

# Improving Empirical Estimation of Demographic Drivers:

## Fertility, Child Mortality & Malaria Ecology

Gordon C. McCord<sup>a</sup>  
Columbia University

Dalton Conley<sup>b</sup>  
New York University

Jeffrey D. Sachs<sup>c</sup>  
Columbia University

a: Corresponding author  
[gm2101@columbia.edu](mailto:gm2101@columbia.edu)  
314 Low Library MC 4327  
535 West 116<sup>th</sup> Street  
New York, NY 10027  
United States  
+1-646-236-1720  
+1-212-854-8702 (fax)

b: [conley@nyu.edu](mailto:conley@nyu.edu)  
Tasco House  
6 Washington Square North, 21  
New York, NY 10003  
United States

c: [sachs@columbia.edu](mailto:sachs@columbia.edu)  
314 Low Library MC 4327  
535 West 116<sup>th</sup> Street  
New York, NY 10027  
United States

**Improving Empirical Estimation of Demographic Drivers:  
Fertility, Child Mortality & Malaria Ecology**

**Abstract**

Much of Africa has yet to go through a demographic transition; this Malthusian crisis of high mortality, high fertility, rapid population growth and chronic extreme poverty has been attributed to factors including the status of women, pro-natalist policies, and poverty itself. Large uncertainty exists among demographers as to the relative importance of these factors, mostly since econometric estimation is complicated by the endogeneity of fertility to other variables of interest. We attempt to improve estimation of the effect of the child mortality variable on fertility by deploying exogenous variation in the ecology of malaria transmission. Results show that child mortality is a powerfully robust driver of fertility behavior. Meeting the Millennium Development Goal of reducing 1990 child mortality rates by 66% in sub-Saharan Africa would translate into a reduction of total fertility rates from around 6.3 in 1990 to 3.3, more than halfway towards achieving replacement fertility levels of 2.1.

Keywords: fertility, mortality, malaria, demography, human ecology

JEL Codes: J13, I10, O15

## *Introduction*

The broad categories of determinants of fertility are generally thought to be reasonably well identified by demographers, sociologists and economists, though the detailed quantitative determinants of fertility levels and changes are much less understood. The relationships between fertility, on the one hand, and economic development, the status of women, access to family planning, pro-natalist or pro-planning policies of government, and mortality (both adult and child), on the other hand, have been elegantly theorized and extensively studied. Likewise, many researchers have gone before us in empirically modeling these relationships in both cross-country and within-country analyses, some of which explicitly tackle endogeneity problems in estimation using instrumental variables.<sup>1</sup>

In quantitative terms, however, we still lack a good understanding of why some countries have experienced significant reductions of fertility rates, while those in Africa remain with very high fertility rates—on average difference of almost three births between Sub-Saharan Africa and the rest of the developing world. (See Figure 1, below, for the distribution of total fertility rates in the developing world.) To what extent are Africa's high fertility rates the result of illiteracy, poverty, high child mortality, or lack of access of the poor to contraception? If the region were to experience a mortality decline, what would the effect on fertility be? This paper seeks to contribute towards a quantitative assessment of these questions. We acknowledge from the start, however, that many of the potential explanatory variables are only imperfect proxies for the household and community-level drivers of fertility and fertility change. Bongaarts et al.

---

<sup>1</sup> The most relevant to this paper are attempts to instrument child mortality: Schultz (1997), Benefo and Schultz (1996), Dreze and Murthi (2001), and Kalemli-Ozcan (2006).

[1] posits that fertility is regulated directly by proximate determinants (e.g., contraception, age at marriage, abortion), while socioeconomic variables (income, education, mortality) affect fertility only indirectly by modifying the proximate determinants. Thus, we can aim only to get a rough quantitative assessment of the role of key categories of determinants, rather than precise point estimates of how specific policy changes would affect fertility rates.

[FIGURE 1: TOTAL FERTILITY RATES IN 2000]

Most theories of the demographic transition have put great stress, and we believe rightly so, on the causal link from high child mortality to high desired fertility. Simply put, when parents do not know whether their children will survive, they respond by having large families. In a high mortality context, cultural patterns – age of marriage, social norms in childrearing, community support structures – also favor high natality. The original model of the demographic transition, indeed, was driven almost solely by child mortality rates. Exogenous changes to child mortality (e.g., the advent of public health, safe drinking water, immunizations, improved nutrition) were seen as the primary precursor to reduced fertility rates—albeit with a lag of one or more generations. This putative lag reflected two things according to the standard analysis: first, the lag in perception of households that mortality rates had indeed come down persistently and reliably; and second, the lag in cultural norms surrounding marriage age, birth spacing, family size, and so forth, all needed to promote the transition from high to low fertility. Many studies found, indeed, that the fertility transition is strongly conditioned by a preceding child mortality transition. That said, recent experience suggests that the lag may be waning and that demographic transitions are happening with increased rapidity

once “triggered”: whereas Western Europe’s transitions took over a century (1800-1930 for Britain), more recent large declines in fertility have happened in as little as twenty or fewer years, such as in Bangladesh, Mauritius, and Iran (see Marandi et al [2] for a discussion of Iran). Most research attributes these declines to changes in access to family planning (Cleland et al. [3] in the case of the Matlab experiment in Bangladesh or Aghajanian [4] in the case of Iran), although some scholars emphasize female education and child mortality (Raftery et al [5] in the case of Iran). Others suggest that with the advent of mass media, cultural changes—such as those related to fertility behavior—spread more rapidly than they once did. These temporal (and other) challenges of identification leave open questions as to the relative magnitude of the role of child mortality reduction in the context of 20<sup>th</sup> century fertility declines.

One issue that continues to plague this basic line of research—and therefore which may be relevant to the African question—is the question of causal directionality between child mortality and fertility choice. Several scholars have shown that reduced family size affects human capital investment (Conley and Glauber [6]; Joshi and Schultz [7]; Mogstad and Wiswall [8]) on the micro level as well as economic growth at the macro level (Hazan and Berdugo [9]; Moav [10]). Likewise, the argument can be made that at least some of the powerful correlation of high child mortality and high fertility represents increased child mortality due to higher fertility due to increased strain on household caloric resources and decreased parental care and supervision with the addition of more children. Moreover, government “family planning efforts” could be reflecting (rather than affecting) demand for reduced family size. With these concerns in mind, in this paper we pursue a strategy of deploying exogenous variation in ecological conditions

to derive estimates of the causal impact of child mortality on fertility. We then discuss the implications of these findings for the lagging demographic transition in Africa.

### *Mortality and Fertility*

Chowdhury [11] identifies three possible relationships between fertility and mortality: a lagged causal relationship from mortality to fertility (the theory of demographic transition and choice theory) whether through child “hoarding” (as a precautionary insurance mechanism to guarantee surviving heirs) (Heer and Smith [12]) or direct replacement (for a discussion see, e.g., Cleland [13]); a causal relationship from fertility to mortality (the Ricardian theory); and an interdependent relationship between mortality and fertility (the modern economic theory of population). Perhaps the most influential recent model of fertility choice among economists is the economic theory of fertility offered by Becker and Barro [14]. Assuming stable wage rates and interest rates, falling child mortality lowers the costs average cost of raising surviving children since a greater proportion of the total investment in childrearing costs realizes a benefit (assuming little or no benefits from non-surviving children). Therefore, the authors argue that fertility rates will initially rise as child mortality declines (and cite evidence to this effect). However, if there is no accompanying change in the parents’ interest or wage rates, they argue that there is no cumulative effect. Thus, in the Becker-Barro model, any reduction in fertility resulting from a decline in child mortality would have to work indirectly through wage or interest rates.

Many have attempted to elaborate on the Barro and Becker model theoretically as well as empirically. For example, Doepke [15] seeks to understand whether “stochastic

outcomes and fertility choice are quantitatively important.” (p. 337) He differentiates between total fertility rate and net fertility rate, that latter being the average number of children per woman surviving to age five. Doepke concludes that child mortality is causally related to declining fertility rates, but that other factors are responsible for declines in net fertility. With *replacement* as the mechanism, each family has a target number of children. The death of one child induces the family to replace that child and, as a result, mortality directly affects total fertility rate. However, for mortality to affect the net fertility rate, the *hoarding* motive would have to take place. Parents would preemptively increase their fertility to protect against potential loss. If this mechanism is present, then a decline in mortality would result in a decline in net fertility. Doepke tests three models. He reports that “all three models are consistent with declining total fertility rates (i.e. number of births) in response to falling mortality. However, we are left without a clear-cut prediction for the relationship of child mortality to net fertility (i.e. the number of survivors).” (p. 344)

In his attempts to test for causality, Chowdhury [11] finds no consistent results across his thirty-five country sample. Fourteen of his cases support the hypothesis that infant mortality causally impacts fertility, while only two cases support the opposite hypothesis. The remaining cases indicate feedback between the two variables, or the absence of a relationship between fertility and mortality. His results provide stronger support for the hypothesis that mortality effects fertility, but they are notably (and admittedly) inconclusive. Meanwhile, Zakir and Wunnava [16] found that fertility rates impact mortality rates, and not vice-versa. Their model employs simple GLS regression

on cross-sectional data and fails to acknowledge the endogeneity of fertility and mortality, leading to a model that is notably mis-specified.

Some strategies for isolating the child mortality-fertility causal pathway have relied on using adult (male) mortality as an instrument. For example, Galloway, Lee and Hammel [17] use adult male mortality to instrument infant mortality in a historical analysis of Prussian data from 1875 to 1910. They claim that this measure of mortality may reflect other societal influences on child mortality (i.e. standard of living, nutrition, etc.), but should not reflect particular influences on breastfeeding associated with fertility. However, the common factors affecting both adult male mortality and child mortality are most likely correlated with fertility; furthermore, Lorentzen, McMillan, and Wacziarg [18] have argued that adult mortality should have an independent effect on fertility (and child human capital investment) by changing the discount rate of mothers and fathers.

Seeking to better address the endogeneity of child mortality and fertility, Schultz [19] instruments mortality using calorie availability. However, the validity of his instrument is open to question, as one can imagine a direct, negative causal impact of calorie availability on fertility. Moreover, the instrumented child mortality variable is significant only in cross section and not after including country dummies. In another attempt to deal with endogeneity concerns, Benefo and Schultz [20] instrument child mortality using variables for community health services and environment. (These include: living farther from a market; living close to a clinic; amount of rainfall; malaria/measles.) However, the authors discover that child mortality is only statistically significant if treated as an exogenous variable, but when instrumented by community health services and environment, mortality is not a statistically significant determinant of

fertility. Likewise, Dreze and Murthi [21] instrument mortality using the variable access to safe drinking water, claiming that the later variable should be unrelated to fertility except through its effect on mortality—again a questionable assumption since safe drinking water is related to economic development, which should have an independent effect on fertility. Finally, Kalemli-Ozcan [22] examines the impact of AIDS mortality on halting the demographic transition in Sub-Saharan Africa by using circumcision prevalence as an instrument for AIDS risk. She finds that the AIDS mortality crisis has indeed led to high fertility rates and less human capital investment in offspring. It is worth noting that all of these studies find a relatively small elasticity of child mortality to fertility. If we convert their coefficients to percentages using sample means, Benefo and Schultz [20] find an elasticity of 0.02-0.03, and admit that, “...we estimate that four to fifteen fewer child deaths are associated with a reduction of only one birth. We have no good explanation for the small size of this estimate of the fertility response to child mortality.” (p.152) Similarly, Kalemli-Ozcan [22] finds an elasticity of 0.02 (after converting to percentages), while Dreze and Murthi [21] find an elasticity of 0.11 in the specification using IV, although the instrument’s validity might be open to question. Finally, Schultz [19] finds an elasticity of .64 (converted to percentages) when instrumenting child mortality with caloric intake, but only achieves significance in cross-section and is thus open to more omitted variable bias.

### *Malaria Ecology*

To preview our empirical approach, we argue that the strength of malaria transmission as a function of ecological factors is exogenous to fertility, and use it as an

instrument for child mortality. Malaria is currently the fourth leading cause of death (after neonatal disorders, diarrhea, and pneumonia) for children under five in low income countries (see Black et al. [23]) and is responsible for at least one in every five child deaths in sub-Saharan Africa.<sup>2</sup> Estimates of malaria mortality in Africa range from one million to three million deaths per year. Malaria mortality, in turn, is highly sensitive to ecological conditions, as explained below. Employing malaria ecology as an instrument does not imply, of course, that malaria is the only disease affecting child mortality. However, its strong link to ecology allows us to exploit it in order to remove the endogenous part of child mortality rates and isolate its causal affect on fertility.

Specifically, we deploy an ecological index of malaria transmission (used elsewhere as well; see, e.g., Sachs [24]; Carstensen and Gundlach [25] for the use of a time-static version, and McCord [26] for the time-varying version) that combines ecological factors—rainfall and temperature—with biological ones such as the human biting rate of the mosquito species that serves as the vector for the transmission of malaria to develop an index of malaria strength (see the Appendix for more details on the construction of the index). The 1960-2005 average distribution of this malaria ecology index is mapped in Figure 2, below. While the underlying factors determining malaria transmission may be endogenous to human population movements over the course of thousands of years (through co-evolution with mosquito species), we assert that from the point of view of the current demographic transition in recent decades the biophysical ecology of malaria transmission is exogenous. Moreover, we expand over Conley et al.

---

<sup>2</sup> Since infection with malaria leaves an individual more vulnerable to morbidity and mortality from other infections, malaria is an indirect as well as a direct killer. It is very likely implicated in more than one fifth of all deaths. In some malaria control trials, the reduction of malaria has reduced all-cause under-5 mortality by as much as 40 percent.

[27] by employing the time-varying version of malaria ecology: this allows us to pursue a longitudinal strategy within countries over time, which, in turn allows for a fixed-effects approach to factor out time-invariant country-specific determinants of fertility.

[FIGURE 2: 1960-2002 AVERAGE MALARIA ECOLOGY INDEX]

It is important to note that this malaria ecology index does not include the number of malaria cases or deaths, but only the potential strength of transmission, which we argue should be exogenous to fertility. Nevertheless, a review of the medical literature shows that malaria may have a direct effect on fertility through malaria-related severe anemia, as well as through increased incidence of hypertensive diseases of pregnancy, spontaneous abortion, and maternal mortality [28, 29, 30, 31].<sup>3</sup> Since these prevent a live birth, and since TFR counts only live births, then this effect should work in the opposite direction of our putative causal model, thereby biasing any net effect of child mortality on fertility toward zero. Likewise, maternal malaria during pregnancy is also associated with low birth weight and increased neonatal and infant mortality—which is in line with our models. That said, there is some evidence that malaria may reduce lactation period (Bates et al. [32]), which might increase fertility through decreased child spacing.

As an added motivation for the adequacy of the malaria ecology instrument, we tested its relationship to national level malaria incidence and mortality data as reported by WHO [33]. We run country-level regressions of the annual incidence (cases per thousand people) on the malaria ecology for that year, using country dummies, country-specific time trends, and weighing observations by population. The results are in Table 1 below,

---

<sup>3</sup> See also [http://www.rbm.who.int/cmc\\_upload/0/000/015/369/RBMInfosheet\\_4.htm](http://www.rbm.who.int/cmc_upload/0/000/015/369/RBMInfosheet_4.htm) for overview information on malaria in pregnancy.

and show that higher values for the index are associated with higher malaria incidence and mortality after including country and period dummies.

[TABLE 1 HERE]

In order for the malaria ecology instrument to be valid, it is important that the instrument affect fertility only through child mortality. Since temperature and precipitation are components of the index and since they could affect fertility through other channels (such as agricultural production as in Jones and Olken [34], and thus the return to children's labor on the field), note that the climatic variables enter the index in a specific nonlinear form emerging from the epidemiological dynamics and so unlikely to be strongly correlated to agriculture or other non-epidemiological systems. Figure 3 below plots the malaria ecology index against temperature and precipitation to illustrate the nonlinearity. Since the human biting rate of the local dominant anopheles vector also goes into the index, the figure shows the value of the function at all levels of precipitation and temperature for a given HBI of 0.5, as well as the actual value of several countries on an average month (these countries are not on the surface because they have HBI different from 0.5). Correlation between the index and agricultural yields is -.35, and malaria ecology has no effect on yields after controlling for temperature, precipitation, and country-specific time trends (this table is available upon request). Moreover the analysis below will demonstrate that the estimation of the child mortality effect is robust to including temperature and precipitation as controls, as well as agricultural yield.

[FIGURE 3: MALARIA ECOLOGY, TEMPERATURE AND PRECIPITATION FOR

HBI = 0.5]

*Data*

We compiled a national-level, cross-country dataset covering the period 1960-2005. The demographic data (fertility, infant mortality and child mortality) come from the U.N. Population Division [35]. All other data sources are described below.

A note on the time series: the time series data we use in this dataset is divided into quinquennia, beginning in 1960 and ending in 2005. This is determined largely by the fact that the U.N. Population Division, our main source for demographic data, uses five year averages for several of the key demographic variables considered in our analysis, such as TFR and child mortality. Whenever we have used yearly time series data, we compute five year averages for the appropriate quinquennia.<sup>4</sup>

The variables used in the analysis are the following:

*TFR*: Total fertility rate, or number of children per woman of reproductive age. Data from the U.N. Population Division reported as 5 year averages.

*Under-5 Mortality Rate*: Data from the U.N. Population Division on the mortality rate in children under the age of 5, collected in averages over 5 year periods. Note that we use the natural logarithm of both total fertility and child mortality; the natural limits to a childbearing during a woman's reproductive lifetime suggest a nonlinear relationship which calls for a log-log approach. Figure 4 shows the bivariate relationship between fertility and child mortality with and without logarithms; evidently the log-log form is appropriate in the linear regression context.

[FIGURES 4A & 4B: TOTAL FERTILITY RATES AND CHILD MORTALITY RATES]

---

<sup>4</sup> The quinquennia are the following: 1960 through 1964; 1965 through 1969; 1970 through 1974; 1975 through 1979; 1980 through 1984; 1985 through 1989; 1990 through 1994; 1994 through 1999; 2000 through 2004.

*Index of malaria ecology.* Malaria Ecology is an ecologically-based spatial index of the stability of malaria transmission based on the interaction of climate with the dominant properties of anopheline vectors of malaria that determine vectorial capacity (Kiszewski et al., [36]). The index is constructed on a 0.5 degree spatial grid to derive the climatic characteristics of individual months, and then averaged over a 12-month period for every year (McCord, [26]). For a complete description of the ME variable see the appendix. Note that the index does not rely on disease incidence or human mortality or morbidity, but instead only on biophysical factors, and therefore provides a variable that is ideally exogenous to human intervention.

[FIGURE 5 HERE]

*Temperature and Precipitation.* The Malaria Ecology variable is constructed using temperature and precipitation data from the Climate Research Unit in East Anglia. The data is available monthly at the 0.5 degree grid cell level, from 1901-2006. (described in New et al., [37]).

*Log of GDP per capita,* measure at purchasing-power-parity, in 2000 international dollars. Yearly data are taken from the World Bank's World Development Indicators database [38]. We calculate our own averages for each 5 year period.

*Average years of females' secondary schooling:* Data are from Barro and Lee [39]. We calculate our own averages for each five-year period. We chose this age band to measure educational attainment since it was presumed that post-puberty adolescence and young adulthood were the critical time periods during which "fertility engines" would ignite (Wu and Martin [40]) and thus the correct population to measure.

*Region:* We use the World Bank regional classifications, with a dummy variable for Sub-Saharan African countries. Please see the World Bank's *World Development Indicators 2009* [38] for more information regarding how specific countries are classified.

*Population:* Total population by country. Five year averages calculated from U.N. Population Division yearly data. We use this variable as a weight for weighted least squares.

*Cereal Yields per Hectare:* Kilograms per hectare of harvested land of wheat, rice, maize, barley, oats, rye, millet, sorghum, buckwheat, and mixed grains [38].

Means and standard deviations for the sample of country-years are presented in Table 2, below. We present un-weighted values, as well as values obtained using weights by population size. Weighting by population size diminishes mean TFR in our sample by less than 1 child per woman (4.71 to 3.84), and reduces the mean child mortality rate by around 15 deaths per 1,000 live births.

[TABLE 2 HERE]

### *Methods and Findings*

In order to evaluate the effects of child mortality on fertility we employ several statistical models on our panel. We first use reduced form OLS, then OLS with fixed effects, and then test which of those two models is preferred. We then move to an instrumental variable framework (instrumenting for child mortality) using malaria ecology. We test for the endogeneity of the child mortality variable and the strength of the malaria ecology instrument in the first stage.

[TABLE 3 HERE]

We begin the reduced form OLS analysis looking simply at regional deviations from a baseline time trend (regression (i) in Table 3). Since we are interested primarily in sub-Saharan Africa, we collapse the other World Bank regions into the suppressed category in order to interpret the difference between sub-Saharan Africa and the rest of the developing world. Note that high-income countries are excluded, since they have completed the demographic transition and intertemporal variation in fertility is likely to be following very different dynamics. China is dummied beginning in 1980 because its fertility rate is artificially suppressed due to the one-child policy in place since 1979. From regression (i) we see that sub-Saharan Africa has a TFR that is roughly 80% higher ( $e^{-.56}$ ) than the rest of the developing world. Regression (ii) then looks at child mortality, the variable of interest. The regression surely suffers from endogeneity (we have discussed above how high fertility can impact child mortality, and how previous studies have found child mortality to be endogenous after formal tests), so while it is not a perfect model and does not allow us to deduce causality or interpret coefficients, it is enlightening to see the strength of the correlations between these variables and fertility. Regressions (iii) – (iv) add other variables which we will use as controls in the analysis (income and women’s education). In this simplified framework, all of the variables come out as significant and with predicted signs (high child mortality is associated with higher fertility, while increasing income and increasing women’s education are associated with lower fertility). Note that there are several observations for every country in the sample, so we estimate Huber-White standard errors robust to clustering within countries in regressions (i) – (iv). Regression (v) is estimated using fixed-effects. All variables

maintain their significance and sign. Note that the sub-Saharan Africa dummy with the controls in these “naïve” models comes down from .56 in every case, and as low as 0.21 when child mortality is included. This would suggest that the high African fertility rates relative to the rest of the world are largely (though not exclusively) explained by high child mortality levels. When the control variables are all included in this naïve regression, the variables are all significant, but the coefficient on income switches sign, owing most probably to the multicollinearity between the variables. Regression (v) moves to a fixed-effects framework by inserting country dummies (this now discards all cross-country variation and uses the controls to explain within-country variation in TFR). All the variables are now significant when included jointly. A Hausman test to compare the two specifications (an OLS and fixed-effects version of regression (v)) rejected the null that the OLS model is valid, leaving us with the fixed effects specifications as preferred.

We tested the robustness of the FE regression (v) by omitting the income variable, since the purported endogeneity of income might be biasing other coefficients. Both the child mortality and female education variables remain significant and of comparable magnitude.

[TABLE 4 ABOUT HERE]

We then move to the instrumental variables framework, where the time-varying malaria ecology variable instruments for child mortality. All regressions include country dummies to reduce omitted variable problems, all include global period dummies to flexibly control for global trends, and all weight the observations by population. Note that following Bertrand et al [41], the serial correlation in our dependent variable

(fertility) calls for clustering by country in all regressions. Regression (i) is the first stage of the basic regression, showing that the exogenous malaria ecology is leading to increases in child mortality. The second stage is (ii), where the instrumented child mortality variable has a coefficient of 0.65 (larger than the estimates in Table 3). In order to evaluate the lagged effect of changes in child mortality on fertility, we ran regression (ii) and its first stage using the contemporary and one of 5-year, 10-year, 15-year, and 20-year lagged values of child mortality instrumented by the lagged malaria ecology (these results are not reported in the table). Only with the 10-year lag did the child mortality coefficient remain significant, and the magnitude of the contemporary and lagged coefficients were about equal and half of the coefficient in regression (ii). As an added robustness test, we included different leads to the regression with a 10-year lag, and in all cases a lead caused all coefficients to lose significance, perhaps due to lack of power or to collinearity in variables over time. Regression (iv) and its first stage (iii) add temperature and precipitation to show that the malaria ecology instrument and the estimate for child mortality are robust to those controls (the coefficients hardly change), and in fact malaria ecology is capturing the nonlinear relationship between climate and disease transmission. To further assuage concerns that the malaria ecology instrument might be affecting fertility through a channel other than child mortality, we control for cereal yields per hectare in regression (vi) and find that instrument is still significant in the first stage and the coefficient on child mortality drops to 0.55. Adding females' years of secondary schooling in (viii) reduces the child mortality coefficient to 0.59 (though the education variable itself isn't significant in the second stage). Including income in (x) does not change the results from the basic regression. As a robustness check, regression

(xi) limits the sample to low-income countries and the coefficient on child mortality increases to 0.86.

We proceed to run some tests on the malaria ecology instrument. First we test the endogeneity of the child mortality variable (the null of exogeneity can be rejected at a one percent alpha level) and for the strength of the instrument (the F-tests for the instrument in the first stage are reported in the table; most are above Staiger and Stock's [42] rule of thumb value of  $F=10$ , so we can conclude that the instrument is strong). To further dispel concerns of a weak instrument, we follow Stock and Yogo [43] and compute the Cragg-Donald statistic. In all cases, the statistic exceeds the critical value of 16.38, suggesting that significance test distortions due to a weak instrument are not a concern here. Next, we run all specifications with a limited information maximum likelihood estimator (Fuller [44]) with a modification parameter of  $\alpha = 1$ , which is robust than 2SLS with a weak instrument (Hahn, Hausman, and Kuersteiner [45], Hausman, Stock and Yogo [46]). In all cases, the coefficient child mortality remained significant and of similar magnitude compared to the 2SLS: it increased to 0.97 in (ii), did not change in (iv), (vi), (viii) and (x), and decreased to 0.37 in (xi).

All regressions were re-estimated without the population weights; results were qualitatively unchanged (the coefficient on mortality ranges between 0.55-0.75 and is always significant; in addition, the female education variable is significant in the second stage without the population weights). Finally, we dispense with clustering and analytical weights in order to estimate the basic model in (i) and (ii) using first differences instead of fixed effects. The results are qualitatively similar: the child mortality coefficient drops from 0.54 under FE to 0.45 under FD. The coefficient on the

serial correlation of errors is 0.2, which does not strongly warrant one model above the other. We opt to report FE results throughout so that population weights and clustering of standard errors can be employed.

The effect of child mortality on fertility is overwhelmingly robust across specifications. The income and education variables are less robust: after child mortality is instrumented, they are not robust to clustering by country. We do not focus on interpreting their coefficients due to the potential endogeneity problems (fertility might affect these control variables), but they serve to give a range of plausible estimates on the child mortality variable given potential omitted variable bias. The coefficients on child mortality have ranged between 0.10 in the non-IV fixed-effects specification to 0.89 in the IV fixed-effects specification. The preferred specifications, however, would be the IV regressions with country dummies, population-weighted observations, and clustering by country, which gave a range of 0.55 – 0.86. This range corresponds to having a decrease of TFR of 37% - 57% (or a decrease of 2.3 – 3.6 births from the 1990 level of 6.3) if child mortality were to decrease by 66% from its 1990 value in sub-Saharan Africa (the Millennium Development Goal for child mortality). Even the smallest coefficient in the IV specification yields a powerful effect: a TFR reduction of 2.3 if child mortality is reduced by 66%. This points to the fact that child mortality still plays a powerful role in fertility choice in the 20<sup>th</sup> century, an entire order of magnitude larger than the effect identified in some less well-identified studies mentioned previously.

Before concluding, it is worth pausing to discuss the generalizability of the child mortality coefficient estimated through IV. Some authors have expressed concern with interpreting coefficients from IV too generally: the coefficient from IV only is the local

average treatment effect for the sub-population for which the instrument is well-correlated to the endogenous variable. In order to make sure that we can draw policy implications for sub-Saharan Africa, we check the first stage regression of under-5 mortality on malaria ecology using only sub-Saharan African observations. We find that the coefficient on malaria ecology is positive and significant, indicating that sub-Saharan Africa is a "complier" in the sense that the sub-sample exhibits a strong first stage. In the second stage, the estimated coefficient on child mortality rises to 1.11. We opt for focusing on a coefficient identified from the larger sample not only to be conservative but because it is likely that the responsiveness of fertility to changes child mortality is likely to go down as the demographic transition progresses. Therefore, to think about the average elasticity of child mortality across the whole transition and not just in pre-transition countries, we report estimates using the entire developing world.

A second concern is that our child mortality coefficient might only be valid for children dying due to malaria. However, there is nothing in the demography literature or theory that identifies a differentiated fertility response to a child death depending on the cause of death (though one might imagine different fertility responses between health-related deaths and conflict-related deaths, for example). Nonetheless, whether there is a differentiated fertility response depending on cause of child death is beyond the scope of this paper, but in the absence of such evidence we maintain that the child mortality coefficient that we have estimated is a step forward in the quantitative study of fertility dynamics.

### *Conclusion*

We are only the latest in a long research tradition to test the link between child mortality and fertility. Where we have added to the literature is through the deployment of a novel instrument to estimate this putatively important relationship: Malaria ecology. Instrumenting child mortality in this way yields a result that appears robust across a number of models that include or exclude various control variables, time and country fixed effects and with or without population weights. In all our models, we find that child survival is a quantitatively important and robust driver of fertility. This is where the theory of the demographic transition started: save children and families will choose to have fewer children. The finding is consistent with the fact that the demographic transition has proceeded in the widest range of social settings: rural and urban, male-dominated and gender equal, impoverished and middle-class (Cleland [2001]). Since the child mortality transition has also proceeded in a wide range of settings, it is plausible that fertility has been driven largely by changes child mortality.

This is not to say, however, that child mortality is the only important driver of a demographic transition. As explained in Bongaarts et al [1], family planning through contraception, sexual behavior and abortion, are proximate determinants of fertility which are affected by socioeconomic variables such as income, education, and health. If socioeconomic and policy variables are to successfully trigger a transition, then the change in desired fertility must be allowed to translate into a change of actual fertility through the proximate variables (especially contraception use). Though we have focused on better quantifying the role of reducing child mortality, the importance of other variables affecting fertility such as female education or access to family planning should not be understated.

If this conclusion is correct, it heralds the possibility of a rapid fertility transition in Africa. Child survival can be dramatically improved in a short period of time (UN Millennium Project [47]). It is indeed possible, even in a very low income setting and within a five-year period, to reduce the child mortality rate from 150 per 1,000 or above, to well below 100 per 1,000. Similarly, other determinants of fertility, such as increases in agricultural yields and public policies in support of family planning, may change in just a few years, in contrast with many socioeconomic variables. Sub-Saharan Africa's TFR in 1990 was 6.3, and in 2007 it was 5.1 (World Bank, [38]). If the Millennium Development Goals are achieved and child mortality is decreased by two-thirds compared to 1990, our results would predict a decrease in TFR of around 3 babies to 3.3, or more than half of the reduction necessary to reach replacement fertility of 2.1. If in addition female education is increased, access to family planning is improved, and farm yields increased, the resulting fertility reductions would correspond to a new TFR in sub-Saharan Africa of around 2.1. While these predictions are very crude, nevertheless the prospect for a rapid, policy-supported transition to lower fertility in Africa, say a TFR of less than 3.0 by 2015, looks reasonable. The results also point to the continued importance of child mortality in fertility declines: while other variables emphasized in research and policy (such as family planning or female education) are important contributors to decline in fertility, child mortality declines seem to account for around half of the fertility decline and its importance in demographic dynamics should not be ignored.

More research is needed to continue exploring the causes of rapid fertility transitions seen in some countries in the last few decades. In particular, our models do

not attempt to solve the endogeneity issues with the control variables, nor attempt to understand the lag structure in changes to the independent variables (though, as mentioned above, changes in child mortality are found to have lagged effects on fertility). Future work could look intensively at case studies of rapid fertility declines and attempt to better model lagged effects of the determinants of TFR. In the meantime, however, this paper provides a more reliable range of estimates for how changes in child mortality causally translate into changes in fertility behavior. To our knowledge, no paper has employed a time-varying exogenous instrument to solve the endogeneity problem between fertility and child mortality. The results are also encouraging: whereas previous estimates of the elasticity of fertility to child mortality using IV have failed to find statistical significance or have found perplexingly small magnitudes, our result finds that child mortality is likely to account for over half of the fertility reduction during the demographic transition.

### **Acknowledgements**

The authors thank Samuel Freeman for excellent research assistance, as well as seminar participants at the 2006 NBER Health Economics Summer Institute, the Duke Global Health Seminar, the Watson Institute Colloquium at Brown University, Sociology Department Colloquia at Johns Hopkins University and the University of Florida, the Development Economics Seminar at New York University, the Joint Harvard-MIT Economic Sociology Seminar, the Population Studies Colloquium at Pennsylvania State University, the Sustainable Development Seminar at Columbia University and at the 2010 World Congress of Environmental and Resource Economists for useful comments.

## References

- [1] J. Bongaarts, O. Frank, R. Lesthaeghe, The proximate determinants of fertility in sub-Saharan Africa. *Population and Development Review*, v. 10, no. 3 (Sept), pp. 511-537. 1984
- [2] S. A. Marandi, A. Mehryar, M. J. Abbasi-Shavazi, Z. Majdfar, M. H. Shamsian. 2006. "Iran: From Family Sizes of Six to Replacement Levels in 15 Years," in Tobias, Michael, Bob Gillespie, Elizabeth Hughes, and Jane Gray Morrison. No Vacancy: Global Responses to the Human Population Explosion. Pasadena, CA: Hope Publishing House.
- [3] J. Cleland, J. F. Phillips, S. Amin, and G.M. Kamal. "The Determinants of Reproductive Change in Bangladesh: Success in a Challenging Environment." World Bank: Washington DC, 1994.
- [4] A. Aghajanian. 1991. "Population Changes in Iran, 1966-86: A Stalled Demographic Transition?" *Population and Development Review*. Vol. 17, No. 4, pp. 703-15.
- [5] A. E. Raftery, S. M. Lewis and A. Aghajanian. 1995. "Demand or Ideation? Evidence from the Iranian Marital Fertility Decline." *Demography* Vol. 32, No. 2, pp. 159-82.
- [6] D. Conley and R. Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources*.
- [7] S. Joshi and T. P. Schultz. 2007. "Family Planning as an Investment in Development: Evaluation of a Programs' Consequences in Matlab, Bangladesh." Economic Growth Center Discussion Paper 951. Yale University, New Haven, Connecticut.
- [8] M. Mogstad and M. Wiswall. 2009. Family Size and Children's Education: How Linear Models Can Mask a Non-Linear Relationship. NYU Economics Working Paper.
- [9] M. Hazan and B. Berdugo. 2002. Child Labour, Fertility, and Economic Growth. *The Economic Journal*, v. 112, pp. 810-828.
- [10] O. Moav. 2005. Cheap Children and the Persistence of Poverty. *The Economic Journal*, v. 115, pp. 88-110.
- [11] A. Chowdhury. 1988. The Infant Mortality-Fertility Debate: Some International Evidence. *Southern Economic Journal*, v. 54, no. 3, pp. 666-674.
- [12] D. M. Heer. and D. O. Smith. 1968. Mortality Level, Desired Family Size, and Population Increase. *Demography*, v. 5, no. 1, pp. 104-121.

- [13] J. Cleland. 2001. "The Effects of Improved Survival on Fertility: A Reassessment." *Population and Development Review*, Vol. 27, Supplement: Global Fertility Transition, pp. 60-92.
- [14] G. Becker. and R. Barro. 1988. A Reformulation of the Economic Theory of Fertility. *The Quarterly Journal of Economics*, v. CIII, no. 1, pp. 1-25.
- [15] M. Doepke. 2005. Child mortality and fertility decline: Does the Barro-Becker model fit the facts? *Journal of Population Economics*, v. 18, pp. 337-366.
- [16] M. Zakir and P. Wunnava. 1999. Factors affecting infant mortality rates: evidence from cross-sectional data. *Applied Economics Letters*, v. 6, pp. 271-273.
- [17] P.R. Galloway, R. Lee and E. A. Hammel. 1998. "Infant mortality and the fertility transition: macro evidence from Europe and new findings from Prussia." in *From Death to Birth: Mortality Decline and Reproductive Change*. Montgomery, Mark R. and Barney Cohen, eds. Washington, D.C., National Academy Press.
- [18] P. Lorentzen, J. McMillan, and R. Wacziarg. 2005. "Death and Development." NBER Working Paper No. 11620.
- [19] Schultz, T. Paul. 1997. Demand for Children in Low Income Countries. in Rosenzweig, M. R. and Stark, O. (eds.) *Handbook of Population and Family Economics*.
- [20] Benefo, Kofi and Schultz, T. Paul. 1996. Fertility and Child Mortality in Cote d'Ivoire and Ghana. *The World Bank Economic Review*, v. 10, no. 1, pp. 123-158.
- [21] Dreze, Jean and Murthi, Mamta. 2001. Fertility, Education, and Development: Evidence from India. *Population and Development Review*, v. 27, no. 1, pp. 33-63.
- [22] Kalemli-Ozcan, Sebnem. 2006. "Aids, 'Reversal' of the Demographic Transition and Economic Development: Evidence from Africa." NBER Working Paper Series, w12181,
- [23] Black R.E., Morris S.S., and Bryce J. 2003. Where and why are 10 million children dying every year? *Lancet*, v. 361, pp. 2226-2234.
- [24] Sachs, J.D. 2003. Institutions Don't Rule: Direct Effects of Geography on Economic Development. *NBER Working Paper 9490*.
- [25] Carstensen, Kai and Gundlach, Erich. 2006. The Primacy of Institutions Reconsidered: Direct Income Effects of Malaria Prevalence. *The World Bank Economic Review*, v. 20, no. 3, pp. 309-339.
- [26] McCord, Gordon C. 2010. Improving Ecological Modeling of Malaria Risk using Time-Variant Malaria Ecology. Unpublished working paper.

- [27] Conley, D., McCord, G.C., and Sachs, J.D. "Africa's Lagging Demographic Transition: Evidence from Exogenous Impacts of Malaria Ecology and Agricultural Technology." *National Bureau of Economics Research Working Paper No. 12892*. February 2007.
- [28] Etard, J.F., Kodia, B. and Ronsmans, C. 2003. Seasonal Variation in Direct Obstetric Mortality in Rural Senegal: Role of Malaria? *American Journal of Tropical Medicine and Hygiene*, v. 68, no. 4, pp. 503-504.
- [29] Maubert, Bertrand, Fievet, Nadine, Tami, Germaine, Cot, Michel, Boudin, Christian, and Deloron, Philippe. 1999. Development of Antibodies against Chondroitin Sulfate A-Adherent *Plasmodium falciparum* in Pregnant Women. *Infection and Immunity*, v. 67, n. 10, pp. 5367-5371.
- [30] Sartelet, H., Rogier, C., Milko-Sartelet, I., Angel, G., and Michel, G. Malaria associated pre-eclampsia in Senegal. 1996. *Lancet*, v. 347, n. 9008, p. 1121.
- [31] Guyatt, H.L., and Snow, R.W. 2001. Malaria in pregnancy as an indirect cause of infant mortality in sub-Saharan Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, v. 95, n. 6, pp. 569-576.
- [32] Bates I, Fenton C, Gruber J, Laloo D, Medina Lara A, Squire SB, Theobald S, Thomson R, Tolhurst R. 2004. Vulnerability to malaria, tuberculosis, and HIV/AIDS infection and disease. Part 1: determinants operating at individual and household level. *Lancet Infectious Disease*, May; 4(5): 267-77.
- [33] World Health Organization. 2008. *World Malaria Report 2008*. Geneva.
- [34] Jones, Benjamin F. and Benjamin A. Olken. 2010. "Climate Shocks and Exports." NBER Working Paper Series, No. 15711.
- [35] United Nations Population Division. 2004. *World Population Prospects: the 2004 Revision*.
- [36] Kiszewski, Anthony, Mellinger, Andrew, Spielman, Andrew, Malaney, Pia, Sachs, Sonia Ehrlich, and Sachs, Jeffrey. 2004. A Global Index Representing the Stability of Malaria Transmission. *American Journal of Tropical Medicine and Hygiene*, v. 70, no. 5, pp. 486-498.
- [37] New, M. Lister, D., Hulme, M. and Makin, I. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* 21.
- [38] World Bank. 2009. *World Development Indicators*. Accessed from <http://publications.worldbank.org/WDI/>

- [39] Barro, Robert J and Jong-Wha Lee. A New Data Set of Educational Attainment in the World, 1950-2010. NBER Working Paper 15902. April 2010.
- [40] Wu, Lawrence and Stephen Martin. 2002. "Is There an Engine of Nonmarital Fertility?" University of Wisconsin, CDE Working Paper No. 2002-14.
- [41] Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*. Vol. 119, No.1, pp. 249-75.
- [42] Staiger, Douglas and James H. Stock. 1997. "Instrument Variables with Weak Instruments." *Econometrica* 65(3): 557-86.
- [43] Stock, James H. and Motohiro Yogo. 2002. "Testing for Weak Instruments in Linear IV Regressions." *NBER Technical Working Paper*. No. 284. National Bureau of Economics Research. Cambridge, MA.
- [44] Fuller, Wayne A. 1977. "Some Properties of a Modification of the Limited Information Estimator." *Econometrica* 45(4): 939-54.
- [45] Hahn, Jinyong, Jerry Hausman and Guido Kuersteiner. 2004. "Estimation with Weak Instruments: Accuracy of Higher-Order Bias and MSE Approximations." *Econometrics Journal* 7(1):272-306.
- [46] Hausman, Jerry, James H. Stock, and Motohiro Yogo. 2005. "Asymptotic Properties of the Hahn-Hausman Test for Weak Instruments." *Economics Letters* 89(3):333-342.
- [47] United Nations Millennium Project. *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals*. Earthscan: New York, 2005.

**Table 1: Malaria Ecology and Malaria Incidence & Mortality**

<u>Independent Variables</u>	(i)	(ii)
	ln( Malaria cases per 1,000 population )	ln( Malaria deaths per 1,000 population )
Malaria Ecology	0.25* (1.57)	0.37*** (2.53)
N =	1480	565
Within R-squared =	0.09	0.25

t-statistic in parentheses, \* indicates  $\alpha = 0.15$ , \*\*\*  $\alpha = 0.05$

Regressions include country and year dummies and a constant (not reported)

Regressions weigh observations by population

Within R-squared computed without clustering

**Table 2: Means and Standard Deviations of Variables in Sample**

<u>Variable</u>	<u>Mean</u>		<u>Standard Deviation</u>	
	Un-weighted	Population weighted	Un-weighted	Population weighted
Total Fertility Rate	4.71	3.84	1.93	1.70
Under 5 mortality rate (per 1,000)	108.8	93.64	80.3	62.98
Malaria Ecology Index	2.73	1.49	3.97	3.01
Average Temperature (°C)	20.4	17.38	7.36	9.2
Average Precipitation (mm)	101.5	86.27	70.0	56.7
GDP per capita, PPP, constant \$	4,224	2,942	3,838	2,985
Females' Years of Secondary Schooling	1.11	1.01	1.10	0.84
Average Population (millions)	28.9	475	114	472
Cereal Yields per Hectare (kg/ha)	1,682	2,459	1,054	1,293
Year	1986.5	1988.9	12.5	12.0

**Figure 1**

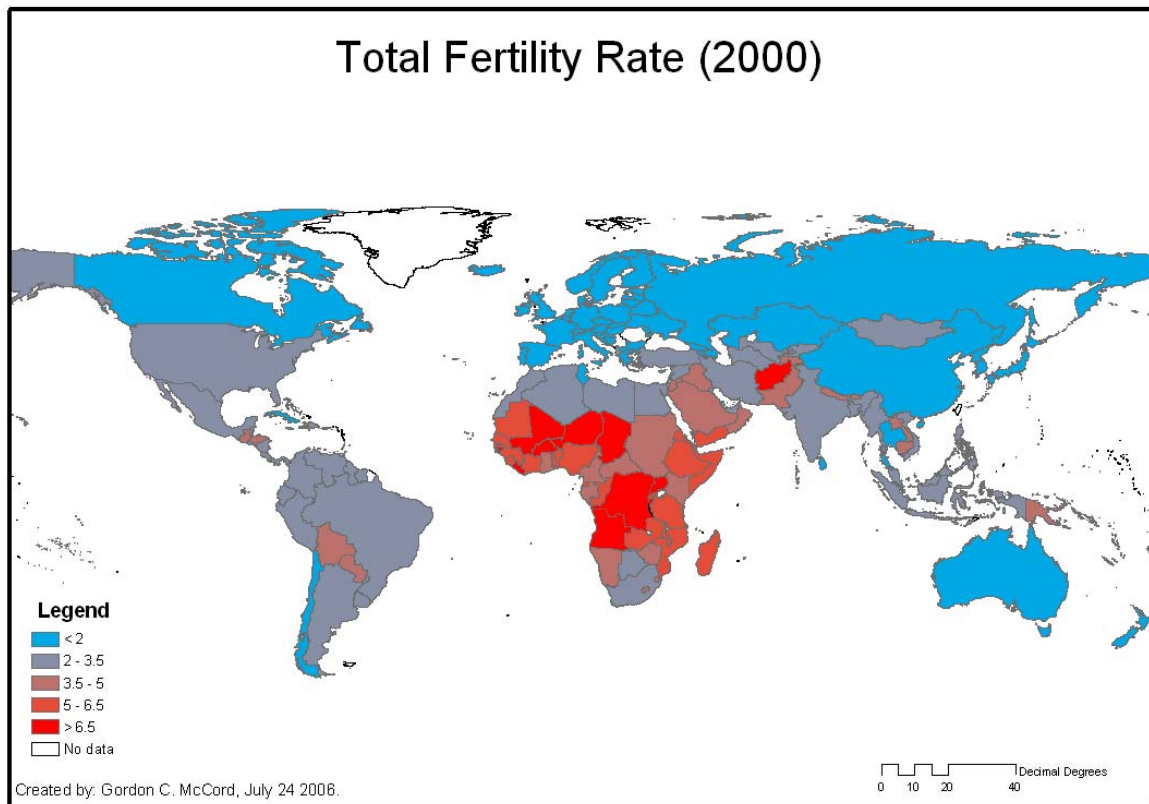
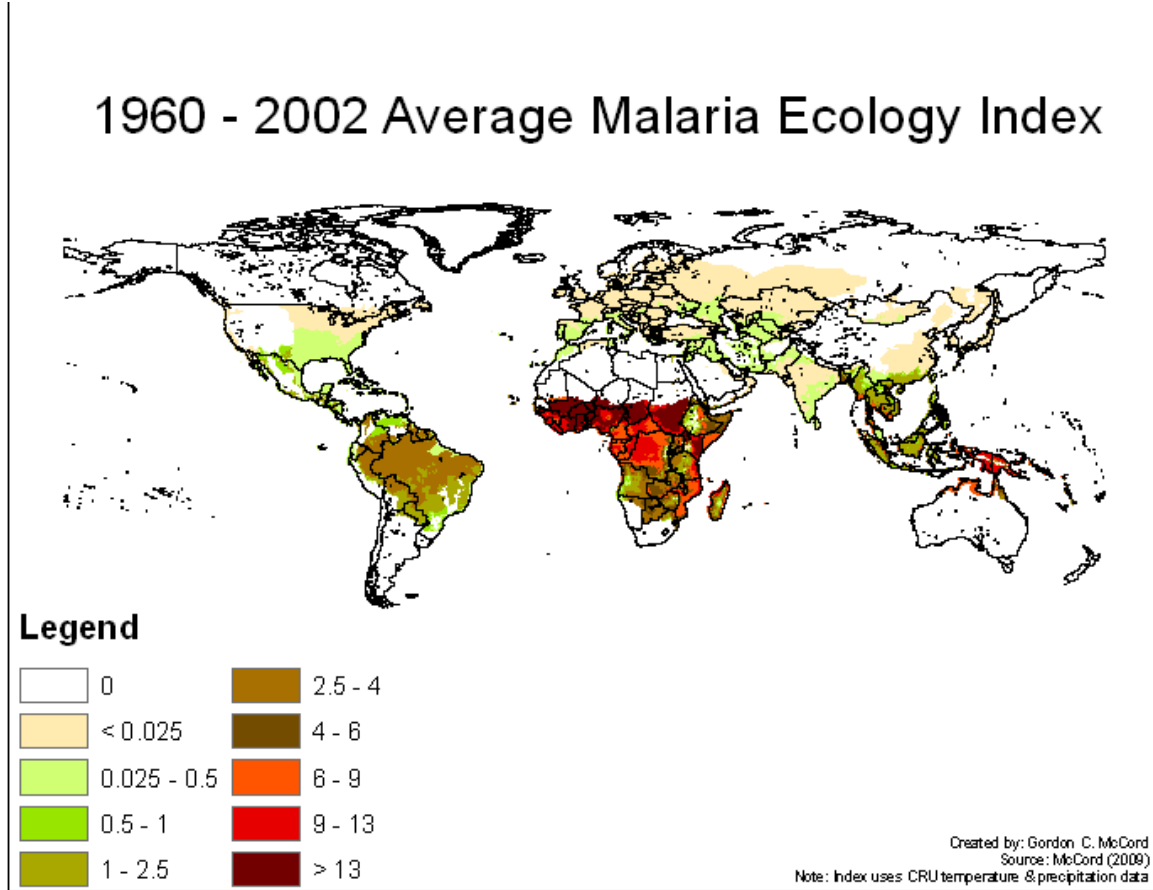


Figure 2



**Figure 3: Malaria Ecology & Temperature, Precipitation**

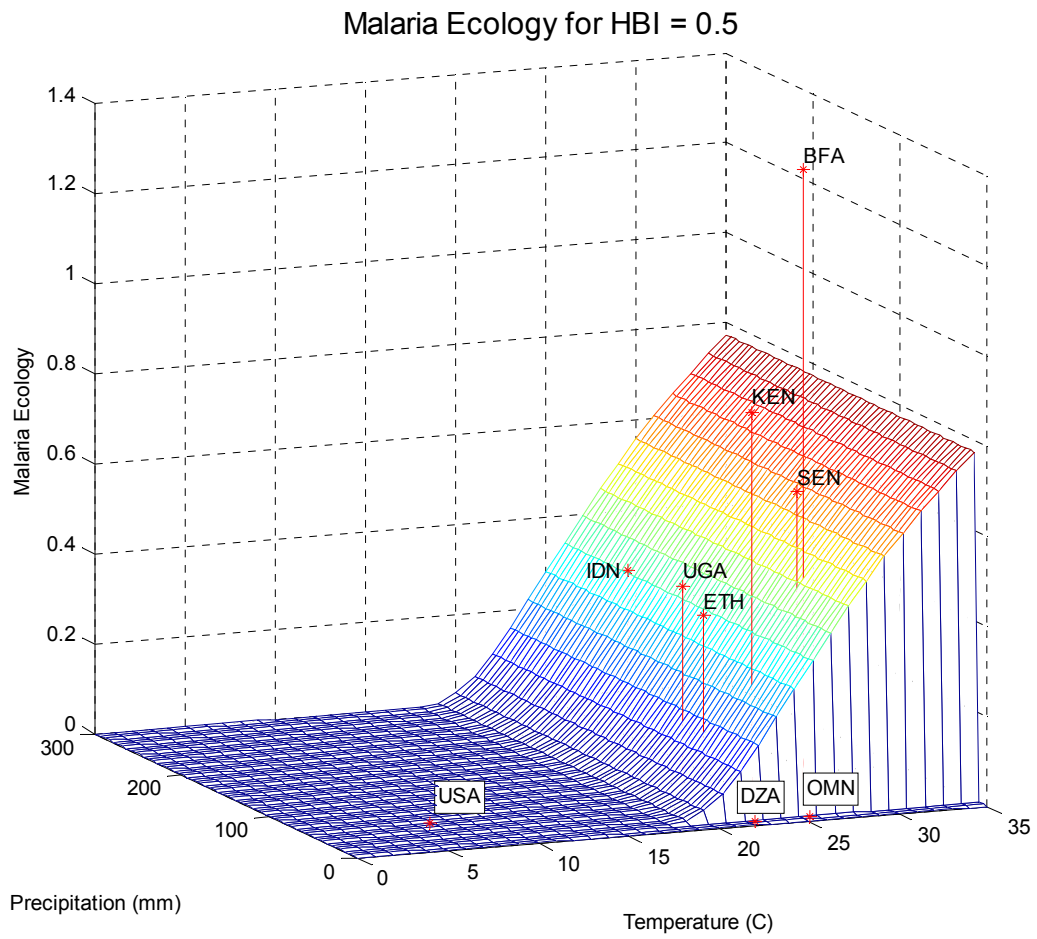
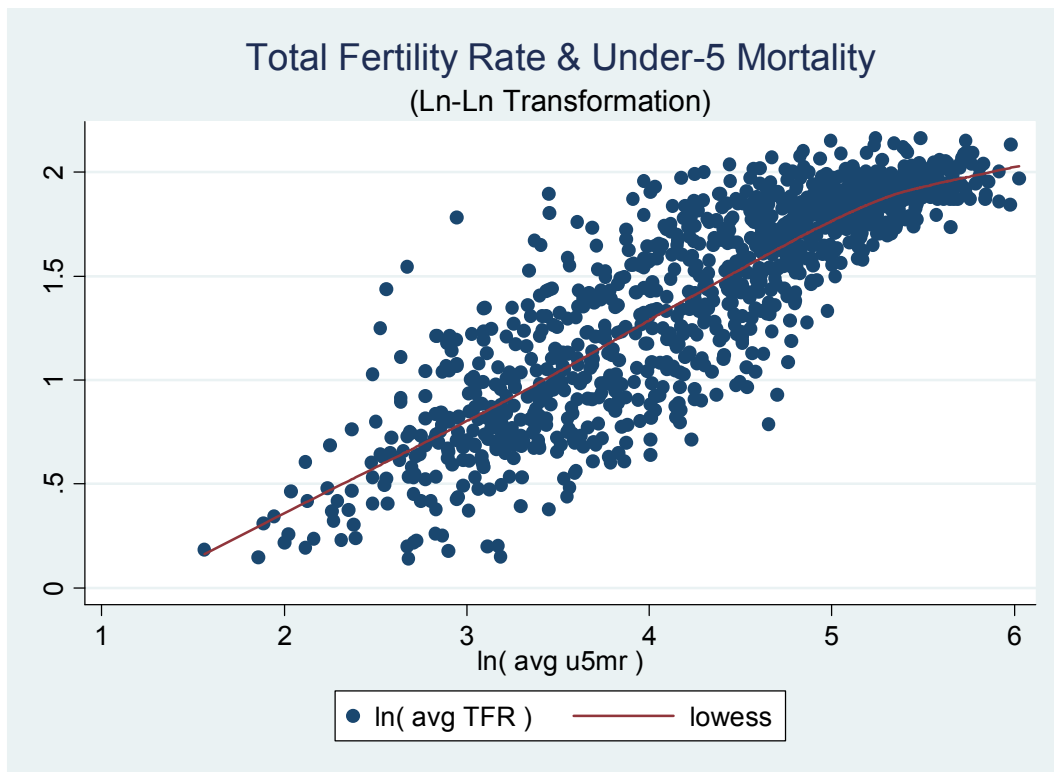
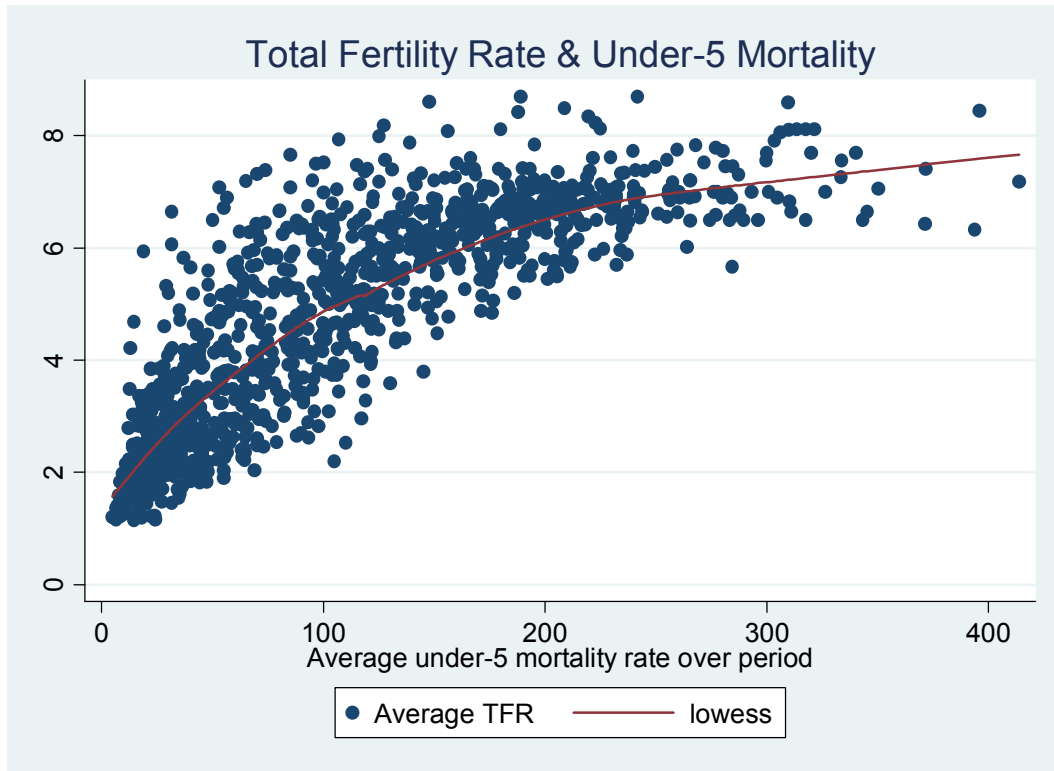


Figure 4: Fertility and Child Mortality



**Table 3: OLS and Fixed-Effects Regression of Total Fertility Rates**

	(i)	(ii)	(iii)	(iv)	(v)
Dependent Variable:	ln( TFR )	ln( TFR )	ln( TFR )	ln( TFR )	ln( TFR )
<u>Independent Variables</u>					
Sub-Saharan Africa dummy	<b>0.56</b> (7.89)	<b>0.21</b> (3.24)	<b>0.37</b> (8.27)	<b>0.53</b> (7.49)	
ln( under-5 Mortality Rate )		<b>0.41</b> (7.30)			<b>0.10</b> (1.80)
Females' Years of Secondary Schooling			<b>-0.26</b> (-5.80)		<b>-0.07</b> (-1.58)
ln( GDP per capita, in PPP prices & constant \$)				<b>-0.12</b> (-2.49)	<b>0.06</b> (1.89)
China One-Child-Policy Dummy	<b>-0.47</b> (-7.35)	<b>-0.28</b> (-6.01)	<b>-0.44</b> (-11.47)	<b>-0.53</b> (-12.17)	0.02 (0.72)
N =	1255	1142	896	735	536
R-squared =	0.63	0.85	0.76	0.71	0.83

t-statistics in parentheses; bold indicates significant to 95% confidence.

All regressions include year dummies and a constant, which are not reported.

All regressions weigh observations by population and report robust standard errors clustered by country

Regression (v) includes country dummies

All regressions exclude high-income countries

R-squared for (v) is within R-squared, calculated by running regression without weights or clustering

**Table 4: Instrumenting Child Mortality using Malaria Ecology**

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)
Dependent Variable:	ln( u5MR )	ln( TFR )	ln( u5MR )	ln( TFR )	ln( u5MR )	ln( TFR )	ln( u5MR )	ln( TFR )	ln( u5MR )	ln( TFR )	ln( TFR )
Independent Variables											
Malaria Ecology	<b>0.29</b> (9.13)		<b>0.29</b> (9.09)		<b>0.27</b> (7.99)		<b>0.20</b> (4.82)		<b>0.22</b> (6.00)		
ln( under-5 Mortality Rate )		<b>0.65</b> (5.10)		<b>0.71</b> (5.15)		<b>0.55</b> (3.97)		<b>0.59</b> (2.82)		<b>0.68</b> (3.04)	<b>0.86</b> (6.51)
Females' Years of Secondary Schooling							<b>-0.09</b> (-2.92)	0.03 (0.51)			
ln( GDP per capita, in PPP prices & constant \$)									0.006 (0.26)	0.026 (0.54)	
Average temperature over period			-0.03 (-0.98)	-0.08 (-1.77)							
Average precipitation over period			-0.001 (-1.12)	0.011 (0.84)							
ln( Cereal Yield per Hectare )					<b>-0.18</b> (-3.33)	-0.12 (-1.17)					
China One-Child-Policy Dummy	<b>-0.12</b> (-4.58)	<b>-0.21</b> (-6.13)	<b>-0.11</b> (-3.87)	<b>-0.18</b> (-4.93)	<b>-0.06</b> (-2.19)	<b>-0.19</b> (-7.87)	<b>-0.08</b> (-2.66)	<b>-0.22</b> (-7.54)	<b>0.1</b> (2.37)	-0.03 (-0.68)	
N =	1134	1134	1134	1134	1001	1001	857	857	717	717	477
F-test on instrument =	14.34		14.18		11.96		4.4		9.71		30.2
Cragg-Donald Statistic =	83.3		82.59		63.84		23.23		35.95		107.43
Within R-squared =	0.77	0.75	0.77	0.73	0.78	0.74	0.80	0.82	0.71	0.72	0.53

t-statistics in parentheses; bold indicates significant to 95% confidence.

All regressions include country dummies, time period dummies and a constant (not reported).

Regressions (i), (iii), (v), (vii), and (ix) are first stage for (ii), (iv), (vi), (viii), and (x) respectively. First stage for (xi) not shown, but the malaria ecology is significant.

Regressions (i) - (xi) weigh observations by population

All regressions exclude high-income countries, and regression (xi) only includes low-income countries.

All regressions reports robust standard errors clustered by country

Within R-squared computed without clustering or analytical weights

## **Appendix: Time-Varying Malaria Ecology Index**

The time-varying malaria ecology index used in this paper was developed by McCord (2010) as an extension of the Malaria Ecology index developed by Kiszewski et al. (2004). Malaria Ecology is an ecologically-based spatial index of the stability of malaria transmission based on the interaction of climate with the dominant properties of anopheline vectors of malaria that determine vectorial capacity. Malaria is a disease of climate because a key part of the life cycle of the parasites (sporogony) depends on a high ambient temperature and because vectors require sufficient rainfall to provide breeding sites. Additionally, the intensity of malaria transmission depends on the specific mosquito species that are present and their relative attraction to humans versus animals. The Malaria Ecology variable published by Kiszewski et al. measures the effects of ambient temperature (using a monthly average from 1901-1990) on the force of transmission of malaria, as expressed through the length of the extrinsic incubation period, and therefore the proportion of the vector population able to survive long enough to become infectious. The index is constructed on a 0.5 degree spatial grid to derive the climatic characteristics of individual months, and then averaged over a 12-month period. The first step is to identify the distribution of anopheline species across the world using observation records and satellite-based vegetation maps to identify likely habitats where observations have not been recorded.

A dominant species is identified for each spatial zone, and for each month (in cases where there is a seasonal pattern to the dominant species). An ecological screen was created for the presence or absence of a vector during particular months. (For those vectors that breed mainly in temporary water, a minimum precipitation threshold of 10mm per month, lagged one month, is used to judge when the vector would be present in the site during a given month. Vectors that mainly exploit permanent or semi-permanent bodies of water were considered to be independent of seasonal fluctuations in rainfall unless empiric evidence indicated otherwise. In temperate or altitudinous regions, temperature thresholds are used to determine whether parasites can develop in mosquito vectors in a particular month, assuming that malaria parasites cannot develop when the mean monthly temperature remains below 15°C). Note that the mosquito presence screen is ecology-based and not affected by human activity; indeed, it is worth keeping in mind that public health interventions against malaria serve to break the transmission cycle, but do not eliminate the presence of the vector itself (even until today, *Anopheles* mosquitoes capable of transmitting malaria can be found throughout the US and Europe, places where malaria has been largely eradicated).

The basic formula for Malaria Ecology combines climatic factors, the presence of different mosquito vector types and the human biting rate of the different mosquito vectors. The index expresses the factors that most powerfully and perennially influence the intensity of malaria transmission. It uses, therefore, a subset of the vectorial capacity equation without terms for mosquito abundance, vector competence, or recovery rate for infected people. To calculate the duration of the extrinsic incubation period “E,” the

index (1) was calculated for each month, and biting activity was designated based on the average monthly temperature and Moshkovsky's degree-day-based formulae:

$$\sum_{m=1}^{12} a_{i,m}^2 p_{i,m}^E / -\ln(p_{i,m})$$

where:

m = month (1-12)

i = identity of dominant vector

a = proportion biting people (0-1)

p = daily survival rate (0-1)

E = length of extrinsic incubation period (in days) =  $111 / \text{temperature} - 16$

Because it is built upon climatological and vector characteristics on a cell-by-cell basis, Malaria Ecology is exogenous to public health interventions and economic conditions, and thus can serve as an instrumental variable in regressions of economic performance on malaria risk. The Malaria Ecology index correlates strongly to malaria incidence, especially in the absence of public health interventions. Even a cursory comparison of the map of the geographic extent of malaria risk and Malaria Ecology show clear similarities: the high latitudes of the northern hemisphere are both where malaria was eliminated first, and also the malaria ecology index is lowest without being zero.

McCord (2010) re-calculates the index month-by-month using the same methodology using monthly data for temperature and precipitation. The index is aggregated to the country level without weighting by population to prevent potential endogeneity if humans can migrate internally to adapt to malaria prevalence.