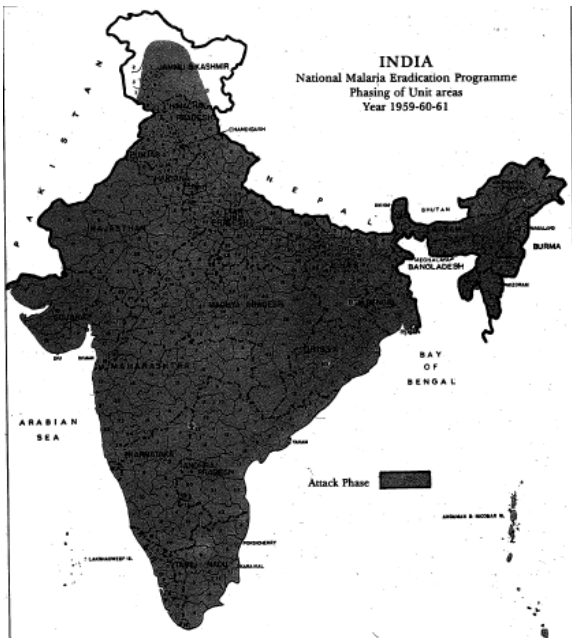

FIGURE 2: DECLINES IN MALARIA DEATHS FROM VITAL STATISTICS DATA

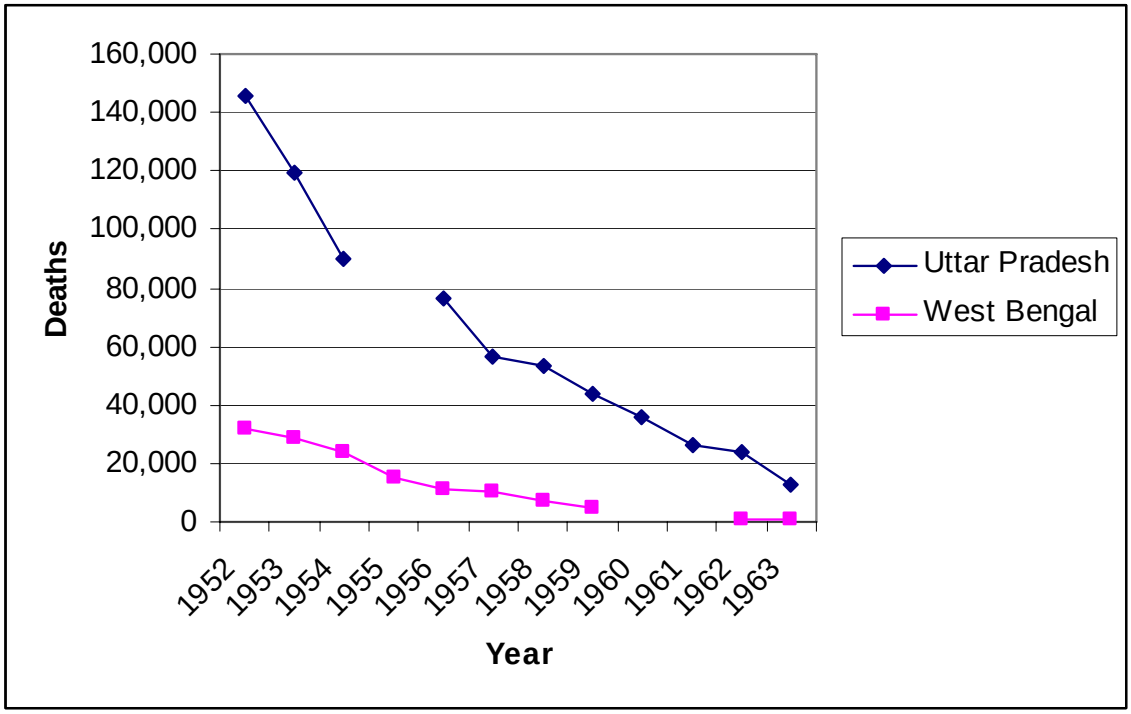

FIGURE 3: MALARIA ENDEMICITY MAP

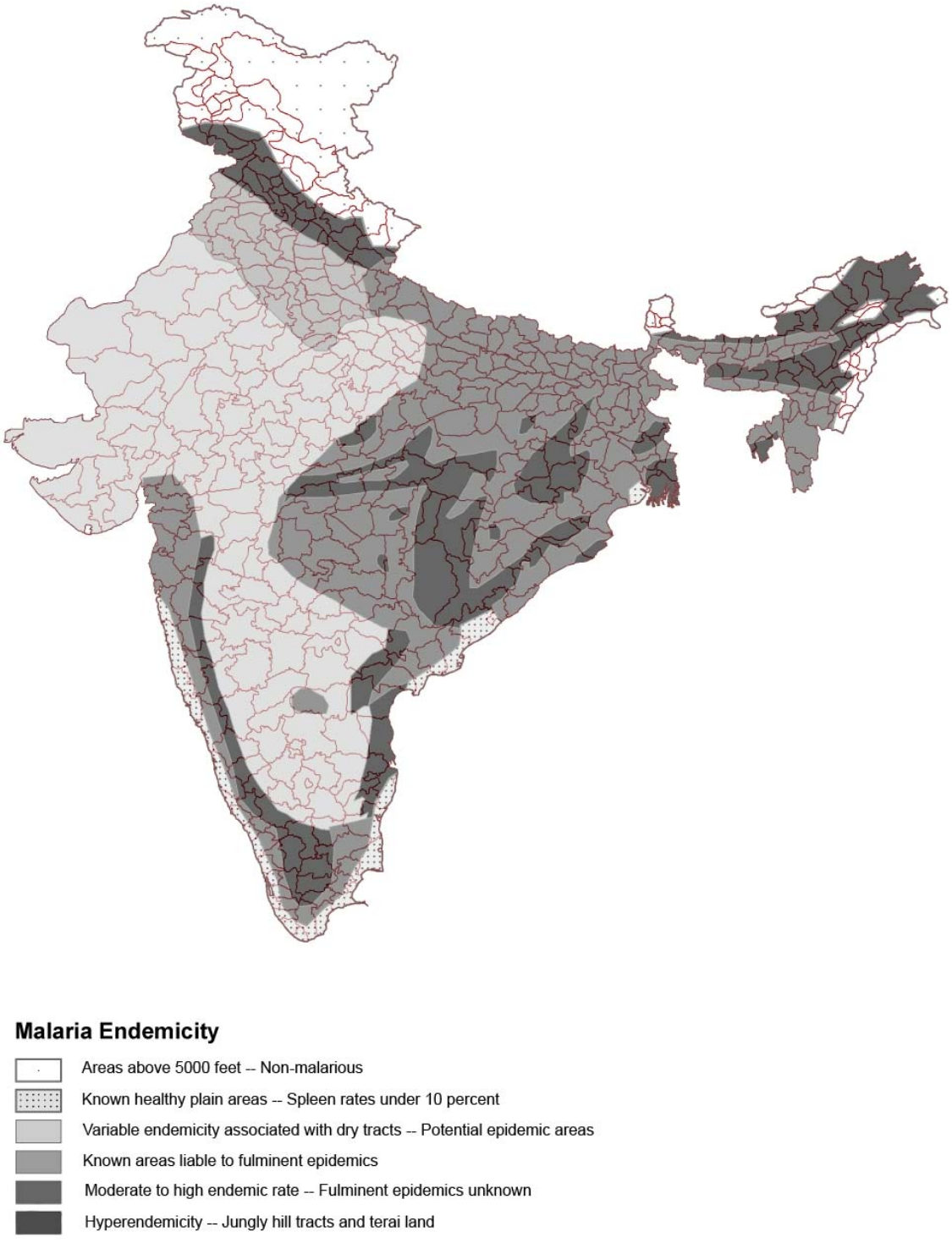

TABLE 1
Expansion of the National Malaria Control Program

	Number of Units Established	Number of Units Established (Cumulative)	Population Protected (Cumulative, in millions)	Percentage Population Protected
1953-54	84.00	84.00	49.50	13.2
1954-55	26.75	110.75	79.90	20.9
1955-56	23.00	133.75	112.00	28.7
1956-57	35.50	169.25	144.50	36.3
1957-58	23.25	192.50	165.57	40.8

Source: NMCP (1986). The population statistics used to calculate the percentage of the population protected are from the United Nations Demographic Yearbook Historical Supplement. We use the 1953 midyear population figure for 1953-54, etc.

TABLE 2
Malaria Prevalence over Time

	Population (in millions)	Malaria cases (in millions)	Percentage of population with malaria each year	Number of deaths
1947	344	75.0	21.8	800,000
1965	483	0.1	0.02	0
1976	616	6.5	1.1	59
1984	735	2.2	0.3	247
1994	900	2.5	0.3	1122

Source: NMCP (1986), NMCP (1996). The population statistics used to calculate the percentage of the population with malaria are from the United Nations Demographic Yearbook Historical Supplement.

TABLE 3
Malaria Prevalence Indicators during the NMCP

	Child Spleen Rate	Child Parasite Rate	Infant Parasite Rate
1953-54	15.7	3.9	1.6
1954-55	12.4	4.2	2.0
1955-56	7.7	1.8	0.7
1956-57	6.0	1.1	0.5
1957-58	4.2	0.8	0.6
1958-59	3.2	0.5	0.2
1959-60	1.4	0.16	0.14

Source: NMCP (1986)

TABLE 4
Summary Statistics for 43rd National Sample Survey (NSS)

	Mean	Standard Deviation
% Rural	0.776	(0.42)
% Female	0.411	(0.49)
% Married	0.539	(0.50)
Age	27.669	(17.85)
Household size	6.178	(2.96)
% Scheduled caste	0.176	(0.38)
% Scheduled tribe	0.093	(0.29)
% Hindu	0.833	(0.37)
% Muslim	0.110	(0.31)
% Christian	0.023	(0.15)
Number of states	31	
Number of regions	77	
Number of districts	431	
Number of households	291,648	
Number of observations	1,005,863	

Source: NSS (1987)

TABLE 5
Summary Statistics for Pre- and Post-Eradication Cohorts

	Non-Malarious Regions			Potential Epidemic Areas			Malarious Regions		
	Pre-cohort (1)	Post-cohort (2)	Difference (3)=(2)-(1)	Pre-cohort (4)	Post-cohort (5)	Difference (6)=(5)-(4)	Pre-cohort (7)	Post-cohort (8)	Difference (9)=(8)-(7)
Literate	0.363	0.534	0.171	0.268	0.483	0.215	0.331	0.572	0.241
	(0.481)	(0.499)	[26.808]	(0.443)	(0.500)	[55.579]	(0.471)	(0.495)	[95.946]
Primary	0.247	0.454	0.207	0.163	0.402	0.239	0.205	0.468	0.263
	(0.431)	(0.498)	[31.275]	(0.369)	(0.490)	[69.161]	(0.404)	(0.499)	[118.471]
Middle	0.136	0.317	0.181	0.088	0.221	0.133	0.106	0.235	0.129
	(0.343)	(0.466)	[28.067]	(0.284)	(0.415)	[56.185]	(0.308)	(0.424)	[107.564]
Secondary	0.066	0.165	0.099	0.045	0.112	0.067	0.053	0.118	0.065
	(0.248)	(0.372)	[14.677]	(0.207)	(0.315)	[29.188]	(0.224)	(0.322)	[53.958]
College	0.020	0.031	0.011	0.008	0.019	0.011	0.012	0.021	0.009
	(0.139)	(0.173)	[1.225]	(0.091)	(0.137)	[3.184]	(0.107)	(0.143)	[2.761]
% Hindu	0.449	0.433	-0.016	0.903	0.896	-0.007	0.854	0.842	-0.012
	(0.497)	(0.496)	[-3.934]	(0.296)	(0.305)	[-2.734]	(0.353)	(0.098)	[-8.769]
% Muslim	0.395	0.429	0.034	0.051	0.058	0.007	0.087	0.098	0.011
	(0.489)	(0.495)	[5.192]	(0.220)	(0.233)	[3.892]	(0.281)	(0.298)	[10.447]
% Scheduled Caste	0.070	0.061	-0.009	0.199	0.200	0.001	0.188	0.192	0.004
	(0.255)	(0.240)	[-0.194]	(0.399)	(0.400)	[0.268]	(0.391)	(0.394)	[2.451]
% Scheduled Tribe	0.040	0.028	-0.012	0.080	0.086	0.006	0.112	0.123	0.011
	(0.197)	(0.166)	[-2.820]	(0.271)	(0.281)	[2.927]	(0.316)	(0.329)	[8.725]
# observations	8,813	5,659		44,037	32,789		164,404	112,371	

Note: The number in parentheses is the standard deviation, and the number in square brackets is the t-statistic. Summary statistics are for rural sample only. Pre-eradication cohorts are defined as 1912-1952. Post-eradication cohorts are defined as 1962-1972 for literate and primary school completion outcomes, and 1962-1967 for other educational outcomes.

TABLE 6
Baseline Results: DD Regression Using 1948 Malaria Endemicity Map

	(1) Literate	(2) Primary	(3) Middle	(4) Secondary	(5) College	(6) Years of schooling
Post * Malarious	0.125	0.129	0.056	0.027	0.009	0.892
	(0.040)***	(0.052)**	(0.028)**	(0.014)*	(0.006)	(0.384)**
Post * Potential Epidemic	0.117	0.117	0.055	0.026	0.007	0.784
	(0.040)***	(0.053)**	(0.028)*	(0.014)*	(0.007)	(0.386)**
Female	-0.306	-0.226	-0.144	-0.076	-0.015	-1.815
	(0.005)***	(0.004)***	(0.003)***	(0.002)***	(0.001)***	(0.036)***
Scheduled Caste	-0.181	-0.154	-0.103	-0.060	-0.013	-1.255
	(0.007)***	(0.006)***	(0.004)***	(0.003)***	(0.001)***	(0.051)***
Scheduled Tribe	-0.227	-0.180	-0.114	-0.064	-0.015	-1.438
	(0.010)***	(0.009)***	(0.006)***	(0.004)***	(0.001)***	(0.070)***
Hindu	0.002	-0.005	-0.014	-0.006	-0.001	-0.085
	(0.018)	(0.018)	(0.013)	(0.010)	(0.003)	(0.151)
Muslim	-0.095	-0.122	-0.102	-0.055	-0.013	-1.056
	(0.022)***	(0.021)***	(0.015)***	(0.011)***	(0.003)***	(0.169)***
Christian	0.092	0.075	0.060	0.041	0.009	0.715
	(0.028)***	(0.026)***	(0.020)***	(0.013)***	(0.004)**	(0.216)***
Jain	0.324	0.305	0.152	0.088	0.050	2.439
	(0.061)***	(0.054)***	(0.039)***	(0.036)**	(0.023)**	(0.474)***
Constant	0.570	0.202	0.156	0.073	0.022	1.720
	(0.021)***	(0.021)***	(0.011)***	(0.006)***	(0.002)***	(0.137)***
Observations	366801	366801	304048	304048	304048	304048
R-squared	0.27	0.24	0.16	0.09	0.02	0.22

Note: Map categories 1 and 2 (omitted category) are defined as non-malarious regions. Map categories 3 and 4 are potential epidemic areas. Map categories 5 and 6 are malarious regions. Analysis focuses on rural sample only. Total sample excludes 1953-1961 cohorts born during the eradication era. Pre-eradication cohorts are defined as 1912-1952. Post-eradication cohorts are defined as 1962-1972 (age 15 or above) for columns (1) and (2), and are defined as 1962-1967 (age 20 or above) for the other columns. Control variables include district and year of birth dummies, female, scheduled caste, scheduled tribe, and household religion. Robust standard errors clustered by district. ** significant at 5%; *** significant at 1%.

TABLE 7
Robustness Tests: Empirical Specification and Sample Selection

	(1) Baseline specification	(2) Exclude controls	(3) Urban	(4) Include eradication-era cohort	(5) Pre-eradication = 1942-52	(6) Pre-eradication = 1912-47	(7) Pre-eradication = 1912-42
Dependent Variable: Literate							
Post * Malarious	0.125	0.116	0.031	0.125	0.096	0.128	0.136
	(0.040)***	(0.037)***	(0.043)	(0.040)***	(0.043)**	(0.037)***	(0.038)***
Post * Potential Epidemic	0.117	0.105	0.011	0.117	0.091	0.116	0.127
	(0.040)***	(0.037)***	(0.044)	(0.040)***	(0.043)**	(0.037)***	(0.038)***
Eradication era * Malarious				0.073			
				(0.015)***			
Eradication era * Potential Epidemic				0.058			
				(0.016)***			
Observations	366801	366801	184282	458680	258557	319246	279698
R-squared	0.27	0.15	0.24	0.28	0.14	0.28	0.28
Dependent Variable: Primary							
Post * Malarious	0.129	0.122	0.053	0.129	0.095	0.142	0.148
	(0.052)**	(0.050)**	(0.041)	(0.052)**	(0.047)**	(0.055)**	(0.057)***
Post * Potential Epidemic	0.117	0.109	0.045	0.118	0.087	0.126	0.133
	(0.053)**	(0.050)**	(0.042)	(0.053)**	(0.047)*	(0.055)**	(0.057)**
Eradication era * Malarious				0.109			
				(0.026)***			
Eradication era * Potential Epidemic				0.089			
				(0.026)***			
Observations	366801	366801	184282	458680	258557	319246	279698
R-squared	0.24	0.15	0.23	0.24	0.14	0.24	0.25

Note: Baseline specification follows Table 6. Column (2) excludes the set of control variables. Column (3) includes cohorts from all three periods (pre-eradication, eradication, and post-eradication). Column (4) narrows the pre-eradication window from 1912-52 to 1942-52. Columns (5) and (6) shift the pre-eradication window earlier, away from 1953 which is the first year of the eradication program. All regressions except Column (2) include district and year of birth dummies, as well as control variables including female, scheduled caste, scheduled tribe, and household religion. Robust standard errors clustered by district. ** significant at 5%; *** significant at 1%.

TABLE 8
Robustness Tests: Map Categorization of Malaria Endemicity

	(1) Baseline specification	(2) Restrict to "unambiguous" districts	(3) Restrict to "unambiguous" and "majority" districts	(4) Choosing the least malarious category possible	(5) Choosing the most malarious category possible
Dependent Variable: Literate					
Post * Malarious	0.125	0.126	0.116	0.072	0.125
	(0.040)***	(0.038)***	(0.038)***	(0.017)***	(0.041)***
Post * Potential Epidemic	0.117	0.102	0.102	0.047	0.116
	(0.040)***	(0.038)***	(0.038)***	(0.017)***	(0.041)***
Observations	366801	241682	294553	366801	366801
R-squared	0.27	0.26	0.27	0.27	0.27
Dependent Variable: Primary					
Post * Malarious	0.129	0.135	0.124	0.095	0.128
	(0.052)**	(0.047)***	(0.047)***	(0.018)***	(0.053)**
Post * Potential Epidemic	0.117	0.104	0.105	0.068	0.112
	(0.053)**	(0.047)**	(0.047)**	(0.018)***	(0.053)**
Observations	366801	241682	294553	366801	366801
R-squared	0.24	0.22	0.23	0.24	0.24

Note: Baseline specification follows Table 6. All regressions include district and year of birth dummies, as well as control variables including female, scheduled caste, scheduled tribe, and household religion. Robust standard errors clustered by district. ** significant at 5%; *** significant at 1%.

TABLE 9
Robustness Tests: District Convergence

	(1) Baseline specification	(2) Include post*(1912-52 literate)	(3) Include post*(1912-52 primary)	(4) Include post*(1912-52 literate) & post*(1912-52 primary)	(5) Include post*region	(6) Include post*state
Dependent Variable: Literate						
Post * Malarious	0.125	0.128	0.127	0.124	0.102	0.067
	(0.040)***	(0.037)***	(0.037)***	(0.038)***	(0.046)**	(0.056)
Post * Potential Epidemic	0.117	0.114	0.112	0.112	0.100	0.047
	(0.040)***	(0.037)***	(0.037)***	(0.038)***	(0.048)**	(0.055)
Post * (1912-52 literate)		-0.129		0.138		
		(0.033)***		(0.078)*		
Post * (1912-52 primary)			-0.189	-0.340		
			(0.039)***	(0.093)***		
Observations	366801	366800	366800	366800	366801	366801
R-squared	0.27	0.27	0.27	0.27	0.28	0.28
Dependent Variable: Primary						
Post * Malarious	0.129	0.135	0.131	0.135	0.158	0.092
	(0.052)**	(0.039)***	(0.043)***	(0.039)***	(0.059)***	(0.066)
Post * Potential Epidemic	0.117	0.110	0.109	0.110	0.130	0.091
	(0.053)**	(0.040)***	(0.043)**	(0.040)***	(0.060)**	(0.066)
Post * (1912-52 literate)		-0.292		-0.280		
		(0.033)***		(0.087)***		
Post * (1912-52 primary)			-0.321	-0.015		
			(0.039)***	(0.101)		
Observations	366801	366800	366800	366800	366801	366801
R-squared	0.24	0.24	0.24	0.24	0.25	0.24

Note: Baseline specification follows Table 6. Column (2) focuses on urban sample only. Columns (3) through (5) include post*base year education dummies. Base year education is defined as the average literacy or primary school completion rate by district from 1912 to 1952. Column (6) allows for differential regional trends by adding in post*region dummies. Column (7) allows for differential state trends by adding in post*state dummies. Robust standard errors clustered by district. ** significant at 5%; *** significant at 1%.

TABLE 10
Heterogeneity in Treatment Effects

	(1) Rural	(2) Rural Female	(3) Rural Male	(4) Rural SC / ST	(5) Rural non-SC & non-ST
Dependent Variable: Literate					
Post * Malarious	0.125	0.141	0.104	0.158	0.121
	(0.040)***	(0.058)**	(0.029)***	(0.028)***	(0.044)***
Post * Potential Epidemic	0.117	0.173	0.068	0.159	0.110
	(0.040)***	(0.059)***	(0.030)**	(0.028)***	(0.045)**
Observations	366801	146255	220546	110477	256324
R-squared	0.27	0.28	0.19	0.23	0.27
Dependent Variable: Primary					
Post * Malarious	0.129	0.136	0.114	0.097	0.122
	(0.052)**	(0.061)**	(0.046)**	(0.041)**	(0.060)**
Post * Potential Epidemic	0.117	0.172	0.073	0.092	0.110
	(0.053)**	(0.061)***	(0.046)	(0.042)**	(0.060)*
Observations	366801	146255	220546	110477	256324
R-squared	0.24	0.24	0.19	0.19	0.24

Note: Baseline specification follows Table 6. Robust standard errors clustered by district. ** significant at 5%; *** significant at 1%.