

NBER WORKING PAPER SERIES

LIFE EXPECTANCY AND HUMAN CAPITAL INVESTMENTS:
EVIDENCE FROM MATERNAL MORTALITY DECLINES

Seema Jayachandran
Adriana Lleras-Muney

Working Paper 13947
<http://www.nber.org/papers/w13947>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2008

We are grateful to Ran Abramitzky, Angus Deaton, Ilyana Kuziemko, Ofer Malamud, Jordan Matsudaira, Ben Olken; seminar participants at Berkeley, Bocconi, Brown, City University of New York, Cornell, Duke, Harris School, Harvard-Boston University-MIT, Minnesota, NBER Summer Institute, Notre Dame, Penn, Princeton, Stanford, SUNY Buffalo, UCL/LSE, University of Chicago, University of Illinois, Champaign-Urbana, UCLA, University of Maryland, and USC. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Seema Jayachandran and Adriana Lleras-Muney. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Life Expectancy and Human Capital Investments: Evidence From Maternal Mortality Declines
Seema Jayachandran and Adriana Lleras-Muney
NBER Working Paper No. 13947
April 2008
JEL No. I10,I20,O15

ABSTRACT

Longer life expectancy should encourage human capital accumulation, since a longer time horizon increases the value of investments that pay out over time. Previous work has been unable to determine the empirical importance of this life-expectancy effect due to the difficulty of isolating it from other effects of health on education. We examine a sudden drop in maternal mortality risk in Sri Lanka between 1946 and 1953, which creates a sharp increase in life expectancy for school-age girls without contemporaneous effects on health, and which also allows for the use of boys as a control group. Using additional geographic variation, we find that the 70% reduction in maternal mortality risk over the sample period increased female life expectancy at age 15 by 4.1%, female literacy by 2.5%, and female years of education by 4.0%.

Seema Jayachandran
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305-6072
and NBER
jayachan@stanford.edu

Adriana Lleras-Muney
Department of Economics
320 Wallace Hall
Princeton University
Princeton, NJ 08544
and NBER
alleras@princeton.edu

1. Introduction

A person will make a greater investment if the stream of returns from the investment lasts longer, all else equal. An important implication is that improvements in life expectancy should increase investment in human capital, which in turn may spur economic growth. A large literature has explored this idea theoretically (Ben Porath 1967, Kalemli-Ozcan, Ryder and Weil 2000, Soares 2005, Murphy and Topel 2005).

Most previous empirical research measures the cross-country relationship between life expectancy and education or growth (Shastry and Weil 2003, Lorentzen, McMillan and Wacziarg 2005, Acemoglu and Johnson 2007).¹ One limitation of that approach is that much of the cross-country variation over time in life expectancy is driven by changes in infant mortality. Deaths that occur before schooling begins do not affect the expected period over which returns to education are earned. In addition, the identifying variation in mortality is primarily from reductions in tuberculosis, malaria, and pneumonia, which also dramatically lowered morbidity rates. Another branch of the literature presents evidence on the effects of health on education, but does not aim to disentangle incentive effects of a longer life from direct effects on children's ability to attend school. For example, Miguel and Kremer (2004) and Bleakley (2007) show that deworming interventions led to increased school attendance. The main interpretation of their findings is that sickness had been preventing children from attending or succeeding in school.

This paper's main contribution is to examine how health affects individuals' investment in education specifically through the channel of increased life expectancy. We focus on reductions in maternal mortality, a source of variation that has several advantages for empirical identification. First, maternal mortality occurs after major human capital investments are made; investment decisions will depend on this risk, unlike infant mortality risk which is realized before educational decisions are made. Second, maternal mortality occurs early in adult life, so an averted maternal death translates into a large life-expectancy gain. Third, maternal mortality directly affects only women, so men can serve as a natural comparison group. Finally, maternal mortality, unlike examples used in previous papers such

¹ The evidence is mixed on the quantitative importance of improvements in life expectancy on growth and human capital. Shastry and Weil (2003) and Lorentzen, McMillan and Wacziarg (2005) find large effects, while Acemoglu and Johnson (2007) find small effects. See also Weil (2007), Young (2005), Rodriguez and Sachs (1999), and Barro and Sala-i-Martin (1995).

as hookworm and malaria, affects only adults and is not associated with high morbidity in the school-age population, allowing one to better isolate the incentive effects of longer life expectancy.^{2,3,4}

We examine a 70% decline in maternal mortality that occurred in Sri Lanka between 1946 and 1953.⁵ The maternal mortality ratio (MMR) fell from 1.8% (1.8 maternal deaths per 100 live births) to 0.5%. The decline (1) was very large not just in percentage terms but also in levels and thus had a substantial effect on life expectancy for women, (2) was concentrated in this seven-year period, allowing us to separate the effects caused by the MMR decline from slower, secular changes in education, (3) exhibited significant geographic variation across Sri Lanka's 19 districts. Our estimation strategy is a difference-in-difference-in-difference (DDD) across districts, time and gender.

We find that the reduction in maternal mortality risk between 1946 and 1953 increased female life expectancy at age 15 (censored at 65) by 1.5 years, or 4.1%, and can account for all of the gain in women's life expectancy relative to men during this period.⁶ The declines in maternal mortality increased female literacy among the affected cohorts by 2.5% (one percentage point) and increased years of schooling by 4.0% (0.17 years). The implied elasticity of human capital with respect to life expectancy is between 0.6 and 1.0.

² Complications from pregnancy and childbirth can result in subsequent disability or illness such as fistulae. Reductions in maternal mortality could be associated with higher or lower incidence of these complications.

³ Lucas (2005) studies several countries that underwent malaria eradication and finds that cohorts born after malaria eradication are more educated than those born before. This pattern could reflect the higher incentive to invest in human capital when life expectancy is higher, but also could reflect morbidity reductions.

⁴ Using HIV/AIDS to examine how education responds to longevity is confounded by HIV morbidity among children, and the larger problem of parents' health being affected along with the child's expected health. A child's HIV risk, as proxied by current HIV rates, is correlated with her parent's actual and anticipated morbidity and mortality from HIV. MMR has several advantages over HIV. Income effects from parents' affliction will be higher for HIV both because a long illness usually precedes an HIV death, and because HIV, unlike maternal mortality, also affects fathers. Also, in high-prevalence settings, HIV has more severe community-wide effects that could affect education (e.g., effects on teacher supply). In principle, variation in the individual's risk, conditional on parental or community risk, would address these problems, but such a strategy is not empirically viable with most data sets or in most settings. Fortson (2007) finds a negative effect of HIV prevalence on school attendance in Africa. She interprets the results as partly explained by direct effects from morbidity and income effects due to sick parents, but due mainly to lower expected lifetime returns to education due to mortality risk.

⁵ We investigated other countries and time periods for which MMR declined rapidly (and mortality statistics exist to study the decline). In the other cases, MMR improvements increased life expectancy by much less. For example, MMR fell dramatically in the US around 1936 after the introduction of sulfa drugs (Thomasson and Treber 2004). However, life table calculations suggest that the elimination of maternal mortality in the US could have at most increased female life expectancy by less than 0.5 years (Retherford 1972). Today MMR averages 0.4% in developing countries, with several poor countries facing rates over 1%.

⁶ We use expected years of life between ages 15 and 65 because these are the years over which returns on human capital are likely to mainly accrue. A more complete description is given later, and the exact formulae used for computations are in the data appendix.

The estimation strategy makes three main assumptions. First, the underlying behavioral model is that people's choices depend on their subjective life expectancy, and our analysis assumes that people rationally update their subjective life expectancy based on changes in objective life expectancy.⁷ The total fertility rate in Sri Lanka was 5 births over a lifetime, so MMR of 1.8% implies a lifetime risk of dying in childbirth on the order of 9% (Langford 1981). It seems likely that people would have noticed the rapid near-elimination of a large risk, particularly since complications from childbirth are a readily identifiable cause of death.

Second, the analysis compares outcomes for females and males. Our estimates would be biased if maternal mortality were correlated with other factors that had gender-specific effects on educational attainment. Maternal mortality declines in Sri Lanka were driven by greater availability of health care, improved transportation to hospitals at the time of delivery and eradication of malaria, which is a risk factor for maternal death. The estimation strategy allows for gender-neutral effects of these policies within a district-year, but the assumption is that their differential effect on females is due to their large impact on maternal mortality risk. We present suggestive evidence on the education response to other health improvements in Sri Lanka, which lends support to the assumption of gender-neutral effects. Using males as a comparison group also assumes there are no appreciable spillover effects on boys when girls' life expectancy increases. For example, effects such as a family budget constraint causing an increase in a girl's education to crowd out her brother's education are assumed to be small. Our findings suggest that if anything these spillovers cause us to underestimate our effects.

Another potential concern is that children's education can be affected when their mother is less likely to die in childbirth. In our framework, a decline today in maternal mortality affects girls' education because expectations about the girl's future maternal mortality risk are revised, and the mother is also at lower risk of dying in childbirth now. If effects of mothers' survival on children's education are gender-neutral, they do not threaten the validity of our estimates, but one might worry that there are gender-specific effects. However, as we show in the paper, the decrease in orphanhood is much too small to explain our estimates.

The paper is organized as follows. The next section describes the theoretical predictions that we test in the data. Section 3 discusses the historical context of the MMR

⁷ Hurd and McGarry (2002) have shown evidence of such updating elsewhere. There are no data on subjective expectations to verify the finding in our setting.

decline in Sri Lanka. Section 4 describes the data. In Section 5 we discuss our empirical strategy. Section 6 quantifies the effects of maternal mortality on life expectancy. Section 7 examines the effects of MMR on human capital, and Section 8 presents the effects on fertility. Section 9 concludes.

2. Conceptual framework

This paper tests the hypothesis that changes in life expectancy affect human capital investments and other choices. To lay out the empirical predictions, we present a simple model of schooling and fertility choices, and examine comparative statics when mortality rates change.

Consider a unitary household consisting of a woman and man who make two decisions, whether to have a child and how much schooling to give their child. The decisions depend, in part, on the risk of maternal mortality. For the fertility decision, the risk of maternal mortality is a cost of childbearing, and also affects the utility derived from a daughter. For the schooling decision, a daughter's maternal mortality risk will affect her returns to schooling. In the model, and also in the empirical analysis, it is assumed that a reduction in mortality risk today changes beliefs about both current risk and future risk. For the mother (older cohorts), education is predetermined but fertility is affected, and for the daughter (younger cohorts), her educational attainment is also affected.

We use a standard Mincerian model of returns to education: each year of schooling leads to a certain percentage increase in earnings.⁸ Earnings are just one, and perhaps not the most important, benefit of education for females, particularly in the context we study (Haveman and Wolfe 1984). Other potential benefits of education are that it improves her health; enables her to match with a higher “quality” husband; increases her bargaining power in the household; improves her effectiveness in using contraceptives and controlling her fertility; and improves the quality (e.g., education or health) of her children (Rosenzweig and Schultz 1989, Thomas, Strauss, and Henriques 1991, Glewwe 1999, Peters and Siow 2002).

⁸ The empirical analysis and hence the model focus on education, but the reasoning could apply to health investments as well. As one mortality risk (maternal mortality) declines for daughters, parents would have an incentive to invest in preventing other competing mortality risks or to make health investments that give a flow of payoffs throughout their daughter's life. Dow, Philipson, and Sala-i-Martin (1999) discuss theoretically how health investments respond to increased longevity and provide evidence that other health measures improved in response to vaccination availability. Oster (2007) finds that HIV/AIDS avoidance behavior is more pronounced when there is lower non-HIV mortality, e.g., due to lower malaria prevalence or maternal mortality risk.

We model earnings because it is the most standard outcome to model, but the model should be thought of as also encompassing other benefits of education that provide a stream of utils during post-schooling years.

Consider a unitary household that maximizes expected utility. The subscript w denotes wife, h denotes husband, b denotes boy and g denotes girl. The household makes a binary choice C_w about whether to have a child and then chooses the years of schooling s of the child, after observing the child's gender. The only cost of schooling is its opportunity cost in terms of foregone earnings. We treat the next generation's childbearing decision as exogenous; a daughter will have a child with probability C_g .⁹ We assume that utility is linear in consumption, that the discount rate equals the interest rate, and, for simplicity, that there is no opportunity cost of childbearing. Thus, the utility-maximizing household maximizes the present discounted value of income of the wife, the husband, and the (potential) child.

Having a child occurs at age τ in the woman's lifetime. The decision problem we model occurs at this moment of childbearing for the wife (time $t=0$). Childbearing results in the mother's death with probability μ , which is the only uncertainty in life expectancy. Conditional on surviving childbirth, the wife lives until age T_w . The husband faces no longevity uncertainty and lives until T_h . Households have a discount rate δ . The return to schooling is γ , and (instantaneous) income is y for someone with no schooling.

The household's maximization problem is

$$\max_{s_b, s_g, C_w} \left[Y_w(C_w) + Y_h + \frac{C_w}{2} (Y_b(s_b) + Y_g(s_g)) \right]$$

where

$$Y_w = (1 - C_w \mu) \int_0^{T_w - \tau} e^{-\delta t} y e^{\gamma s_w} dt, \quad Y_h = \int_0^{T_h - \tau} e^{-\delta t} y e^{\gamma s_h} dt,$$

$$Y_g = \int_{s_g}^{\tau} e^{-\delta t} y e^{\gamma s_g} dt + (1 - C_g \mu) \int_{\tau}^{T_g} e^{-\delta t} y e^{\gamma s_g} dt, \text{ and } Y_b = \int_{s_b}^{T_b} e^{-\delta t} y e^{\gamma s_b} dt.$$

Discounted utility is the integral of instantaneous income over the time period during which income is earned. For the husband and wife, income is earned from $t=0$ until death, which occurs either at $t=0$ in the case of maternal mortality or otherwise at $t=T - \tau$ (age T). For the

⁹ We make this assumption, rather than considering infinite generations, because it simplifies the model without qualitatively changing the comparative statics that the model lays out.

child, who is born at $t=0$, the income stream begins when schooling is complete, or from time (and age) s_g or s_b until death. The factor of $\frac{1}{2}$ in total utility represents the (approximately accurate) assumption that there is equal probability of having a boy or a girl. Note that the schooling level is already determined for the wife and husband. We abstract from the foregone earnings of parents who are raising a child since that would not affect the comparative statics of interest.

Working backwards, conditional on having a girl, the schooling decision is as follows:

$$\max_{s_g} \int_{s_g}^{\tau} e^{-\delta t} ye^{\gamma s_g} dt + (1 - C_g \mu) \int_{\tau}^{T_g} e^{-\delta t} ye^{\gamma s_g} dt$$

The optimal schooling level is¹⁰

$$s_g^* = \frac{1}{\delta} \left(\ln \frac{\gamma - \delta}{\gamma} - \ln \left[C_g \mu e^{-\delta \tau} + (1 - C_g \mu) e^{-\delta T_g} \right] \right),$$

yielding the comparative static that girls obtain more schooling when the risk of maternal mortality falls (μ decreases):¹¹

$$(2.1) \quad \frac{\partial s_g^*}{\partial \mu} = -\frac{C_g}{\delta} \frac{e^{-\delta \tau} - e^{-\delta T_g}}{C_g \mu e^{-\delta \tau} + (1 - C_g \mu) e^{-\delta T_g}} < 0$$

As modeled, the reduction in maternal mortality risk does not affect boys' education.

$$(2.2) \quad \frac{\partial s_b}{\partial \mu} = 0$$

Under different assumptions, one might find a positive or negative effect on boys' education. For example, if we incorporated credit constraints and extended the model to allow for multiple children per household, then higher returns to a daughter's education might crowd out her brothers' education.

Maternal mortality risk also affects the decision to have a child:¹²

$$(2.3) \quad \frac{\partial C_w}{\partial \mu} \leq 0$$

¹⁰ We assume an interior solution where $0 < s < \tau$.

¹¹ Another result is that the effect given in (2.1) is larger when mortality reductions occur earlier in adulthood (when τ is lower). We do not empirically test the cross derivative, but it is a reason why maternal mortality, which occurs early in post-schooling years, is advantageous in terms of statistical power. The earlier in productive life the mortality risk occurs, the larger the incentive effects on investment from reduced risk.

¹² The derivative for a discrete variable is of course not defined. The notation is shorthand for the comparative static that for any two values of the mortality risk μ and μ' , if $\mu' > \mu$, then holding all else equal, $C_w' \leq C_w$.

The household will have a child if the benefit (utility from the child) outweighs the cost (risk of utility loss from the mother's death). Maternal mortality risk affects this tradeoff through two channels, which operate in the same direction. First, higher μ raises the cost of childbearing because of the risk to the mother. Second, it lowers the benefit of childbearing since, if the child is a girl, she will have a shorter expected life and generate less utility for the household. This second effect illustrates an important point that fertility choice will be affected by *any* change in the expected longevity of children (Soares 2005). In our case, the change in longevity (change in μ) also applies to those making the fertility decision, conditional on having a child.

Other effects

As modeled, choosing between zero and one child presents no “quantity-quality” tradeoff. Households respond to the maternal mortality decline by (weakly) having more children because maternal mortality risk, in essence, raises the price of quantity. If one also modeled the intensive margin for fertility, households might shift from quality to quantity, having more children and educating each of their children less. This effect, though, would not necessarily change gender differentials in education, if quality falls for both boys and girls. Since the expected lifetime returns to female education increase when MMR falls, parents might shift instead from quantity to quality, having fewer children but educating each one more (Becker, Murphy, and Tanamura 1990, Galor and Weil 2000, Bleakley and Lange 2006).¹³ This incentive would create an offsetting effect to the higher fertility induced by greater maternal survival modeled above. But again this would not necessarily imply differences in how parents educate boys versus girls.

3. Historical context of the MMR decline in Sri Lanka

Sri Lanka's reduction in maternal mortality was driven by several policies related to health. Improved access to maternal health care services is commonly cited (World Bank 2003). The number of hospitals, clinics and health centers rose considerably, and many of these facilities were specifically for maternal and child care. The number of trained birth attendants also increased. Importantly, most of the services were provided for free.

¹³ Bleakley and Lange (2006) find that fertility declined in response to the eradication of hookworm in the U.S. South, consistent with a shift from quantity to quality.

Transportation to health facilities was also improved: a system of free ambulances was developed, and if ambulances were unavailable, then transportation in cases of emergencies would be reimbursed by the government (World Bank 2003). Sri Lanka also adopted new technologies from the West such as sulfa drugs and penicillin (Loudon 2000a, 1991, 1988, Paxton et al 2005, Van Lerberge and De Brouvere 2001). Figure 1 shows Sri Lanka's increase in ambulances, health centers, government midwives, and hospital beds. The proportion of women delivering at health clinics or hospitals, rather than at home, increased from 20% in 1945 to 55% in 1960, suggesting that access to care indeed improved (World Bank 2003).

The expansion of health care services was funded mainly from the central government budget, and it was particularly rapid in the post-WW2 years because resources had accumulated during the war (World Bank 2003). As governing power was transferred to Sri Lanka from the British over the 1930's and 1940's, health as a share of government expenditure increased, while spending on police and defense declined, consistent with the high priority given to social welfare for which Sri Lanka is known (Jones 2004). The emphasis on basic care and health clinics was influenced by the Rockefeller Foundation, which was involved in the campaign against hookworm. The Foundation urged Sri Lanka to invest in basic care, based on its success with this system in the southern United States (Jones 2004). In terms of the specific focus on maternal health, in 1940 a special Ministry of Health commission on infant and maternal health was established. The release of the commission's report in 1943, coincidentally just after the appointment of a new Minister of Health, pushed maternal and infant health high on the priority list for policy makers (Jones 2004).

Another key factor that contributed to MMR declines was malaria eradication. In 1945 Sri Lanka began a successful malaria control program centered on DDT spraying, and malaria death rates fell sharply. Reports on maternal mortality at the time linked malaria to maternal mortality (De Silva 1943). Malaria is thought to cause anemia and increase the likelihood of death from hemorrhage at birth (WHO 2007). As shown in Appendix Table 1, malaria rates explain much of the cross-sectional variation in MMR in 1946. The general level of economic development, as proxied by male literacy or urbanization rates, explains little of the initial

cross-sectional variation in MMR. There are no district-level data on health facilities with which to test the importance of these factors.¹⁴

The improvements in MMR were larger in places with initially higher levels of MMR. This pattern is shown in Figure 2, which plots for Sri Lanka's 19 districts the decline in MMR between 1946 and 1953 versus the 1946 level. (The data will be described in detail in the next section.) The slope of the relationship is -0.7. This strong convergence combined with the across-the-board declines in MMR imply that the initially-high-MMR districts essentially caught up to the initially-low-MMR districts by 1953.¹⁵

Sri Lanka during 1946-1953 also had higher educational participation and gender equality in education than most poor countries. The education system was organized into primary school (ages 5-11), secondary school (ages 12-18) and higher learning. In 1945, fees in state schools, which made up the majority of schools and were open to both genders, were abolished. There was also a shift away from English toward local languages (Sinhalese or Tamil) as the medium of instruction. School enrollment increased between 1946 and 1953: the proportion of children ages 5 to 14 in school went up from 58% to 72% (Jayaweera 1969).¹⁶ Decisions about how much education to obtain should respond to longevity only if there are returns to education that accrue over time. While no solid causal estimates of the returns to education exist for the cohorts we study, Mincerian estimates for Sri Lanka suggest a return to a year of education of 7% for both males and females (Psacharopoulos 1994). Consistent with the hypothesis that education also has benefits outside of the labor market, unreported OLS regressions using the 1987 Demographic and Health Survey for Sri Lanka suggest that more education for a woman is associated with marriage to a more educated man and with lower infant mortality among her children.

¹⁴ The consensus in the literature is that most other factors that may affect maternal mortality (such as malnutrition, parity, or mother's age) are in fact not as important in magnitude as access to proper care at the time of delivery. Prenatal care does not seem to be a major determinant of maternal mortality because most complications at birth cannot be predicted (Maine et al 1991). Recent work is also skeptical about the relationship between nutrition and maternal deaths (Loudon 2000a, Maine 2000). Maternal mortality is highest for very young and old mothers, and it is also higher for first born and for high-parity births, so changes in the number and timing of births may affect maternal mortality. However, Trussell and Pebley (1984) calculate that, in most settings, eliminating all births by women under 20 and over 39, as well as all births parity six or higher would reduce maternal mortality by only about 25%. Thus, even large changes in fertility behavior could not explain the dramatic declines in MMR in Sri Lanka.

¹⁵ Note that each data point is a three-year running average, which should minimize measurement error and the extent to which the pattern could be due to mean reversion.

¹⁶ The enrollment rate for 5-14 year olds is available for Sri Lanka as a whole. District-gender data on enrollment is aggregated for 5-24 year olds.

Finally, note that Sri Lanka became independent from Britain in 1948. In 1931 self-governance was instituted, so the transition to independence was peaceful and the ruling politicians continued to hold power after independence. Therefore, although independence was of course momentous, it was not associated with dramatic shifts in social policy in 1948 (Peebles 2006). For the purposes of our empirical analysis, which uses district-gender-year variation, countrywide effects of independence are not a confounding factor, nor are district-level or gender-level effects. The same point applies to other national events that occurred, such as the changes in education policy mentioned above.

4. Data and descriptive statistics

The data for the analysis were collected primarily from annual Vital Statistics reports and the Census of Population for 1946 and 1953. The data are disaggregated geographically into 19 districts. Sri Lanka's vital statistics are based on a registration system and have been shown to be close to complete (United Nations 1976, World Bank 2003). Vital statistics data on the maternal mortality ratio (ratio of maternal deaths to births) are available by district and year from 1941 onwards.¹⁷ MMR data for 1925 to 1939 (which we do not use in the empirical analysis but show in Figure 3a) are from De Silva (1943) who calculated them from vital statistics data. Figure 3a shows MMR by district and Figure 3b, MMR averaged across districts, from 1925 to 1964. The paper analyzes the period from 1946 to 1953. As can be seen, MMR declined sharply after 1946, with most of the decline occurring by 1950. A test for a trend break in MMR between 1930 and 1960 identifies 1947 as the break year.¹⁸ Also evident is a spike in MMR in 1946, thought to be caused by a malaria epidemic that year. To ensure that the trend break in 1947 is not spuriously created by the spike, we repeat the trend break test excluding 1945 and 1946 (or only 1946), and again the data choose a break in 1947. We also estimate trend breaks for each district. For 11 out of 19 districts the data identified a break in the 1946-1950 period (or 12 out of our 19 districts when we exclude the years corresponding to malaria epidemics).

¹⁷ MMR data are also in the 1940 Vital Statistics but we were unable to obtain the 1940 Vital Statistics.

¹⁸ Using district-year level data from 1930 to 1960, we regress MMR on a continuous year variable, a year dummy, and the interaction of year*dummy. We repeat the procedure, allowing the year dummy to vary from 1939 to 1953. The highest R-squared is obtained for 1947 (even if we drop 1935 and 1946 or 1935, 1945 and 1946, years with malaria epidemics). The spike in MMR in 1936 is also attributed a malaria outbreak.

For variables available annually, the regression analysis uses three-year running averages (e.g., for 1946, we average 1945 to 1947 and for 1953 we average 1952 to 1954) to reduce measurement error. Table 1a shows that the mean of three-year-averaged MMR is 1.80 maternal deaths per 100 births or 1.8% in 1946 and 0.5% in 1953. We also use MMR lagged by 3 years, which also declined by 1.3 percentage points over the study period.

Total deaths are available by district, year, gender and age (5-year groups). By matching deaths to interpolated population counts from the Census, we construct age-specific death rates, shown in Table 1b. Death rates exhibit the usual age profile, with high infant mortality followed by decreasing mortality until age 40. The ratio of female to male deaths is the largest for ages 15 to 45, a noteworthy pattern since these are the childbearing years. Mortality rates at all ages fell significantly during the period. The mortality decline in combination with an increase in the birth rate resulted in a population increase from 6.7 million in 1946 to 8.3 million in 1953.

The mortality data are believed to be of excellent quality and allow us to construct life expectancy measures for each district-year-gender (World Bank 2003).¹⁹ We construct measures of life expectancy at age a censored at age b , which we denote as $e(a-b)$ following the notation in demography. For example, our main measure is $e(15-65)$, which is life expectancy at 15 censored at 65, or the expected years alive between ages 15 and 65, conditional on survival until 15. We censor life expectancy at age 65 because death rates in the early vital statistics are not reported for older ages, and life expectancy calculations are very sensitive to assumptions about the distribution of deaths among those censored.

The life expectancy measures are calculated in the standard way from age-specific mortality rates, measured at a given point in time. The data appendix describes the procedure in detail. Ages 15 to 65 are a plausible span over which returns to education are earned, so we use $e(15-65)$ as the life expectancy metric to which educational decisions should be responsive. Table 1a shows that there was a large improvement in life expectancy from 1946 to 1963, with $e(15-65)$ increasing by 7.4 years for women and 6.6 years for men. Life expectancy also shows convergence across gender: the difference between men and women was 2.3 years in 1946 but 1.5 years by 1953. Published life expectancy measures by gender

¹⁹ Sri Lanka's registration system began in 1867 and is highly regarded (Levine 2007). Our life expectancy measures are considerably more accurate than country-level measures, which often are based solely on infant mortality (Deaton 2006).

for the entire country show very similar patterns (Nadarajah 1983, United Nations 1976). We also construct e(45-65) and e(0-15) to check that MMR reductions are not significantly correlated with relative female life expectancy gains outside of childbearing years.

Data on deaths broken down by cause are available from the vital statistics by district, year, and gender, but not age. Cause of death is reported in fine categories that we aggregate up to broader causes of death for use as control variables. Table 1b reports the means for the diseases used in the analysis. We arrived at the list of diseases using two distinct criteria. First, we selected diseases that *a priori* seem likely to have been affected by the major health interventions in the period, which centered on malaria, nutrition and helminths. Second, we collected cause-specific mortality rates for almost all causes of death, and then selected the diseases with the highest death rates among school-age children, under the assumption that diseases with high mortality rates also had high morbidity rates.²⁰ The two criteria yielded almost the same set of diseases; these diseases—malaria, vitamin deficiencies, diarrhea, helminths, and anemia—are used as control variables.²¹

Vital statistics also report the total number of births by district and year. Birth registration was almost 100% complete (United Nations 1976). As shown in Table 1a, the birth rate (births per 1000 females ages 15 to 45) increased from 179 to 202 between 1946 and 1953. (The birth rate decreased from 1953 to 1963, so Sri Lanka appears to have entered its fertility transition in the late 1950s.) We also use the number of births, combined with mortality data, to calculate the infant mortality rate. The 1946 infant mortality rate was 17 infant deaths per 100 live births, and by 1953 it had fallen by roughly 50%.

Data on population, literacy, and student population are available from the census in 1946 and 1953. Literacy rates, which are for five-year age cohorts, by gender and district, are the main educational outcome we examine. A person was considered literate if “the person is able to write a short letter and read the reply to it.” As seen in Table 1a, literacy rates rise at young ages, which reflects the fact that a child’s likelihood of becoming literate (almost

²⁰ We collected data on the causes of death that were large in 1946 and for which consistent series could be obtained. These constitute 78% of all deaths in 1946. For selecting which diseases were most common for children, we use data from 1950, which is the earliest year with cause-specific mortality rates by age (for the country as a whole).

²¹ The one exception is pyrexia which was a leading cause of death of school-age children. Pyrexia is a catch-all category: the decedent had a fever and the cause was otherwise unknown. We exclude pyrexia as a covariate because maternal mortality was frequently classified as pyrexia; puerperal pyrexia is one of the common causes of maternal death (De Silva 1943, Loudon 2000b, Deneux-Tharaux et al 2005).

necessarily) increases as she becomes older (age effects). Literacy rates then fall for older ages, reflecting a secular increase in literacy across birth cohorts (cohort effects). The turning point is a good way to infer the ages at which people are on the margin of becoming literate, a piece of information that is important to our empirical analysis. The turning point for males is 25-29 years and, for females, 20-24 years, suggesting that people become literate up until age 24 or age 19. It may seem surprising that teenagers are still becoming literate, but in many developing countries, a child might attend school for several years without becoming literate, and many children start school late and interrupt their schooling.

The empirical analysis tests for effects on “treated” age cohorts. We define the treated group as individuals who were ages -2 to 11 in 1946, just prior to the decline in MMR. These individuals satisfy several conditions. First, they are young enough to still become literate, as discussed above. Second, they are old enough that we observe their literacy in 1953; they are ages 5-19 then, and literacy is not measured in the Census for those under age 5. Lastly, most individuals in this group are very unlikely to give birth during the 1946-1953 period; the fertility rates for ages below 15 was almost 0 and was very low for ages 15-19, which is consistent with the mean age of marriage being 21 for females in 1946 (United Nations 1976). For the 5-19 year old cohorts, male literacy was 58% in 1946 and 69% in 1953. For females, it was 44% in 1946 and 58% in 1953. Female literacy increased by 3 percentage points more than male literacy during 1946-53.

The census did not ask directly about school enrollment. Published statistics instead report the fraction of 5 to 24 year olds that reported their occupation as “student” (statistics are not available for finer age categories). We use this “percent in school” as a proxy for enrollment. The percent in school increased more for males than females, in contrast to literacy which increased more for females. This pattern most likely reflects relative female gains in primary-school enrollment and relative male gains in secondary and post-secondary enrollment.

Some variables of interest are unfortunately unavailable. Years of education or other measures of completed schooling were not recorded in the 1946 and 1953 censuses. To measure how health investments respond to life expectancy, data on height or vaccination rates would have been valuable, but we have been unable to obtain them. Further breakdowns of the data within districts, either geographically or by characteristics besides age and gender,

were not published. Lastly, we attempted to obtain district-level data on education and health services, such as the number of schools, or the number of hospitals and ambulances. These data, however, are not available for the period we study.²²

We supplement the main analysis, which uses the 1946 and 1953 censuses, with estimates based on a 10% sample of the 1971 census. Table 1c reports descriptive statistics for the 1971 data, aggregated into cells defined by district, gender and 5-year age group. The main advantage of the 1971 census is that it has data on educational attainment. One limitation of using these data is that one cannot conduct a before-after comparison using the single cross-section. Instead, the analysis makes a cross-cohort comparison. For the treated cohorts (ages 5 to 19 in 1953), females had on average 5.7 years of schooling, and males, 6.3 years. For the control cohorts (ages 25 to 45 in 1953), average education for females was 4.2 years and for males, 5.1 years. We also have data on literacy, which can be compared to literacy rates reported in the earlier censuses. For the treated cohorts, literacy rates reported in 1971 are higher than those reported in 1953, most likely reflecting the fact that some in this group had not completed their schooling in 1953. For the control cohorts, literacy rates are somewhat lower than those reported in 1953 especially for women.

A serious limitation of the 1971 Census is that the schooling data are missing for 23% of observations (whereas literacy is never missing). The potential for sample selection bias is thus large. Another potential problem is that there is a significant amount of attrition between 1953 and 1971, which could be due to mortality, international migration, and census quality. Also, for 1971 we match individuals to their district of birth, whereas the 1953 data are tabulated by district of residence, so immigrants are excluded from the 1971 data but are included in the 1953 data. Table 1c reports the “survival rate,” constructed as the ratio of the number of enumerated individuals in 1971 to the number of enumerated individuals in 1953. Survival rates average 78.5% and vary both by gender and age. We address in the empirical analysis the potential selection bias that missing data and attrition introduce.

5. Empirical strategy

Our empirical strategy uses differential declines in maternal mortality across districts in Sri Lanka as a source of variation in life expectancy. This approach is a difference-in-

²² Gray (1974) reports that during the post-war period, “with the exception of the innovations in maternity homes, there is no evidence for an unbalanced improvement [across districts] in medical services.”

difference-in-difference (DDD). The first difference is over time, since maternal mortality fell between 1946 and 1953. The second difference is across geographic areas; the magnitude of the MMR declines varied considerably across Sri Lanka's 19 districts. The third difference is between genders; maternal mortality is quite unique among major causes of death in that it exclusively pertains to women.

The data comprise 76 observations corresponding to a gender (2), district (19) and year (2). Consider life expectancy, denoted by e , as the dependent variable. The estimating equation is

$$(5.1) \quad e_{dgt} = \beta_0 + \beta_1 MMR_{dt} * female_g + \mu_{dg} + \gamma_{dt} + \nu_{gt} + \varepsilon_{dgt}$$

where d denotes district, t year and g gender. The coefficient of interest is β_1 , which measures the effect of MMR on women relative to men, using variation across districts in how much MMR declined between 1946 and 1953. The specification includes a full set of double interactions, namely district-gender (μ_{dg}), district-year (γ_{dt}) and gender-year fixed effects (ν_{gt}), and ε_{dgt} is a random disturbance term. Since MMR differs in 1946 and 1953, we do not need to include a dummy for "post" to obtain the DDD estimate. The identifying assumption is that there are no unobserved district-gender specific changes that (1) are correlated with changes in maternal mortality in the district and (2) are correlated with district-gender specific changes in the outcome.

We examine two distinct types of outcomes: those that MMR affects mechanically, such as life expectancy; and those such as educational attainment and fertility that MMR is hypothesized to affect through behavioral responses. For the mechanical effects, MMR should have immediate effects and thus the specification uses current MMR, while with behavioral outcomes, the correct specification will depend on how quickly people update beliefs and change their behavior. We allow for a three-year lag between the declines and when people respond, and we test the sensitivity of our results using other lags.²³ Another important reason for using a lag is to minimize the potential for reverse causality. The 3-year lag also implies that the variation in MMR used excludes the 1946 spike.

²³ A practical consideration is that our second observation is for 1953, so to observe effects of the steep MMR declines that occurred from 1946 to 1950, three years is the maximum behavioral lag for which our data allow us to observe the full effects. In principle, people could have anticipated the MMR decline in which case a lead would be appropriate, but it seems unlikely that the declines were anticipated.

We have age-specific literacy data, so we can examine whether MMR has effects specifically on age cohorts who during 1946-53 are in the age range when people become literate. Recall that we defined this “treated group” as individuals who are ages -2 to 11 in 1946 (ages 5-19 in 1953), or those who are (1) young enough to respond to MMR declines and become literate, (2) old enough that their literacy is observed in 1953, and (3) are unlikely to contribute to the declines in MMR through their fertility. For this group, we estimate the following regression, where the unit of observation is a district, gender, year, and 5-age category:

$$(5.2) \quad literacy_{atdg} = \beta_0 + \beta_1 LaggedMMR_{dt} * female_g + \mu_{dg} + \gamma_{dt} + \nu_{gt} + \varepsilon_{atdg}$$

β_1 gives the effect of MMR on females relative to males, in 1953 relative to 1946. The specification includes a full set of double interactions of district, gender and year, namely district-gender (μ_{dg}), district-year (γ_{dt}) and gender-year fixed effects (ν_{gt}), and ε_{atdg} is a random disturbance term. The regression also includes double interactions between age and district, year and gender, which are omitted in (5.2) for brevity. The hypothesis is that $\beta_1 < 0$, or that when MMR falls, female literacy sees relative increases. We also estimate regressions controlling for year-, gender and district-specific mortality rates from diseases that affected children. Ideally we would control for morbidity rates, but these data are not available. Using mortality as a proxy is based on the assumption that the gender difference in morbidity effects is correlated with the gender difference in mortality effects.

As a falsification test, we also verify that MMR has no effect on older cohorts—ages 25 to 44 in 1953—whose literacy was already determined at the start of the study period.²⁴ A significant coefficient for the older ages would imply that either there is reverse causality (if, for example, female education lowered MMR) or there are pre-existing trends in literacy correlated with MMR declines. We also estimate a DDDD regression using the older cohorts as a fourth difference, where the coefficient of interest is for *laggedMMR*female*treated*, and all sets of triple interactions are included.

6. Effect of MMR on life expectancy

This section quantifies the impact of MMR reductions on female life expectancy. This exercise is useful later for interpreting the magnitudes of MMR’s effect on education.

²⁴ We exclude ages 45 and above since their literacy rates appear to be affected by selective mortality.

Examining life expectancy also allows us to test for omitted variable bias by examining whether MMR declines are correlated with relative female gains in life expectancy outside of childbearing ages. Figure 4 graphically previews the results. For each district, the relative female gains in $e(15-65)$ are plotted versus the 1946 level of MMR. As predicted, female life expectancy improved relatively more in districts where MMR was initially high, or where MMR declined the most.

One can quantify the effect of MMR declines on $e(15-65)$ using the standard demographic method, which is to calculate what $e(15-65)$ would have been if MMR declined by a given amount, but all other death rates remain unchanged. Maternal deaths declined by 67% from 1946 to 1953. We do not know the age distribution of these deaths, except that they occurred between ages 15 and 44, and we know that they accounted for 26% of deaths in 1946 for this age group. If one recomputes $e(15-65)$ for 1946 with mortality rates for women ages 15 to 44 reduced by 17% (.67*.26), one finds that declines in MMR increased female life expectancy by 1.4 years (or 4%).

Another way to quantify the effect of maternal mortality on life expectancy is to estimate equation 5.1. The direct calculation above must make assumptions about MMR by age and about independence between MMR risk and other mortality risks, whereas the regression analysis does not. More importantly, regression analysis allows us to probe for omitted variables by testing for an “effect” of MMR on life expectancy outside of childbearing ages. The results are reported in Table 2. The first column shows the results from the main specification. The effect of MMR on relative female $e(15-65)$ —the life expectancy measure that MMR *should* affect—is negative and significant. When MMR fell, life expectancy rose. Since MMR fell by 1.3 percentage points between 1946 and 1953, the estimate implies that MMR declines resulted in an increase in female life expectancy of 1.5 years. Reassuringly, this estimate is very close to the direct calculation of 1.4 years presented above. Female $e(15-65)$ increased by 7.4 years over the period, so MMR declines can explain about 20% of this increase. Female $e(15-65)$ increased by 0.8 years more than male $e(15-65)$ from 1946 to 1953, which is less than the female-male convergence predicted by maternal mortality declines. There seem to have been other factors causing relative male life expectancy gains, and absent them, MMR declines would have enabled women to catch up even more in terms of life expectancy.

Specification checks

The next rows show the effect of maternal mortality on other measures of life expectancy. The coefficient of $MMR * female$ on $e(45-65)$ is close to zero and insignificant, and for $e(0-15)$ it is marginally significant but small (about 5 percent of the change in female $e(0-15)$ during the period). Both of these age ranges are outside of the primary childbearing ages, so large negative and significant effects would have suggested other female-specific health gains that were correlated with MMR declines.

Next, we examine the correlation between MMR and the infant mortality rate (IMR). There are at least two distinct reasons that IMR might be correlated with MMR. First, health programs for mothers and children plus the malaria eradication efforts that contributed to MMR declines are likely to have improved IMR. Second, maternal mortality could have a causal effect on infant mortality, since motherless infants may be more at risk. The primary concern for our identification strategy is a correlation between MMR and *gender differentials* in IMR. Nevertheless, it is helpful to begin by discussing the correlation between MMR and total IMR. In a regression of changes in IMR on changes in MMR (N=19), with both variables measured in percentage points, the coefficient is 3.73 and statistically significant at the 1% level. A 1 point reduction in MMR is associated with a 3.7 point reduction in IMR. A back-of-the-envelope calculation suggests that the upper bound on what the *causal* effect of MMR on IMR could be is 0.8, implying that much of the correlation between MMR and IMR is due to third factors such as health care improvements.²⁵ This calculation uses the upper bound in the literature that IMR could be as much as 6 times higher for motherless infants as for infants with mothers (Loudon 1991). Note, though, that the key issue is whether health programs had gender-specific effects, and whether maternal mortality presents a higher mortality risk for female infants. Table 2 shows the effect of MMR on IMR for girls relative to boys. The coefficient on $MMR * female$ of 0.13 is small and insignificant.²⁶

One potential threat to the research design is that declines in MMR could be correlated with other health improvements that affected outcomes such as literacy. For example,

²⁵ A 1 point decline in MMR would decrease the number of motherless infants by 1 percentage point. One can use the IMR ratio of 6 to back out the initial-period IMR for those orphaned due to MMR (93.6%) and non-orphans (15.6%) that gives back the population-average IMR, and then simulate the change in IMR from a 1 point decline in the proportion of orphans.

²⁶ The 0.13 percentage point decrease in relative female IMR that is associated with a 1 percentage point decline in MMR represents less than a 1% effect.

improved maternal and child health programs is one reason that MMR declined, and a concern is that these programs directly improved child health. The strength of the identification strategy, in this regard, is that such improvements likely helped both boys and girls, and we identify MMR effects based on differential improvements for girls. Another potential confounding factor is the successful malaria control program. We are aware of two other programs that targeted health and may have reduced morbidity among children, nutrition programs, such as free milk distribution, and intestinal worm eradication. We address this concern by controlling for gender- and district-specific malaria and nutrition-related death rates (we refer to anemia, diarrhea, vitamin deficiencies and helminths jointly as nutritional diseases hereafter; our controls include gender- and district-specific death rates for each of these diseases separately).

These controls address potential omitted variable bias, but they are not our preferred specification for three reasons. First, the control variables could be endogenous in some specifications. For example, school enrollment rates could determine nutrition-related diseases since the government provided food in school. Second, we could be over-controlling by including these diseases because malaria and nutrition deficiencies could increase the likelihood of maternal deaths. Third, some of the control diseases, especially helminths and vitamin deficiencies, predominantly cause under-5 mortality (based on post-1950 data when the breakdown by age becomes available), so the mortality rate aggregated across ages is a far from an ideal proxy for post-age-5 mortality and morbidity. Nonetheless, we view the results with these controls as useful checks. As seen in columns 2 to 4 of Table 2, the controls do not have an appreciable impact on the coefficients for e(15-65). The coefficients for e(0-15) and IMR are somewhat more sensitive to the inclusion of the control variables, which is sensible since these diseases were chosen because they are prevalent among children.

More detailed results are presented in Appendix Table 2, which examines the change in age-specific death rates. $MMR * female$ is positively and significantly associated with death rates for ages 15-19, 20-24, 25-29 and 30-34 year-olds, with the largest effects at age 20-24, which is consistent with birth rate patterns. The effects on death rates below age 15 are sometimes statistically significant, but the magnitudes are small. The results for ages 35 and older are small and insignificant except for one case.

Maternal mortality resulted in an increase in female $e(15-65)$ of 1.5 years between 1946 and 1953. A helpful way to think about this change in time horizon is to calculate the change in a constant hazard rate of mortality that would result in the same change in life expectancy. The constant hazard rate that results in the observed value for $e(15-65)$ in 1946 is 1.10%. Lowering the hazard rate to 0.93% yields the observed 1.5 year increase in $e(15-65)$. Thus, declines in MMR can be thought of as equivalent to a 0.17 percentage point decline in the mortality rate (a 15% decline). In other words, the behavioral responses to the decline in MMR should be the same as if there were a 0.17 percentage point decrease in an individual's discount rate (Preston 1980).

7. Effect of MMR on literacy and schooling

We now turn to the effect of MMR on literacy and schooling. Figure 7 plots the change in the female-male gap in literacy for the treated cohorts versus initial MMR for each district. The treated age cohorts are those whose literacy should be affected by the 1946-53 MMR declines, namely those ages 5 to 19 in 1953. Table 3 presents the regression results. Panel A shows the DDD results for the treated ages. Column 1 is the basic regression with no covariates besides district*year, district*gender, gender*year, age*year, age*district, and age*gender fixed effects (all sets of double interactions). Each observation is a district-gender-5 year age group, and standard errors are clustered at the district-gender level. We use the 3-year lag of MMR to allow for a delay before people notice the mortality decline and adjust their behavior. As predicted, declines in MMR (health improvements) increase female literacy. The coefficient on *laggedMMR*female* is -0.87, with a p-value of 0.06. Controlling for district-, time-, and gender-specific mortality rates from other diseases (column 2) increases the coefficient of interest, suggesting that if MMR declines were correlated with other health interventions, these interventions tended to disproportionately benefit boys rather than girls.

Column 3 presents an alternative specification where MMR changes are instrumented using the 1946 level. The reasoning behind this specification is that it uses initial conditions (pre-period MMR) combined with the timing of the intervention as exogenous, and exploits only the fact that areas with high initial MMR benefited disproportionately from the interventions. As seen in Figure 2, this explains most of the variation in MMR declines, so not

surprisingly, the IV specification gives very similar point estimates to those in column 1: the coefficient is -1.0 and statistically significant at the 5% level.

Column 4 drops the district*year fixed effects and controls for male e(0-65).

Compared to column 1, the precision falls considerably but the magnitude of the coefficient is similar (-.88 versus -1.07). The purpose of this specification is to shed light on the validity of the fertility estimates presented later. To the extent that controlling for male e(0-65) versus district fixed effects gives similar results for literacy, this increases confidence about controlling for male e(0-65) as a way of arriving at causal estimates of the effect of MMR on the birth rate, where district*year fixed effects are impossible to include. Controlling for male e(0-65) seems to do a good job of absorbing district-year-level omitted factors.

Panel B presents identical results for cohorts whose education was completed prior to the MMR declines (ages 25 to 44 in 1953, i.e., 17-37 in 1946). The coefficient on the control cohorts, which is hypothesized to be zero, is statistically insignificant. The coefficient remains insignificant across specifications and is always substantially smaller than the estimated coefficient for the treated ages. Columns 5-7 show additional specifications that use older cohorts in the estimation. In column 5 we estimate a DDDD model where the control cohorts serve as a fourth difference. If there are important gender-specific trends in education that are correlated across districts with MMR declines, the older cohorts will account for them. The coefficient is statistically insignificant in this specification but the point estimate is very similar to those in columns 1-4.

In columns 6 and 7, we present a DDD separately for women and men to determine the extent to which our use of males as a control drives the result. The DDD estimate for women only is -1.3. If MMR declines are correlated with other changes within districts that increased education, or if there are positive spillovers from women to men, then the male DDD coefficient will be positive. On the other hand, if there are negative spillovers (for example, credit-constrained families educate boys less in order to educate their girls more) then the coefficient on males would be negative. Of course, the possibility of district-year level omitted variables, which motivates the identification strategy of using males as a control group in the first place, should be kept in mind when interpreting these results as measuring spillovers. The coefficient on *laggedMMR* for the males-only specification is negative and insignificant. It is also smaller than that for women. This suggests that using males as a

control is either helping us control for omitted variables, or is biasing our estimates toward zero because of positive spillovers. Thus our main estimates would appear to be, if anything, underestimated.

To be conservative, we use the coefficient from Panel A, Column 1 when assessing the effect size. (Most of our estimates are larger.) The coefficient of -0.87 implies that the 1.3 decline in MMR between 1946 and 1953 led to an increase in relative female literacy of 1.1 percentage points. The female-male literacy gap among 5-19 year olds was 14.3 percentage points in 1946 and narrowed to 10.9 percentage points by 1953. MMR declines explain a third of the relative gains that girls made. The results can also be interpreted as an elasticity. Female literacy among 5-19 year olds was 44% in 1946, so MMR declines led to a 2.5% increase in female literacy. As shown in the previous section, MMR led to a 4.1% increase in female $e(15-65)$. The implied elasticity of literacy with respect to $e(15-65)$ is 0.6.²⁷

In Panel C of Table 3, the dependent variable is the percent in school among those ages 5 to 24. More precisely, the variable is the percent who report “student” as their occupation. This is not an ideal measure of completed education, but it is the only variable in the 1946 and 1953 censuses from which one can infer the effect of MMR declines on years of schooling. *LaggedMMR*female* has a negative and marginally significant effect on the percent in school. The coefficient is -0.9, implying that Sri Lanka’s 1.3 decline in MMR increased the probability of attending school by 1.2 percentage points or 3.5%. This estimate is considerably smaller than the lowest cross-country estimates in Lorentzen et al (2005) which would imply a 10% change in enrollment. Multiplied by the 20-year age range covered, the estimate implies that years of schooling would increase by 0.23 years. When disease rates are added as covariates, the coefficient becomes insignificant, although its magnitude remains similar, and in the IV specification, the coefficient is slightly larger.

Table 4 presents the literacy results for each 5-year age group. (School enrollment is unfortunately not broken down by age). The coefficients are less precise, but the patterns are similar. MMR declines are associated with female literacy gains for young cohorts but not older cohorts, except those ages 20-24 for whom the coefficient is negative and significant.

²⁷ Another way to assess the estimated magnitude is to recall that the decline in MMR is equivalent to a 0.17 percentage point decrease in the discount rate and to ask whether the change in literacy observed implies a reasonable distribution of discount rates. Under the assumption that the distribution is lognormal, one can back out that the mean discount rate would be 2.4 % and the median, 8.4%, which is consistent with estimates of the discount rate found in the literature (for example see Frederick, Lowestain and O’Donoghue 2002).

Reverse causality is a likely explanation. These cohorts are too old to have become literate in response to the 1946-53 declines in MMR, but the MMR declines may have been affected by how educated they were, since they were in their childbearing years.

Specification checks

Table 5 presents other robustness checks for the literacy results. In column 2 we correct the literacy rate for migration across districts, which is potentially a problem since internal migration rates are high and many individuals migrated to areas where malaria was eradicated. The appendix describes in detail the migration correction, which accounts for gender gaps in literacy among migrants and among non-migrants (the former gaps are larger). The migration-corrected coefficient is larger and more precisely estimated than in the basic specification. Column 3 drops the 2 districts that appear to be outliers in terms of initial MMR. Although the estimates are imprecise, the point estimate remains around -0.9. (In Appendix Table 3 we test for the influence of outlier districts, running the basic specification dropping one district at a time. The results are not overly sensitive to any one district.) In column 4 we use the 1953 population for each cell as weights, which results in a larger coefficient. Columns 5 and 6 restrict the treated group to ages 5-14 in 1953. While this narrower definition will not capture effects on those who were older than age 7 in 1946, it rules out any possibility of reverse causality. Relative to the estimates with 5-19 year olds, the coefficient on *LaggedMMR*female* is smaller in magnitude in the specification without disease controls and almost identical in magnitude with disease controls. Finally the table shows the results for different lags where the MMR regressor is a single-year measure, rather than the 3-year average. Overall these specification checks confirm our main results.

In summary we find that literacy increased by 1 percentage points (or 2.5%) in response to a 1.5 year (4.1%) increase in adult life expectancy. In addition, it appears that years of schooling increased by 0.23 years (3.3%).

1971 Census results

We next present estimates using the 1971 census, using the following specification:

$$(7.1) \quad education_{dga} = \beta_0 + \beta_1 LaggedMMR_{da} * female_g + \mu_{dg} + \gamma_{da} + \nu_{ga} + \varepsilon_{dga}$$

Education is the average education (literacy or years of education) of individuals born in district d , of gender g , in age cohort a . An observation is a district, gender and five-year age

group. All double interactions are included and standard errors are clustered at the district-gender level. This regression is a DDD estimation, but it differs from our main specification in an important way: the comparison is across gender, district and *cohorts*, rather than across gender, district and *time*. We define treated and control cohorts as before; those ages 5-19 in 1953 are treated, and those ages 25 to 44 in 1953 are controls. We assign the 1946 MMR value to the control group and the 1953 MMR value to the treated group, making the specification as comparable as possible to our previous specification. (In other words, *LaggedMMR* takes on only two values within a gender-district).

Before estimating the effects of MMR on education, we first investigate two potential problems with the data: attrition and missing data for education. The main issue is whether, conditional on the set of double interaction controls, *LaggedMMR*female* predicts attrition or the percentage of observations with missing education. We first estimate equation 7.1 using the “survival rate” as the dependent variable. Column 1 of Table 6 shows that *LaggedMMR*female* is not significantly related to attrition. The magnitude is small: the decline in MMR of 1.3 percentage points that Sri Lanka experienced between 1946 and 1953 would increase attrition by less than 1% of the mean survival rate in the sample. Column 2 shows that literacy rates in 1953 are also not significant predictors of attrition, conditional on the covariates. This set of results suggests that attrition is unlikely to bias our results, and we ignore it in the estimation.

Next we look at whether *LaggedMMR*female* predicts whether education is missing in the 1971 census (literacy is never missing). Unfortunately, the coefficient on *LaggedMMR*female* is positive and significant (column 3). The magnitudes are large: a 1.3 percentage point change in MMR is associated with an 18% change in the proportion of treated females with missing data. In column 4, we find that illiterate individuals are much more likely to have missing education data. Even when we control for literacy, *LaggedMMR*female* continues to be significant (although its magnitude falls), suggesting that selection is not entirely explained by literacy status and that an attempt to impute missing education values based on observables (e.g., using the distribution of education within a district-gender-cohort-literacy cell), will not solve the selection problem. Individuals with missing education data are less educated on average, as seen from the negative correlation between missing data and literacy. The positive correlation between missing data and

*LaggedMMR*female* implies that missing data will bias the coefficient on *LaggedMMR*female* upwards, or toward zero. To address selection bias, our estimates for education apply a Heckman correction.

Column 5 shows the results for literacy. Since there are no missing observations for literacy, and attrition appears random conditional on the covariates, we just estimate an OLS equation as given in (7.1). We find that *LaggedMMR*female* is negative and significant, or that an improvement in MMR leads to an increase in female literacy. The coefficient of -2.7 implies that a 1.3 percentage point decline in MMR results in a 3.5 percentage point increase in literacy. If we estimate a simple DD for women only, we obtain very similar estimates. This estimate is larger than the estimate we obtained using the 1946-1953 data, but our specification is not the same: the 1971 census estimates use the differences between the older and younger cohorts, instead of the changes over time for the same age groups. In order to confirm that the difference in the estimates is driven by the specification rather than the data, we replicate the same model using the 1953 cross-sectional data only. The coefficient (standard error) for *LaggedMMR*female* in the 1953 specification is -2.74 (1.39) which is nearly identical to the 1971 estimate.

Finally, we look at the effect of *LaggedMMR*female* on years of education, which is the main advantage of the 1971 data. A naïve OLS regression yields a positive and insignificant effect, as shown in column 7. As shown above, there is considerable non-random missing data, which is likely to bias the OLS estimates, so we next estimate Heckman selection models which explicitly account for missing values of education.²⁸ These models make strong functional form assumptions (joint normality of the error terms) and also require, or at least benefit from, exclusion restrictions. We use two sets of excluded variables, that is, covariates that predict whether education is missing but are assumed to be uncorrelated with unobserved predictors of education. First, the quality of data collection likely varies by geographic region, so we use district-of-current-residence dummies as excluded variables (the specification controls for district-of-birth dummies). Second, certain individuals may be less

²⁸ The parameters estimated for the Heckman models confirm that *LaggedMMR*female* is positively correlated with missing data, and individuals with missing education data have more education, implying that the OLS coefficient on *LaggedMMR*female* is upward biased. Note that we use the individual census data to estimate these models. To make the results comparable to the aggregate results (where each cell is an equally-weighted observation), we weighted each individual observation by 1/N, where N is the number of observations in that observation's cell.

likely to provide responses in the census, and we use three dummies for whether religion, ethnicity and district of residence are missing as the excluded variables. The results are in the last two columns of Table 6. Here we find that the coefficient on *LaggedMMR*female* is negative and significant. The coefficient of -0.133 implies that a decline in MMR of 1.3 increases education by 0.17 years of schooling, which is a 4% increase relative to the mean for women in the control group (4.2 years of schooling).²⁹ This estimate is very close to that implied by our results using percent in school in Table 4, where we found that an MMR decline of 1.3 increased schooling by 0.23 years.

These results suggest that schooling increases by 0.11 years per additional expected year of life (schooling increased by 0.17 years in response to a 1.5 year increase in $e(15-65)$). In percentage terms, a 4.1% increase in $e(15-65)$ led to a 4.0% increase in years of schooling, giving an elasticity of 1.0.³⁰ With the caveat that missing data requires us to correct for selection and impose additional econometric assumptions, these results suggest that people respond to longer life expectancy by obtaining more schooling.

Potential threats to validity

This subsection discusses other potential threats to the estimation strategy. The first is related to gender-specific elasticities of human capital with respect to health. The declines in MMR were driven by health interventions (e.g., malaria eradication, expansion of health facilities) that might directly affect boys' and girls' health. These other interventions likely had the same health effects for girls and boys, and, as measured by mortality rates for malaria and helminths, for example, they did. However, one concern is that the responsiveness of education to health improvements might differ by gender. If a given improvement in health has a larger effect on the education of girls than boys, then our results could be driven, not by the differentially larger improvement in health for girls due to MMR declines, but by their differential sensitivity to health improvements. Note that this different elasticity by gender could operate through morbidity affecting school-going, life expectancy effects, or even improvements in other people's health, for example, family members'. While we cannot

²⁹ We also computed bounds for the effect of *MMR*female*. Bounds are attractive because they make no distributional assumptions and require no exclusion restrictions, but they tend to be wide. Our most tightly estimated set of bounds was [-1.7; 1.8]. Unfortunately, these bounds are too wide to be very informative.

³⁰ The MMR decline can be viewed as equivalent to a 0.17 percentage point decline in the discount rate. The resulting 0.2 year increase in schooling is similar to Gan and Gong's (2004) calibration results: they report that a 0.4 percentage point change in the discount rate is associated with a 0.59 year change in schooling.

directly rule out this concern, we can test for gender-specific literacy effects for other health improvements that were common to both boys and girls. We allow the coefficient on mortality rates from a particular disease to vary by gender, and ask whether female literacy is more responsive than male literacy to a given health improvement. Appendix Table 4 shows the results, where the diseases we examine are the set of control diseases we use (malaria and nutrition-related diseases) and any other disease that had a large decline (over 50%) from 1946 to 1953. There is no strong evidence of a larger elasticity for girls. In fact, for the majority of the cases, the interaction effect for females is positive, or girls' literacy is less responsive than boys' literacy to a given change in mortality.

A related concern is that the death of a mother might differentially affect girls' education. Mothers might be more "pro-daughter" than fathers, or maternal death might cause a girl, more so than a boy, to drop out of school to take over her mother's work in the home. Under this alternative, MMR has a causal effect on girls' education, but one different from the incentive effect that arises from changes in girls' life expectancy. Back-of-the-envelope calculations suggest, however, that these effects would be small and are unlikely to account for the empirical results. Consider the following rough calculation. Relative to 1946, the risk of a mother dying per childbirth was on average 1.0 percentage point lower during the subsequent seven years up to 1953. The average number of births per woman during the seven-year interval was 1.2 births (17% fertility per year), so the likelihood of a child being motherless fell by about 1.2 percentage points ($1.2 * 1.0$). Assume an extreme scenario where no girl whose mother has died ever becomes literate, while boys' literacy is unaffected by maternal death. In this stark scenario, girls' literacy would increase by 0.5 percentage points.³¹ Our estimate of the increase in girl's literacy caused by the 1946-53 MMR declines is over twice this magnitude, so even under the starker assumptions, the effects of orphanhood cannot explain our estimated effect size.

The potential for spillovers raises the question of whether males are a valid control group. Our assumption is that MMR has a direct effect on only female life expectancy and, in turn, schooling. However, MMR could have a spillover effect on males if, say, men prefer to be more educated than their wives, and an increase in girls' education induces an increase in boys' education (positive spillover). Conversely, a higher return to education for girls might

³¹ The population average literacy for the relevant cohorts of girls was 44%, and $0.44 * 1.2 = 0.5$.

cause credit-constrained families to shift resources away from boys' education (negative spillover). Unfortunately, we cannot empirically assess the relative magnitude of direct versus spillovers effects, but the results in Table 3 are suggestive that spillovers are positive, leading us to underestimate the magnitude of life-expectancy effects.

In addition, Sri Lanka as a whole underwent other changes during the study period. Education-related policies, such as the elimination of school fees, are especially of interest since education is the outcome we examine. If the national education policies had gender- or district-level effects, these would not bias our estimates, which use district-gender variation. The identification only would be threatened if national policies had gender-specific effects that varied by district in a manner correlated with MMR. We are not aware of any major gender-specific or district-specific educational policy changes during the period.

Finally, the underlying behavioral model has people rationally updating their beliefs about life expectancy. This assumption is not unique to our study; it is standard in the literature. If the assumption were wrong, however, it could be one explanation for null results, but since we do not find null results, this concern becomes somewhat less forceful. When calculating the elasticity of behavior with respect to life expectancy, one needs to make the stronger assumption that individuals get the *magnitudes* of life expectancy gains correct. The calculations are still valid as an elasticity with respect to objective life expectancy, but interpreting them as with respect to subjective life expectancy requires stronger assumptions about the updating of beliefs.

8. Effect of MMR on fertility

We next test the theoretical prediction that when the risk of dying in childbirth falls, the propensity to have children increases. Because we cannot make use of gender as a third difference, we estimate a difference-in-difference model across districts and years. One control variable that might capture much of the potentially confounding district-year-specific factors is male life expectancy. Maternal mortality should not have a causal effect on male life expectancy (except through infant mortality as discussed earlier), so a correlation between MMR and male life expectancy is likely driven by unobserved factors such as overall health improvements. Using male e(0-65) as a control variable seems promising because it gets close to the female-male comparison that one can use with dependent variables such as life

expectancy or literacy that permit the DDD design. Nevertheless, the estimates for fertility are tentative because of the greater possibility of omitted variables.

The fertility results are in the first row of Table 6.³² Column 1 controls for male e(0-65), or life expectancy at birth censored at 65. The magnitude of -5 for the coefficient on MMR implies that the decline of 1.3 in MMR over the period caused the birth rate (births per 1000 women age 15-45) to increase by 6.5 from a base of 178.9. In other words, the birth rate seems to have increased by 4% in response to the 67% decrease in MMR. Column 2 shows that the estimates are stable when malaria and nutritional diseases are added as covariates. The point estimates are also stable when IMR (total for males and females) is included (Column 3), though the coefficient is no longer statistically significant. We used IMR as a control variable because it could affect the birth rate. As discussed in Section 6, IMR declines (total, not gender-specific) indeed were correlated with MMR declines. The fact that IMR is a confounding factor makes us cautious in interpreting the fertility results. But the fact that, conditional on male e(0-65), the results are robust to the inclusion of other covariates gives some reassurance. We interpret the results as suggestive evidence that when the risk of dying in childbirth decreases, individuals respond by increasing their fertility. The birth rate increased by 13% between 1946 and 1953, and lower MMR seems to account for one third of the increase. The fact that fertility responded to decreases in maternal mortality lends further credence to the assumption that the MMR declines were visible enough to affect behavior.

9. Conclusion

Over the past fifty years, longevity has improved dramatically, particularly in developing countries, with welfare gains comparable to the welfare gains from income growth (Becker, Philipson, and Soares 2005). Besides being welfare-improving per se, longevity improvements are hypothesized to have the further benefit of spurring human capital accumulation. A longer horizon gives an individual a stronger incentive to invest in education. This paper found empirical support for this hypothesis. We identified these effects using changes in maternal mortality, a cause of death that is particularly well-suited to isolating this life-expectancy channel from other effects that health has on education.

³² MMR could be affecting either the numerator or the denominator of the birth rate; the next two rows present results separately for log of births and log of female population as the dependent variables.

We examined a period of rapid MMR decline in Sri Lanka. MMR fell by 70% between 1946 and 1953, resulting in a 1.5 year increase in female adult life expectancy. We used geographic variation in this decline to estimate the corresponding change in human capital. The MMR decline caused female literacy to increase by 1 percentage point (relative to changes in male literacy), which represents a 2.5% increase. This change resulted from a 4.1% increase in female life expectancy at age 15 (censored at 65), implying an elasticity of literacy with respect to life expectancy of 0.6. Similar results were found for years of schooling. Schooling increased by about 0.2 years of school or 4%, giving an elasticity of schooling with respect to life expectancy of 1.0. Our schooling estimate is considerably less than the effect size of 0.66 years of school suggested by Kalemli-Ozcan et al's (2000) calibrations, but higher than the estimate of zero in Acemoglu and Johnson (2007). We also found suggestive evidence that MMR declines cause fertility to increase, which is as predicted, both because childbearing becomes less risky, and because having a child (daughter) is likely valued more when she is expected to live longer.

The estimates in this paper can be used to examine other contexts, such as the HIV/AIDS epidemic. In South Africa, about 40% of deaths between age 15 and 49 are AIDS-related. Lowering death rates for 15- to 49-year-olds by 40% would increase life expectancy at age 15 (censored at age 65) by 3 years, or 10.2%, based on a calculation using aggregate vital statistics data. Our elasticity estimates suggest that by shortening time horizons, HIV/AIDS reduces schooling by 0.3 to 0.5 years. (Average education is 5 years in South Africa.) A conservative estimate of the returns to a year of schooling in South Africa is 5% (Hertz 2003). Thus, the reduction in life expectancy from HIV/AIDS lowers annual income by between 1.5% and 2.5%. Through this life-expectancy channel alone, health can have sizeable effects on education and income.

The findings of this paper suggest that the increase in human capital accumulation that results from life expectancy gains is an important component of cost-benefit analyses of public health interventions. Conversely, when life expectancy declines, for example because of HIV/AIDS or civil war, the damped incentive to invest is an important cost to consider.

References

- Acemoglu, D. and S. Johnson (2007), "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy* 115(6): 925-985.
- Barro, R. and X. Sala-i-Martín (1995), *Economic Growth*, New York: McGraw Hill.
- Becker, G.S., K.M. Murphy and R. Tamura (1990), "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy* 98(5): S12–37.
- Becker, G.S., T.J. Philipson, and R.R. Soares (2005), "The Quantity and Quality of Life and the Evolution of World Inequality," *American Economic Review* 95(1): 277-291.
- Ben-Porath, Y (1967), "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75(4, part 1): 352-365.
- Bleakley, H. (2007), "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics* 122(1): 73-117.
- Bleakley, H. and F. Lange (2006), "Chronic Disease Burden and the Interaction of Education, Fertility, and Growth," BREAD working paper.
- Deaton, A. (2006), "Global Patterns of Income and Health: Facts, Interpretations, and Policies," National Bureau of Economic Research Working Paper No. 12735.
- Deneux-Tharaux, C., C. Berg; M-H. Bouvier-Colle, M. Gissler, M. Harper, A. Nannini, S. Alexander, K. Wildman, G. Breart and P. Buekens (2005), "Underreporting of Pregnancy-Related Mortality in the United States and Europe," *Obstetrics & Gynecology* 106(4).
- De Silva, M. W. M. (1943), "Report on the Investigation into Infant and Maternal Mortality in Ceylon," Sessional Paper XXIII, Colombo: Ceylon Government Press.
- Dow, W., T. Philipson and X. Sala-i-Martin (1999), "Health Investment Complementarities under Competing Risks," *American Economic Review* 89(5): 1358-71.
- Fortson, J. (2007), "Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa," mimeo, Princeton University.
- Frederick, S., G. Lowenstein and T. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, Vol XL, No. 2 pp.351-401
- Galor, O. and D. N. Weil (2000), "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition, and Beyond," *American Economic Review* 86(3): 375-87.

Gan, L. and G. Gong, "Mortality Risk and Educational Attainment of Black and White Men", NBER Working Paper # 10381, March 2004.

Glewwe, P. (1999), "Why Does Mothers' Schooling Raise Child Health in Developing Countries? Evidence from Morocco," *Journal of Human Resources* 34(1): 124-159.

Gray, R. H. (1974), "The Decline of Mortality in Ceylon and the Demographic Effects of Malaria Control," *Population Studies*, Vol. 28, No. 2, pp. 205-229

Haveman, R. H., and B. Wolfe (1984), "Schooling and Economic Well-Being: The Role of Non-Market Effects," *Journal of Human Resources* 19(3): 377-407.

Hertz, Thomas (2003) "Upward Bias in the Estimated Returns to Education: Evidence from South Africa" *American Economic Review*, Vol. 93, No. 4., pp. 1354-1368.

Hurd, M.D. and K. McGarry (2002), "The Predictive Validity of Subjective Probabilities of Survival," *Economic Journal* 112(482): 966-985.

Jayaweera, S. (1969), "Recent Trends in Educational Expansion in Ceylon," *International Review of Education*, 15(3): 277-294

Jones, Margaret (2004). *Health Policy in Britain's Model Colony : Ceylon (1900-1948)*. New Delhi: Orient Longman.

Kalemli-Ozcan, S., H.E. Ryder, and D.N. Weil (2000), "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics* 62(1): 1-23.

Langford, C.M. (1981), "Fertility Change in Sri Lanka Since the War: An Analysis of the Experience of Different Districts," *Population Studies* 35(2): 285-306.

Levine, R. (2007), *Case Studies in Global Health: Millions Saved*. Sudbury, MA: Jones and Bartlett.

Lorentzen, P., J. McMillan, and R. Wacziarg (2005), "Death and Development," National Bureau of Economic Research Working Paper No. 11620.

Loudon, I. (1988), "Maternal Mortality: 1800-1950. Some Regional and International Comparisons," *Social History of Medicine*, Vol. 1, 1988, 183-228.

Loudon, I. (1991), "On Maternal and Infant Mortality 1900-1960," *Society for the Social History of Medicine* 4: 29-73.

Loudon, I. (2000a) "Maternal Mortality in the Past and Its Relevance to Developing Countries Today," *American Journal Clinical Nutrition* 72(suppl): 241S-6S.

Loudon I. (2000b), *The Tragedy of Childbed Fever*. Oxford: Oxford University Press.

Lucas, A. M. (2005), "Economic Effects of Malaria Eradication: Evidence from the Malarial Periphery," mimeo, Brown University.

Maine, D. (2000), "Role of Nutrition in the Prevention of Toxemia," *American Journal of Clinical Nutrition* 72(suppl): 298S-300S.

Maine, D. et al (1991). *Safe Motherhood Programs: options and issues*. New York: Columbia University.

Miguel, E. and M. Kremer (2004), "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72(1): 159-217.

Murphy, K. M., and R. H. Topel (2005), "The Value of Health and Longevity," National Bureau of Economic Research Working Paper No. 11405.

Nadarajah, T. (1983) "The Transition from Higher Female to Higher Male Mortality in Sri Lanka," *Population and Development Review* 9(2): 317-325.

Oster, E. (2007), "HIV and Sexual Behavior Change: Why Not Africa?" mimeo, University of Chicago.

Paxton, A., D. Maine, L. Freedman, D. Fry and S. Lobis (2005), "The Evidence for Emergency Obstetric Care," *International Journal of Gynecology and Obstetrics* 88: 181-193.

Peebles, P. (2006), *The History of Sri Lanka*, Westport, CT: Greenwood Press.

Peters, M. and A. Siow. (2002), "Competing Premarital Investments," *Journal of Political Economy* 110: 592-208.

Psacharopoulos, G. (1994), "Returns to Investments in Education: A Global Update," *World Development* 22(9): 1325–1343.

Preston, Samuel H. (1980). "Causes and Consequences of Mortality Change in Less Developed Countries in the Twentieth Century." in R. Easterlin, ed., *Population and Economic Change in Developing Countries*. National Bureau of Economic Research, University of Chicago Press, 1980, pp. 289-360.

Retherford, Robert D. (1972) "Tobacco Smoking and the Sex Mortality Differential," *Demography* 9(2): 203-216

Rodriguez, F. and J.D. Sachs (1999), "Why Do Resource-Abundant Economies Grow More Slowly?" *Journal of Economic Growth* 4: 277–303.

Rosenzweig, M. R., and T. P. Schultz (1989), "Schooling, Information, and Nonmarket Productivity: Contraceptive Use and Its Effectiveness," *International Economic Review* 30(2): 457–477.

Shastry, G. K., and D. N. Weil (2003), “How Much of Cross-Country Variation in Income Is Explained by Health?” *Journal of the European Economic Association* 1(2-3): 387-96.

Soares, R. R. (2005), “Mortality Reductions, Educational Attainment and Fertility Choice,” *American Economic Review* 95(3): 580-601.

Thomas, D., J. Strauss, and M-H. Henriques (1991), “How Does Mother’s Education Affect Child Height?” *Journal of Human Resources* 26(2): 183–211.

Thomasson, M. A and J. Treber (2004), “From Home to Hospital: The Evolution of Childbirth in the United States, 1927-1940,” NBER Working Paper No. 10873.

Trussell J, Pebley AR (1984), “The Potential Impact of Changes in Fertility on Infant, Child, and Maternal Mortality,” *Studies in Family Planning*, 15(6/1): 267–280.

United Nations, Economic and Social Commission for Asia and the Pacific (1976), Population of Sri Lanka. Country Monograph Series No. 4. Bangkok Thailand.

Van Lerberghe, W. and V. De Brouwere (2001), “Of Blind Alleys and Things That Have Worked: History’s Lessons on Reducing Maternal Mortality,” *Studies in Health Services Organisation and Policy* 17.

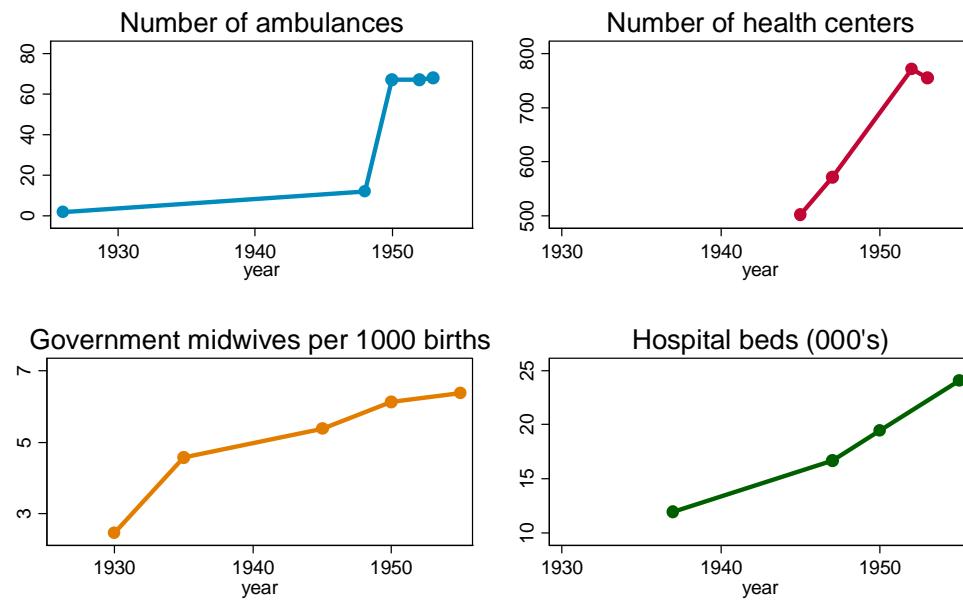
Weil, D. N. (2007), “Accounting for the Effect of Health on Growth,” mimeo, Brown University.

World Bank (2003), “Investing in Maternal Health: Learning from Malaysia and Sri Lanka,” Human Development Network, Health, Nutrition and Population Series, Washington DC.

World Health Organization (2007), *Malaria in Pregnancy. Guidelines for Measuring Key Monitoring and Evaluation Indicators*. Geneva.

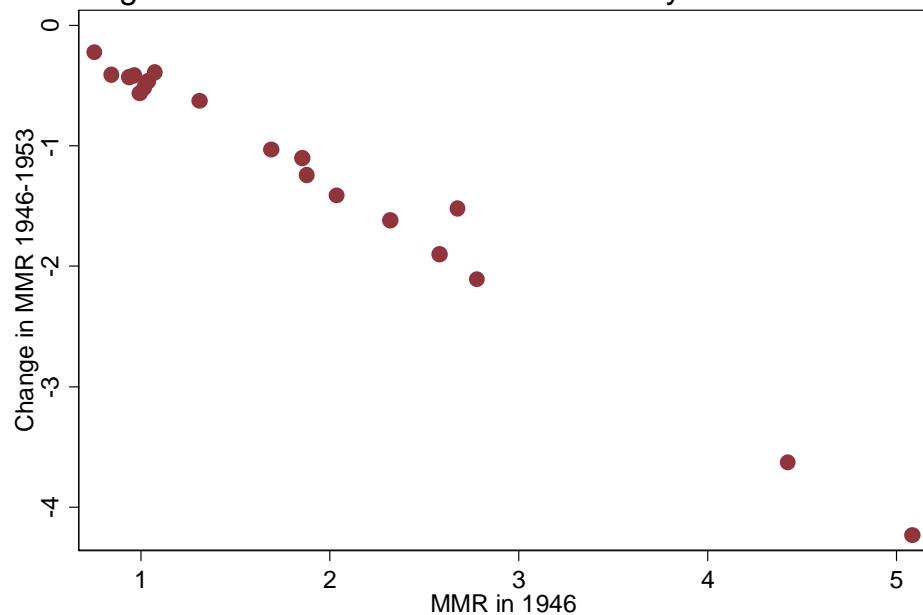
Young, A. (2005), “The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations,” *Quarterly Journal of Economics* 120(2): 423-466.

Figure 1: Expansion of health services in Sri Lanka



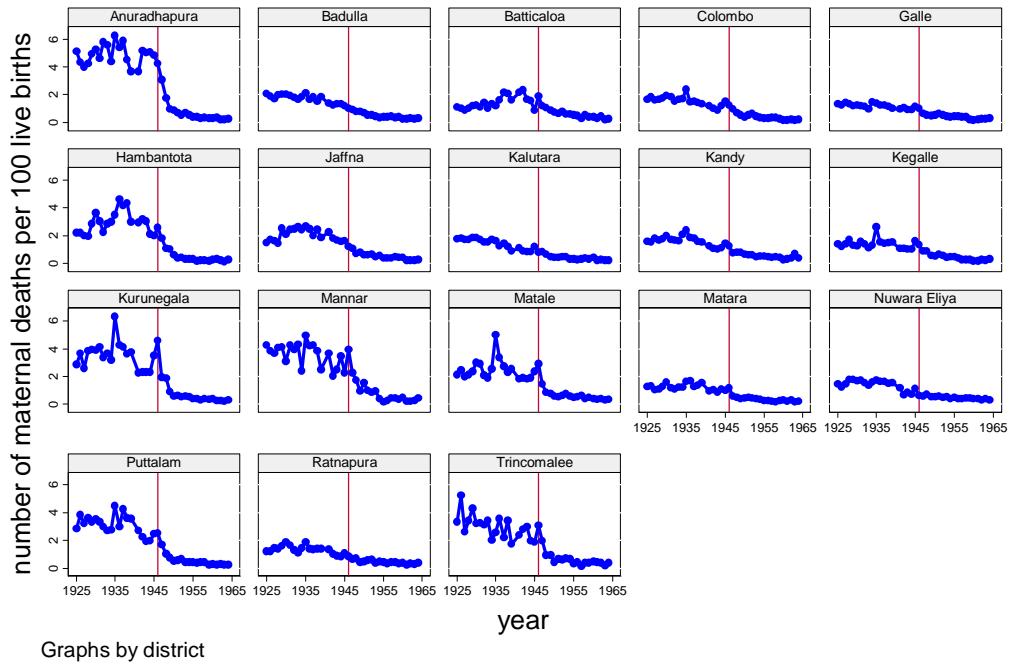
Sources: Data on ambulances, health centers and hospital beds are from the report of the Director of Health Services Administration Reports (various years). Data on government midwives are from World Bank (2003).

Figure 2: Declines in maternal mortality across districts



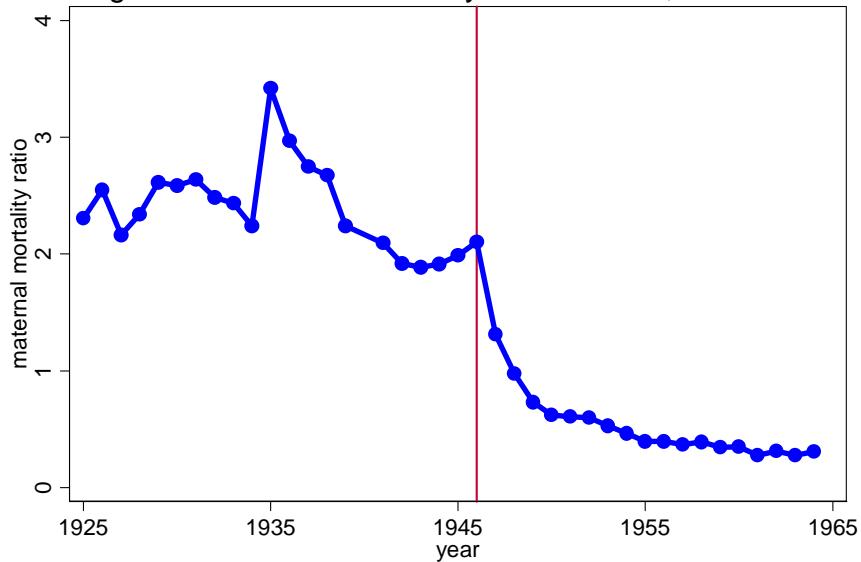
Note: Each dot represents a district. Maternal mortality is the number of deaths per 100 live births. In a univariate regression of maternal mortality changes between 1946 and 1953 on the initial 1946 level, the coefficient on initial MMR is -0.70 and is statistically significant at the 5% level.

Figure 3a: Maternal mortality by district, 1925-1964

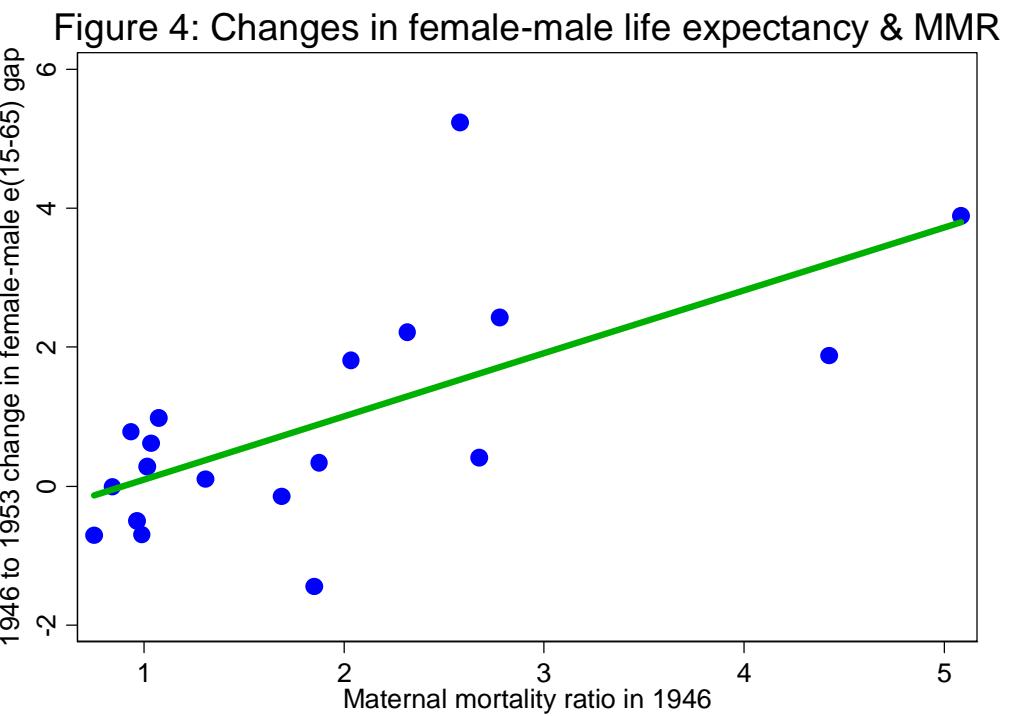


Note: Maternal mortality is the number of deaths per 100 live births. Vavuniya is excluded because of scale: due to small numbers, its female/male ratio is noisy and takes on large values. When plotting all districts on the same scale, including Vavuniya makes the patterns difficult to observe.

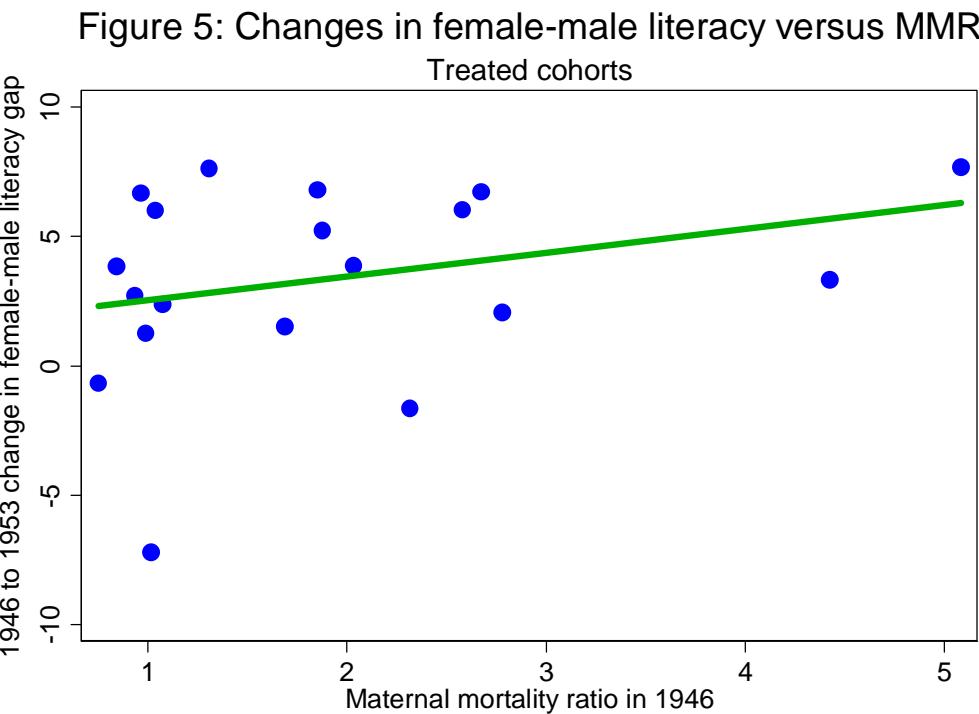
Figure 3b: Maternal mortality for Sri Lanka, 1925-1964



Note: Figure reports unweighted averages across districts. Maternal mortality is the number of deaths per 100 live births.



Note: Maternal mortality is the number of deaths per 100 live births. $e(15-65)$ is life expectancy at age 15, censored at age 65.



Note: Maternal mortality is the number of deaths per 100 live births. Female-male % literate is % literate among females minus % literate among males.

Table 1a: Summary Statistics

	Males		Females	
	1946	1953	1946	1953
Mortality and Life expectancy				
Maternal mortality ratio (current)*			1.80	0.53
Maternal mortality ratio (lagged 3 yrs)			1.91	0.65
e(15-65)**	39.1	45.7	36.8	44.2
e(45-65)	15.7	18.2	16.1	18.3
e(0-15)	10.9	12.8	10.8	12.9
e(0-65)	37.12	50.26	35.27	49.37
Infant mortality rate (current)*	17.7	8.1	16.6	6.8
Fertility***				
Birth rate			178.9	202.4
Log(Number of births)			9.11	9.39
Log(Female population 15-45)			10.83	11.02
Education				
% literate				
Treated cohorts: ages 5-19	58.17	69.35	43.88	58.43
Control cohorts: ages 25-44	78.54	82.32	38.12	46.2
Age 5-9	28.43	44.62	25.17	41.82
Age 10-14	67.55	81.06	51.81	69.26
Age 15-19	78.55	82.37	54.65	64.21
Age 20-24	81.99	85.63	50.36	60.03
Age 25-29	81.83	85.36	44.78	54.99
Age 30-34	80.06	84.37	40.14	49.38
Age 35-39	77.20	80.83	36.04	42.74
Age 40-44	75.07	78.71	31.52	37.76
Age 45-49	72.20	75.50	27.79	34.31
Age 50-54	69.91	73.93	24.53	30.86
Age 55-59	69.28	70.19	23.36	27.67
Age 60-64	64.65	67.64	19.53	23.13
% in school (ages 5-24)****	36.30	44.80	34.60	39.71

N=19 districts. Maternal mortality, life expectancy, fertility and marriage statistics are 3-year averages centered on 1946 and 1953 (e.g. the 1946 birth rate is the average for 1945, 1946 and 1947).

* Maternal mortality ratio and infant mortality rate are number of deaths per 100 live births.

** e(x-y) denotes the expected years of life between ages x and y, conditional on surviving until age x. Note that life expectancy, because it involves conditioning on survival to a certain age is not additive, so e(0-65) is not the sum of e(0-15) and e(15-65).

*** Birth rate=[births/female pop ages 15-45]*1,000. Number of births and female population are in units of 1000 before taking the logarithm.

**** % in school is percent who report their occupation as student.

Table 1b: Summary Statistics
Death rates by disease and by age

	Males		Females	
	1946	1953	1946	1953
Disease mortality rates				
Vitamin	0.99	0.58	1.51	0.79
Malaria	1.44	0.10	1.64	0.12
Diarrhea	0.91	0.39	0.98	0.45
Helminths	0.52	0.38	0.68	0.55
Anemia	0.37	0.21	0.36	0.25
Age-specific death rates				
ages 0-4	72.75	34.92	74.20	33.31
ages 5-9	6.75	3.04	7.72	3.48
ages 10-14	3.57	1.46	4.20	1.56
ages 15-19	5.10	1.68	8.01	2.67
ages 20-24	6.58	2.16	12.87	3.94
ages 25-29	8.19	2.55	13.44	4.91
ages 30-34	8.19	2.78	13.44	5.10
ages 35-39	13.13	3.67	13.44	5.48
ages 40-44	13.13	4.58	13.44	5.44
ages 45-49	21.55	6.42	19.29	6.50
ages 50-54	21.55	9.29	19.29	8.46
ages 55-59	39.78	13.92	35.20	11.91
ages 60-64	39.78	20.68	35.20	19.09
ages 65+	99.46	68.69	107.22	76.96

Disease-specific and cause-specific rates are per 1000. The 1946 data are available only by 10 year age groups for ages 25-34, 45-54 and 55-64. All statistics in this table are averages over three years, centered on 1946 and 1953 (e.g. the 1946 death rate is the average for 1945, 1946 and 1947).

Table 1c: 1971 Census Summary Statistics

Variable	Males		Females	
	Control cohorts	Treated cohorts	Control cohorts	Treated cohorts
Literate	79.5	87.7	41.3	69.8
Years of education	5.1	6.3	4.2	5.7
Survival rate (cell size in 1971/cellsizes in 1953)	0.694	0.868	0.712	0.925
Years of education missing	0.216	0.140	0.531	0.275
District of current residence missing	0.022	0.025	0.022	0.023
Ethnicity missing	0.004	0.005	0.005	0.005
Religion missing	0.003	0.003	0.004	0.003

Literacy rates are per 100. All other proportions are reported on a scale of 0 to 1. Each cell, defined by district, gender, and treatment status, weighted equally Treated cohorts are those ages 5 to 19 in 1953 and control cohorts are those ages 25 to 44 in 1953. The underlying individual census data contain 399,100 individuals.

Table 2: Effect of maternal mortality on life expectancy and infant mortality
 Difference-in-difference-in-difference estimates

Dependent variable:	(1) Basic	(2) Add malaria death rates	(3) Add nutritional diseases death rates	(4) Add nutritional diseases and malaria death rates
e(15-65)				
MMR*female	-1.204*** [0.198]	-1.302*** [0.302]	-1.214*** [0.183]	-1.373*** [0.330]
R-squared	.97	.97	0.98	0.98
e(45-65)				
MMR*female	0.054 [0.089]	-0.033 [0.120]	0.078 [0.119]	-0.043 [0.180]
R-squared	0.94	0.95	0.97	0.97
e(0-15)				
MMR*female	-0.088* [0.050]	-0.081 [0.065]	-0.064* [0.033]	-0.018 [0.055]
R-squared	0.99	0.99	0.99	0.99
IMR				
MMR*female	0.133 [0.164]	0.081 [0.192]	0.265* [0.145]	0.213 [0.265]
R-squared	0.99	0.99	0.99	0.99

All regressions include district*year, district*gender and gender*year fixed effects. Additional controls are measured in changes. The notation e(15-65) is the expected years of life between ages 15 and 65, conditional on surviving until age 15, and so forth. MMR is the maternal mortality ratio, and IMR is the infant mortality rate. Both are measured as deaths per 100 live births and are measured contemporaneously. Nutritional diseases are helminths, anemia, diarrhea and vitamin deficiencies. Standard errors (reported in brackets) are clustered at the district level. Each cell reports the coefficient from a separate regression. N=76 (19 districts, 2 genders, 2 years).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Effect of maternal mortality on literacy and percent in school

	DDD (district, year, gender)				Older cohorts as control		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Basic	Add nutritional diseases and malaria death rates	1946 level as IV	Drop district FE & control for male e(0-65)		DDDD (district, year, gender, cohort)	DDD Females (district, year, cohort)	DDD males (district, year, cohort)
Coefficient reported	lagged MMR *female	lagged MMR *female	lagged MMR *female	lagged MMR *female	lagged MMR *female *treated	lagged MMR * treated	lagged MMR * treated
	-0.879* [0.453]	-1.652** [0.656]	-1.008** [0.470]	-1.07 [1.763]	-0.728 [0.745]	-1.354 [0.797]	-0.626 [0.968]
Obs	228	228	228	228	532	266	266
Panel A: Literacy of treated cohorts ages 5-19							
	-0.151 [0.469]	0.273 [0.450]	-0.149 [0.476]				
Obs	304	304	304				
Panel B: Placebo test, literacy of controls cohorts ages 25-44							
	-0.904* [0.458]	-0.686 [0.995]	-.979** [0.460]				
Obs	76	76	76				

MMR is lagged by 3 years. All regressions include district*year, district*gender and gender*year fixed effects. The regressions also include age*district, age*year and age*gender fixed effects. Nutritional diseases are helminths, anemia, diarrhea and vitamin deficiencies. These diseases are district- and gender-specific. Standard errors clustered within a district-gender are reported in brackets. In Panels A and B each observation is a district-gender-year and 5-year age group (for example in Panel A, column 1, there are 19 districts, 2 gender, 2 years and 2 age groups). In Panel C each observation is a district-gender-year.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Age-specific effects of maternal mortality on literacy
 Difference-in-difference-in-difference estimates

	(1)	(2)
	Basic	Add nutritional diseases and malaria
Dependent variable: % literate by age		
Ages 5-9*		
(lagged MMR)*female	-0.478 [0.574]	-1.373 [1.267]
Ages 10-14*		
(lagged MMR)*female	-0.796 [0.848]	-1.869 [1.359]
Ages 15-19*		
(lagged MMR)*female	-1.363 [1.022]	-1.714 [1.179]
Ages 20-24*		
(lagged MMR)*female	-1.598 [0.995]	-2.592*** [0.937]
Ages 25-29*		
(lagged MMR)*female	-1.387* [0.793]	0.022 [0.754]
Ages 30-34*		
(lagged MMR)*female	-0.962 [0.856]	-0.382 [1.232]
Ages 35-39*		
(lagged MMR)*female	0.008 [0.918]	0.044 [0.797]
Ages 40-44*		
(lagged MMR)*female	1.738 [1.098]	1.407 [1.010]
Ages 45-49*		
(lagged MMR)*female	1.368 [1.166]	0.658 [1.134]
Ages 50-54*		
(lagged MMR)*female	0.419 [1.274]	-0.126 [1.335]

MMR is lagged by 3 years. All regressions include all triple interactions between district, year, age group, and gender. Nutritional diseases are helminths, anemia, diarrhea and vitamin deficiencies. The disease death rates in column 2 are district-, year- and gender-specific, and they are interacted with a dummy for each age group. Standard errors, clustered by district-gender, are reported in brackets. Each observation is a district, gender, year and 5-year age group (19 districts, 2 genders, 2 years and 10 age groups). Coefficients are estimated in a single regression.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Effect of maternal mortality on literacy: Specification checks
 Difference-in-difference-in-difference estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Basic	Corrected for migration	Drop 2 outliers	Population weights	Ages 5-14 as treated group	Ages 5-14 as treated group, controlling for other diseases	MMR lagged 1 year	MMR lagged 2 year	MMR lagged 3 year	MMR lagged 4 year
lagged MMR	-0.879*	-1.067**	-0.922	-1.378**	-0.637	-1.621**	-0.210	-0.683*	-0.731	-1.160**
*female	[0.453]	[0.461]	[1.087]	[0.754]	[0.447]	[0.743]	[0.551]	[0.398]	[0.455]	[0.473]
Observations	228	228	204	228	152	152	228	228	228	228
R-squared	0.36	0.36	0.33	0.13	0.54	0.55	0.36	0.36	0.36	0.36

Each observation is a district (19), gender (2), year (2) and 5-year age group (3). Column 1 reproduces Panel A, Column 1 from Table 3. Column 2 adjusts for inter-district migration between 1946 and 1953 and the gender-specific differences in literacy for migrants and non-migrants (see the data appendix for more detail). Column 3 drops 2 outliers in terms of MMR declines, Anuradhapura and Vavuniya. Column 4 weights each observation by the population for that age-gender-district-year cell. Columns 5 and 6 include observations for only ages 5-14, rather than ages 5-19. In columns 1-6, MMR is a three-year running average lagged by 3 years. In columns 7-10, MMR is the single-year value shown for different lags. All regressions include district*year, district*gender and gender*year fixed effects. The regressions also include age*district, age*year and age*gender fixed effects. Standard errors clustered within a district-gender are reported in brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Effect of MMR on literacy and years of schooling using the 1971 Census

Dependent variable:	Survival rate (cell size in 1971/cell size in 1953)		Education missing in 1971		Literate		Years of education		
	OLS	OLS	OLS	OLS	OLS (DDD)	Women only OLS (DD)	OLS	Heckman selection correction	
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Lagged MMR)*female	0.0106 [0.0137]	0.0127 [0.0131]	0.0246*** [0.0076]	0.0195** [0.0078]	-2.711*** [0.8437]		0.016 [0.028]	-0.133** [0.057]	-0.132** [0.057]
Literacy in 1953		0.0007 [0.0015]		-0.0019*** [0.0007]					
MMR						-2.681** [1.257]			
Excluded variables	n/a	n/a	n/a	n/a	n/a	n/a	n/a	District of 1971 residence dummies	District of residence missing, ethnicity missing & religion missing

N=266. Each observation is a district of birth (19), gender (2) and 5 year age group (7 age groups, three of which are treated and 4 of which are controls). Treated ages (ages 5 to 19 in 1953) are assigned the 1953 MMR value (post-decline) and the control ages (25 to 44 in 1953) are assigned the 1946 MMR value (pre-decline). All regressions control for female*age-group, age-group*district, and district*female dummies. Standard errors (in brackets) are clustered at the district-gender level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Effect of maternal mortality on fertility
Difference-in-difference estimates

	(1)	(2)	(3)
Dependent variable:	Basic: D-D with male e(0-65) as a control	Male e(0-65), malaria and nutritional death rates as controls	Male e(0-65), malaria and nutritional death rates and lagged IMR as controls
<u>Birth rate</u>			
Lagged MMR	-5.15* [2.55]	-4.77* [2.65]	-3.86 [3.01]
R-squared	0.92	0.95	0.95
<u>log(# of births)</u>			
Lagged MMR	-0.08*** [0.02]	-0.10*** [0.02]	-0.09*** [0.03]
R-squared	0.92	0.94	0.94
<u>log (Female population age 15 -45)</u>			
lagged MMR	-0.07*** [0.02]	-0.08** [0.03]	-0.09** [0.04]
R-squared	0.74	0.81	0.82

MMR is lagged by 3 years. The regressions include district and year fixed effects, plus the controls specified in the column heading. Nutritional diseases are helminths, anemia, diarrhea and vitamin deficiencies. Standard errors (reported in brackets) are clustered at the district level. Each cell reports the coefficient from a separate regression. Number of births and female population ages 15 to 45 are in units of 1000. For consistency with the MMR measure, IMR in column 3 is lagged by 3 years. N=38 (19 districts*2 years).

* significant at 10%; ** significant at 5%; *** significant at 1%

Data Appendix

Data sources

Vital statistics: Maternal mortality ratio (calculated by authors as maternal deaths divided by number of births), infant mortality, age-specific mortality rates, disease-specific mortality rates, birth rates (calculated by authors as number of births divided by Census population counts)

Census of Population (1946, 1953): Literacy, % in school, population, % of population in urban areas

Census of Population (1971): Literacy, education

Administration Report of the Director of Medical and Sanitary Services (1945, 1947 1952): Spleen rates, number of ambulances, number of health centers, number of hospital beds

De Silva (1943): Maternal mortality ratio for 1925 to 1939

World Bank (2003): Government midwives

United Nations (1976): Inter-district migration rates

District definitions

Two districts that merged over the study period (Puttalam and Chilaw) are treated as a merged district from the outset (named Puttalam). Also, Colombo and Negombo are treated as one district in the censuses (named Colombo), and therefore in our study, despite being separate administrative districts throughout the period. This yields 19 districts.

Conversion from annual data to 1946 and 1953 time periods

Vital statistics data (births, deaths) are available annually. The values we use for 1946 are the average of 1945, 1946, and 1947, and the values for 1953 are the average of 1952, 1953 and 1954. We construct the other variables in a consistent manner. For example, MMR lagged 3 years from 1946 is constructed as the average of the 1942, 1943, and 1944 rates.

Interpolation between census years

To calculate death rates and birth rates, we use annual vital statistics data on deaths and births in the numerator. For the denominator, we linearly interpolate population between census years.

Life expectancy calculation

To calculate life expectancy from mortality tables, we consider an individual who has survived until age a and calculate the probability of surviving each subsequent year. The death rate data are for a 5-year age band. Our calculation uses 1 year as the age increment that we iterate over, and we assume the death rate is constant for each age in the 5-year-band. The formula is:

$$e(a-b) = \left(\sum_{t=a, a+1, \dots}^b (t + 1/2) * {}_t p_a * q_t \right) + b * {}_b p_a * (1 - q_b) - a$$

where q_t is the probability of dying at age t (the mortality rate for age t), and p_a is the probability of surviving from age a to age t . The summation accumulates the expected years of life at successive ages for individuals who die at that age. The factor of 1/2 is because we treat the death as taking place at the midpoint of the discrete one-year period. The term that follows the summation accounts for individuals who survive past age b , that is, those for whom the censoring at age b is binding. The subtraction of a is because the measure is of *subsequent* years of life, conditional on surviving until a . The original data are death rates, so p_a is itself constructed as follows:

$${}_t p_a = \prod_{\tau=a}^{t-1} (1 - q_\tau)$$

Notes on other variables

IMR: The infant mortality rate is calculated as the number of deaths in a calendar year among those 0 to 1 divided by the number of births in the calendar year. Some individuals born in one calendar year will die as infants in the next calendar year, but this approximation should net out on average.

% Literate: The 1946 and 1953 censuses defined people as literate if they reported that they could both read and write in at least one language. The standard was that “the person is able to write a short letter and read the reply to it.” Literacy was asked of individuals age 5 and over.

% ages 5-24 in school: The census reports the student population among those ages 5-24. Students are those whose response to the occupation question is “student.” The most likely category that would be used for children not in school is “other dependent.” No further instructions are given for the occupation question.

Migration correction

The United Nations' report (1976) provides published statistics of the number of net migrants by district and gender from 1946 to 1953 using the survival ratio method (Table 27, chapter III). They also report the distribution of net migrants by age (although not by district) in Table 29, chapter III. We compute age and gender-specific net migration rates for each district, by assuming that the national age distribution applies to each district. Using the 1971 census, we compute the literacy gap by gender for migrants and non-migrants among those individuals ages 5 to 19 in 1971. The literacy of female migrants was found to be 0.002408 percent higher for women and 0.0269181 percent higher for men. We use these numbers to correct the observed literacy rate in 1953. For women the corrected literacy rate for a given age and district is computed as:

$$\text{Migration-corrected literacy rate} = [(\text{literacy rate} * \text{pop}) - (\text{net migration rate} * \text{literacy rate} * 1.002408)] / (\text{pop} - \text{net migration rate})$$

For men the corrected literacy rate for a given age and district

$$\text{Migration-corrected literacy rate} = [(\text{literacy rate} * \text{pop}) - (\text{net migration rate} * \text{literacy rate} * 1.0269181)] / (\text{pop} - \text{net migration rate})$$

Census 1971: Converting educational attainment to years of schooling

Education attainment as coded in the 1971 census	Assigned years of schooling
0-9 (corresponds to passing grades or standard 0-9)	0-9
10-24 (code for whether passed “O” level exams, number of subjects if available or whether passed the Matriculation or other equivalent examination)	10
25-34 (code for whether passed “A” level exams and number of subjects if available)	12
35: “professional, technical or other qualification below degree level but above G.C.E “A” Level”	14
36: “Academic degree (University degree) e.g. B.A., B.S., etc,”	16
37: “Professional technical or other qualifications equivalent to a degree”	
38: “Academic post-graduate degree, e.g. M.A., M.Sc., Ph.D., etc”	17
39: “Professional, technical or other qualification at the post graduate level”	

Appendix Table 1: Determinants of maternal mortality in 1946

Dependent variable is maternal mortality ratio in 1946 (deaths per 1,000 live births).								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spleen rate 1947	0.066*** [0.009]		0.078*** [0.011]				0.076*** [0.012]	0.073*** [0.011]
Male malaria death rate		0.635* [0.334]	-0.368 [0.228]				-0.276 [0.261]	-0.091 [0.290]
% population in urban areas					-0.008 [0.025]		0.008 [0.015]	0.005 [0.014]
% males literate ages 20 to 24				-0.021 [0.145]				0.137 [0.092]
% males literate ages 25 to 29				-0.231 [0.180]				-0.154* [0.085]
% males literate ages 30 to 34				0.242 [0.146]				0.042 [0.080]
Mean % male literate (ages 20-34)						0.032 [0.042]	0.015 [0.023]	
Observations	19	19	19	19	19	19	19	19
R-squared	0.75	0.18	0.79	0.17	0.01	0.03	0.80	0.85

There are 19 observations, one per district. District-level spleen rates are available only in 1947.

Appendix Table 2: Effect of maternal mortality on age-specific mortality rates

Dependent variable: Age-specific death rate	Basic	Add nutritional diseases and malaria death rates
Age 0-4		
MMR*female	1.672 [1.116]	0.275 [1.447]
Age 5-9		
MMR*female	0.322* [0.174]	0.420*** [0.124]
Age 10-14		
MMR*female	0.578** [0.223]	0.176 [0.343]
Age 15-19		
MMR*female	2.616*** [0.379]	2.491*** [0.668]
Age 20-24		
MMR*female	5.391*** [0.987]	4.513*** [0.560]
Age 25-29		
MMR*female	2.054*** [0.597]	2.537*** [0.419]
Age 30-34		
MMR*female	2.322*** [0.554]	2.966*** [0.483]
Age 35-39		
MMR*female	-0.829 [1.016]	0.227 [0.809]
Age 40-44		
MMR*female	-0.915 [0.846]	0.35 [0.644]
Age 45-49		
MMR*female	0.259 [0.688]	0.689 [1.187]
Age 50-54		
MMR*female	0.625 [1.023]	-0.124 [1.233]
Age 55-59		
MMR*female	-4.303 [2.558]	1.135 [1.659]
Age 60-64		
MMR*female	-5.304** [2.274]	0.248 [2.506]
Age 65+		
MMR*female	-2.861 [2.797]	-3.347 [2.344]

MMR is measured contemporaneously. All regressions include district*year, district*gender and gender*year fixed effects. Nutritional diseases are helminths, anemia, and diarrhea and vitamin deficiencies. Standard errors (in brackets) are clustered at the district-gender level. Each cell reports the coefficient from a separate regression. N=76 (19 districts, 2 genders, 2 years).

* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 3: Test for outliers: results dropping one district at a time
 Dependent variable: % literate by age

		Basic	Drop district 1	Drop district 2	Drop district 3	Drop district 4
Treated ages	(lagged MMR)					
	female	-0.879 [0.453]	-0.459 [0.574]	-1.026** [0.449]	-0.905* [0.456]	-0.516 [0.397]
Placebo ages	(lagged MMR)					
	*female	-0.151 [0.469]	0.021 [0.714]	0.036 [0.474]	-0.126 [0.478]	-0.329 [0.468]
		Drop district 5	Drop district 6	Drop district 7	Drop district 8	Drop District 9
Treated ages	(lagged MMR)					
	*female	-1.048** [0.453]	-0.967** [0.451]	-0.861* [0.459]	-0.930* [0.479]	-0.868* [0.485]
Placebo ages	(lagged MMR)					
	*female	-0.437 [0.437]	-0.337 [0.471]	-0.187 [0.470]	-0.154 [0.501]	-0.057 [0.499]
		Drop district 10	Drop district 11	Drop district 12	Drop district 13	Drop district 14
Treated ages	(lagged MMR)					
	*female	-1.010** [0.461]	-0.973** [0.448]	-0.840* [0.454]	-0.880* [0.454]	-0.827* [0.476]
Placebo ages	(lagged MMR)					
	*female	-0.182 [0.496]	-0.025 [0.455]	-0.122 [0.464]	-0.15 [0.473]	-0.201 [0.489]
		Drop district 15	Drop district 16	Drop district 17	Drop district 18	Drop district 19
Treated ages	(lagged MMR)					
	female	-0.715 [0.484]	-0.876 [0.453]	-0.881* [0.482]	-0.813* [0.456]	-1.216** [0.458]
Placebo ages	(lagged MMR)					
	*female	0.252 [0.406]	-0.155 [0.469]	-0.135 [0.500]	-0.131 [0.468]	-0.448 [0.550]

Each coefficient is from a separate regression. MMR is lagged by 3 years. All regressions include district*year, district*gender and gender*year fixed effects. The regressions also include age*district, age*year and age*gender fixed effects. Nutritional diseases are helminths, anemia, diarrhea and vitamin deficiencies. Standard errors clustered within a district-gender are reported in brackets. Each observation is a district-year-gender-5 age group (19 districts in the first regression, and 18 districts thereafter; 2 years, 2 genders; 3 age groups for the treated, and 4 age groups for the placebo).

* significant at 10%, ** significant at 5%; *** significant at 1%

Appendix Table 4: Gender differences in elasticity of literacy with respect to health – other diseases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Malaria	Anemia	Diarrhea	Vitamin-related diseases	Helminths	Influenza	Dysentery
Mortality rate	-1.514 [1.064]	-21.078*** [6.432]	-3.239 [2.828]	4.231*** [1.365]	-3.48 [6.023]	-7.186*** [1.913]	-5.661 [10.303]
Mortality rate*female	0.002 [1.505]	5.606 [9.096]	1.968 [4.000]	-0.631 [1.931]	-2.998 [8.517]	1.9588 [2.706]	5.907 [14.570]
Average % change in disease rate, 1946-53	-93%	-36%	-55%	-42%	-22%	-61%	-61%
Mortality rate in 1946 (per 1000)	1.64	0.36	0.98	1.51	0.68	0.28	0.22

The regressions also include MMR and MMR*female. There are 228 observations per regression. The diseases are those used as control variables in the main tables (malaria, anemia, diarrhea, vitamin-related diseases, and helminths), plus the other 2 diseases where a decline of over 50% was seen on average between 1946 and 1953, dysentery and influenza.

* significant at 10%; ** significant at 5%; *** significant at 1%