The disease environment, schooling, and development outcomes: Evidence from Ethiopia

Abstract

The disease environment could help explain underdevelopment in Africa. This paper shows that local malaria risk is associated with worse local development outcomes. Combining an Ethiopian household survey with satellite-derived topographical information, the paper shows that malaria incidence is correlated with village elevation, slope, and their interaction. That is, malaria is sensitive to elevation in flatlands, where the habitat is suitable for mosquito breeding, but not in steeper lands. Using topography as a predictor of the disease environment, education levels are found to be negatively correlated with malaria. I find suggestive evidence that some other outcomes are related to malaria risk. Finally, the performance of topography predictors is assessed against other climate-based predictors of malaria.

1. Introduction

Debates over the lack of development in sub-Saharan Africa often mention the role of disease in creating a hostile environment that impedes human and physical capital investments. Gallup and Sachs (2001) have famously used cross-country growth regressions augmented with data on malaria incidence to argue that tropical diseases like malaria keep mortality high, reduce human capital, and pose a constraint to economic growth. Other papers have shown considerable health and human capital benefits of malaria eradication campaigns that have taken place outside the African continent. Currently lacking are measures of the impact of the disease environment within Africa. This paper addresses this by providing new evidence on the relevance of the disease environment along several dimensions of development—including schooling, wealth, and exposure to shocks—within the context of Ethiopia.

Ethiopia is a particularly suitable country to examine this issue for two reasons. First, due to a significant presence of the type of the parasite *Plasmodium Falciparum*, malaria is a leading cause of morbidity and mortality in the country (Ethiopian Ministry of Health, 2006). Second, there are significant and largely predetermined differences in disease environment, such that even villages in close proximity to one another can have starkly different health risk profiles. In particular, Ethiopian villages vary considerably in their exposure to malaria, due to the complex topography of this mountainous country, and the sensitivity of malaria transmission to local differences in elevation and slope. In this paper, I use topographical differences to predict malaria incidence. In particular, I show that the *interaction* between elevation and slope predict differences in malaria incidence in a way that is consistent with our understanding of malaria transmission. I then use predicted malaria to obtain a measure of the correlation of malaria exposure to schooling and other development outcomes.

I study the correlation between local malaria endemicity and individual outcomes by matching a large-scale household survey of rural households with satellite-derived weather and topographic maps, which provided geographical and meteorological information on each village and include elevation, temperature, rainfall levels, and slope. The survey used, the Welfare Monitoring Survey (WMS) of 2004, surveyed approximately 1,000 villages and provides precise measures of schooling, self reported disease, and other household-level outcomes for a random sample of residents. Importantly, the survey was carried out before large-scale malaria interventions took place, starting in 2007. Thus, indoor residual spraying (IRS) and treated or untreated mosquito nets were quite uncommon at the time.[1] With this data, I find that the interaction between elevation and slope captures malaria rates quite well. That is because malaria is sensitive to elevation differences in places (flatlands) where the habitat is suitable for mosquito breeding, and is not sensitive to elevation in unsuitable areas (steep lands). In addition, the variation in topographic characteristics is large enough that it is possible to estimate differences in malaria incidence across villages located within the same province (*wareda*), the smallest administrative unit in Ethiopia after the village. Since most administrative units, local markets, and school administration are centralized at the provincial level, the within-estimator eliminates a large source of heterogeneity in health, education, and other unobserved variation in weather, disease patterns, agricultural practices, and cultural traditions. Using topographical features as instruments for the variation in malaria, I find evidence that village outcomes are correlated with malaria endemicity. In particular, increasing the malaria rate by 5.6 percentage points (which in the data corresponds to moving from a village with no malaria to one with average malaria) is associated with 0.36 fewer years of schooling for children and 0.25 for adults. I also study other village outcomes—such as child labour, food insecurity or wealth accumulation—and find evidence that some but not all outcomes are significantly related to malaria endemicity. As a further benchmark of these results, I re-estimate regressions using a well established climate-based predictor of malaria. I find that topography is a better predictor of malaria incidence, and that climate-based estimates of correlations are not in disagreement with the topography-based estimates.

The correlations found in this paper may not be interpreted causally. A causal interpretation of my findings requires topography to be a driver of long run malaria risk, and be uncorrelated with other possible drivers of local development. It is likely that elevation and slope are correlated with many factors other than disease, including agricultural production, household's wealth, and the returns from education. On the other hand, it is less clear whether the *interaction* between elevation and slope should be correlated with those same factors, after controlling for the direct effect. To gain a better sense of which factors might be important confounders, I also explicitly study the relationship between topography and a number of village variables, including exposure to (non-health) shocks, distance to facilities like schools, and measures of population pressure. I find some evidence that some factors—like droughts and proxies for population pressure—also vary with topography. To the extent that these confounding factors are observable and can be controlled for, I find that they have a negligible impact on my estimates.

Several papers study the effect of malaria on schooling and economic development, but for the most part these papers do not use African data.[2] Barreca (2010) used variation in rainfall and temperature in the 1920s American South and estimates that a standard deviation increase in exposure to in-utero and postnatal malaria reduced education by 0.23 years. Bleakley (2010) and Lucas (2010) use the timing of malaria eradication in the US, Latin America and Sri Lanka to estimate the effect of malaria on education and earnings, finding that education rates increased post eradication. Using a similar methodology, Percoco (2013) finds similar results for Italy, while Cutler et al (2010) do not find improvements in schooling but some modest increases in adult earnings. These estimates identify the effect of childhood exposure of malaria prior to successful eradication campaigns. As such, they capture partial equilibrium effects and exclude those general equilibrium effects that are captured by time invariant instruments such as ones considered in this paper. In addition, there are important differences in the disease environment between Africa and the rest of the world, and it is in light of such differences that it becomes important to find estimates specific to the African continent. Malaria eradication was attempted but did not succeed anywhere in Africa (Webb, 2009). The primary reason given for this is that the disease ecology/environment in Africa is unusually complex relative to the other continents, making eradication difficult and ensuring that the effects of the disease are more pronounced than elsewhere. Successful eradication in Africa would require dealing with all four types of malaria parasite (*P. Falciparum*, *Ovalae*, *Vivax*, *Malariae*) and dozens of mosquito species, some of which are considered extraordinarily effective in transmitting disease and avoiding the standard methods of mosquito control interventions (D'Antonio and Spielman, 2001). By comparison, in places where large-scale, broadly successful eradication or control took place—Southern Europe, the American South, and parts of Latin America—malaria was less entrenched and the mosquito vectors easier to be dispensed with. In addition, the problem of malaria is perceived to be worst on that continent, primarily due to the incidence of *P. Falciparum*, a vector responsible for most malaria mortality. In fact, an estimated 90 per cent of malaria induced child deaths occur in Africa (Cook and Zumla, 2008).

The rest of the paper is structured as follows: the next section provides background information on Ethiopia and malaria. Section 3 presents the data. Section 4 discusses the methodology. Section 5 shows the relationship between the disease environment (and malaria in particular) and topography. Section 6 reports the development correlates of the malaria environment. Section 7 compares my estimates with those derived from climate-based instruments. The conclusion follows.

2. Background

Malaria

Malaria affects approximately 250 million people and is responsible for one million deaths annually (WHO, 2009). This disease is transmitted to humans through bites of female *anopheles* mosquitoes. Once infected, a person can develop chills, very high fevers, anemia, and—especially in children less than five years of age—brain damage, coma, and death. The seriousness of symptoms depends, to a great extent, on the degree of parasitaemia, itself a function of the number of infected bites suffered by the patient (the inoculation rate). In general, the higher the inoculation rate, the higher the chance of severe symptoms or permanent damage.

In Ethiopia, the estimated incidence rate for malaria (that is, the estimated probability of contracting the disease in a year) was 15 per cent in 2005, which is somewhat low relative to the rest of sub-Saharan Africa (where the average incidence rate is 0.33), but higher than any other country outside of sub-Saharan Africa bar Panama, Laos, Myanmar, and the Solomon Islands (Korenromp, 2005). Despite the somewhat low incidence rate, this country is an appealing place to do a study on malaria for at least two reasons. First, malaria is still a very important public health problem: Ethiopia is thought to experience some 10 million cases per year, the fourth highest case number in sub-Saharan Africa (behind Nigeria, the DRC, Tanzania, and Uganda [Korenromp, 2005]). According to an Ethiopian report, clinical malaria accounts for 10 - 40 per cent of all outpatient consultations, 13 - 26 per cent of inpatient admissions at various health facilities, and is responsible for 15-17 per cent of case fatalities in health facilities (WHO, 2006).

A second reason for considering malaria in Ethiopia is that, unlike most other African countries, there is extensive local variation in malaria incidence. Figure 1 and 2 make this clear. The figures show the predicted malaria presence across the continent, according to the MARA model (MARA, 2011) [3] For most countries, malaria incidence has extremely high spatial autocorrelation. Ethiopia, on the other hand, has a very low degree of spatial autocorrelation; malaria in Ethiopia is a localized disease. Studies indicate the presence of spatial autocorrelation in malaria rates for villages located within five to 10 kilometers, dropping rapidly after that (Yeshiwondim et al, 2009). This feature permits the estimation of the impact of malaria incidence by comparing villages in close proximity to one another.

Transmission: slope and temperature

Malaria transmission is affected by land slope and temperature. The effect of temperature is well recognized in the malaria literature: increases in temperature increase the survival rate of mosquito larvae and reduce the length of sporogony—the time it takes for a mosquito which has ingested infected blood to be able to transmit the disease (Cook and Zumla, 2008), thus speeding up transmission. Below a certain temperature threshold (corresponding to approximately 18 degrees Celsius), sporogony is sufficiently slow that it cannot be completed before the mosquito dies, making malaria transmission impossible (Craig, Snow and le Sueur, 1999).

In Ethiopia, air temperature is largely determined by elevation. Figure 3 shows the relationship between elevation and the average daily temperature in the warmest month for Ethiopia. Above 2,500 meters of elevation, temperatures are consistently below the no-malaria threshold, making villages located there malaria-free.[4]

A second determinant of malaria incidence is the presence of mosquito breeding grounds. *Anopheles* mosquito larvae need stagnant water pools to reproduce, and environments with scarcity of bodies of water sustain smaller mosquito populations. On average, villages that are very sloped are less likely to have stagnant water pools, as rainfall converges to river systems downstream. In addition, larvae developing in water pools in sloped areas are more likely to be washed away during downpours, and might be more likely to be targeted and eliminated by drainage activities carried out by local villagers. Thus, sloped areas make for poor mosquito breeding ground, reducing the threat of malaria transmission, especially in the higher elevation areas that are the focus of the present study (Balls et al, 2004).

Since villages that have a high slope gradient are inhospitable to mosquitoes, the incidence rate among villages located in steep lands is low, regardless whether they are located in high elevation or low elevation areas. On the other hand, villages located in flatlands are more likely to have a suitable environment for mosquito breeding; with mosquito populations not being a limiting factor of local disease transmission, elevation is likely to play an important role in malaria transmission, with high elevation villages having significantly lower incidence rates than low elevation villages. In other words, the *sensitivity* of malaria incidence in regards to elevation should increase the flatter the local topography.

To be sure, malaria is not the only vector borne tropical disease that might be influenced by temperature and terrain. Other candidate diseases, however, follow different patterns of transmission. Dengue fever (transmitted by mosquitoes) is generally concentrated in urban areas and is found sporadically in Ethiopia (Aseffa, 1993); Yellow fever and Chikungunya (also transmitted by mosquitoes) are rare in the country;[5] Sleeping sickness (transmitted by tsetse flies) is also found sporadically (Fevre et al, 2006) as is Rift Valley fever (which, in any case, has severe morbidity and mortality rates in less than 1 per cent of patients [WHO, 2010]). Ethiopia is considered at a high risk of meningitis epidemics (Cuevas et al, 2007), and while there is some association between environmental factors (rainfall, forest cover, dust levels, human population density) and the spread of the disease, the relationship between meningitis and either elevation or slope is not understood.[6] Other diseases, like diarrheal diseases or influenza, might be affected by land slope or temperature, but they are unlikely to be affected by the interaction of the two.

3. Data

The Welfare Monitoring Survey (WMS) was conducted by the Ethiopian Statistical Agency between June 24 and July 3 of 2004, and it involved over 2,000 villages across all states in the Ethiopian Federation. It covers basic individual and household characteristics, as well as access to several services. Crucially, it includes information on both schooling and individual level health conditions (including self-reported malaria spells) in the preceding two months. These health outcomes refer to the months May and June, corresponding to the end of the first malaria season (which generally runs from mid-March to June).[7] In particular, for each member of the household, the survey asks whether the individual faced a health problem in the prior two months and, if so, what was the reason for this sickness. The questionnaire provides six mutually exclusive reasons, including malaria. Using this information, I constructed a measure of self-reported recent malaria incidence at a village level as the fraction of the surveyed population who was reported as suffering from malaria.[8] Since this measure of malaria incidence comes from self-reported diagnoses, they are likely to be measured with considerable error. This will justify the instrumental variable approach taken in this paper.

Since it is unlikely that measurement error is as problematic if village *topography* is used in place of self reported village malaria, the WMS was integrated with measures of village topography derived from the use of two additional datasets. The first database is an electronic map developed by the Ethiopian Development Research Institute (EDRI), in collaboration with the International Food Programme Research Institute (IFPRI). Using the EDRI/IFPRI electronic maps, I matched the names of the villages in the electronic map with the names of the WMS villages for the regions of Ahmara, Oromia and SNNP (see figure 4). I was unable to match a number of villages within these regions. In total, around 35 per cent of villages in the three regions were dropped. Dropped villages appear more rural and a little poorer and less educated, but have similar levels of sickness and self reported malaria.[9]

Elevation, temperature, rainfall and slope for the entire country of Ethiopia come from remote sensing data collected by NASA satellites and elaborated by scientists at the Livermore National Laboratory in California. This data shows, for any coordinate point in Ethiopia, its elevation, slope, average rainfall, and minimum, maximum and average daily temperatures for each month of the year. I transferred the remote sensing data to the EDRI/IFPRI map by averaging elevation, temperature and slope over the entire surface of each village.[10]

Table 1 provides summary statistics of the study villages, and separately for villages located below the 2,500 meter threshold. 84 per cent of villages are located between 600 and 2,500 meters above sea level, with the remainder being high altitude villages located up to 3,500 meters in elevation. Average village slope is 6.7 per cent, meaning that there is an average 6.7-meter gain per 100 meters.

Around a quarter of individuals were sick in the prior two months. A quarter of sick respondents reported suffering from malaria, half did not report a specific disease, and the rest were evenly divided among other health problems (diarrhea, tuberculosis, ear, nose and throat problems, other injuries). As one would expect from the discussion, the rate of reported sickness is significantly lower in high altitude villages: only 20 per cent of individuals there reported some sickness, and essentially none reported suffering from malaria. The dependent variables of the paper (years of schooling for children aged between 7 and 19 and adults and other development outcomes) are presented next. Since many children are still enrolled, average schooling for children does not represent completed education, while for adults I have largely completed formal education. On average, both the children sample and the adult sample report having one year of schooling. Other outcomes are reported next. An index of asset accumulation, obtained from the principal component analysis of the assets available in the questionnaire, indicates that households interviewed are generally quite poor and report few assets. Child labour and food insecurity are also quite prevalent: On average, 53 per cent of children reported working, and families report 0.8 instances of hunger in the prior five years. The rest of the table describes some of the control variables used in the study.

4. Methodology

The objective of this paper is to explore the correlates between local malaria incidence and local development outcomes by exploiting the correlation between malaria incidence and topography. The strategy is to first demonstrate how self-reported village malaria (measured as the fraction of individuals who reported having malaria in the previous two months) is explained by village elevation, village slope, and the interaction between the two. I then use an instrumental variable approach to explore the correlation between malaria rates and individual development outcomes, including schooling, asset accumulation, and child labour. Formally, I model village malaria incidence in village *v* in the following regression:

$$malaria_v = \alpha_v + \beta_0 Elevation_v + \sum_{j=2,\dots,5}\left[\beta_1^j + \beta_2^j \times Elevation_v\right] \times Slope_v^j + \eta X_v + \varepsilon_v \quad (1)$$

$Elevation_v$ indicates village height (in hundreds of meters), $Slope_v^j$ is a dummy indicating villages in the *j*-th quintile of slope,[11] α_v is a province fixed effect, and X_v is a set of village-level controls. In addition to malaria, I use the same regressors to show existing correlations between topography and other health outcomes, including other sickness and mortality. β_0 estimates the correlation between elevation and the outcome variable in areas that are very flat; β_2^j measures the added correlation of elevation and the outcome in villages located in slope *j;* while β_1^j measures the direct effect of slope *j.* Alternatively, the effect $\beta_0 + \beta_2^j$ measures the correlation between elevation and the health outcome variable for those villages located in slope quintile j. Based on this model, malaria incidence should be negatively related to elevation in flatlands as well as slope (that is, β_0 and $\beta_1^j < 0$). Moreover, the steeper the area, the less sensitive malaria incidence is to increases in elevation. Thus the overall correlation between elevation and health outcomes should become smaller and smaller, and $\beta_0 + \beta_2^j \to 0$ as j increases. This implies that the interaction term β_2^j should be positive, and $\beta_2^{j+1} > \beta_2^j > 0$.

The above analysis and interpretation is applied to a subsample comprising of villages with some positive risk to malaria--that is, places located below 2,500 meters. As an additional robustness test, I will also model malaria incidence with an augmented regression that includes the full sample--thus including villages above 2,500 meters. A dummy variable *above 2,500m* is added and fully interacted with both elevation and slope quintiles such that each β_0, β_1^j, and β_2^j is identified separately for villages located above and below the 2,500 meter threshold.

Having established the relationship between topographical features and the local malaria environment, I use topographical instruments in a 2SLS regression where the second stage is:

$$y_{iv} = \alpha_v + \gamma_1 \widehat{malaria}_v + \delta_0 Elevation_v + \sum_{j=2,\ldots,5} \delta_1^j Slope_v^j + \xi X_{iv} + \mu_{iv}, \qquad (2)$$

where y_{iv} is an outcome of interest for person i in village v such as years of schooling, average asset ownership, and exposure to shocks. Note that the second stage includes elevation and slope quintiles, but excludes the interaction term. All specifications include province fixed effects α_v, which absorb the unobserved variation within provinces, and a set of control variables X_{iv}. In regressions involving children's outcomes, controls included are average village rainfall and child age, education of household head, education of spouse, ownership of oxen and livestock, household asset ownership, distance to primary schools and health centres, whether child lives in a female headed household, plot size of agricultural household, and drought prevalence. The control set X_{iv} for regressions on adults include age, distance to facilities, land sizes and drought prevalence.[12] Finally, since sickness and education are likely correlated at the village and province level, I cluster errors at the province level.

The parameter of interest is γ_1, which measures the correlation between malaria incidence and outcome *y*. Note that entering *malaria* directly in regression 2 is problematic for a number of reasons. To begin with, self reported malaria in a short period of time is an imperfect measure of the expected local incidence; that is, my measure of malaria is likely to suffer from substantial classical measurement error. A second problem is reverse causality: villages with higher investments in education or higher levels of socioeconomic outcomes might be more likely to invest in malaria prevention schemes and lower malaria rates.[13] These elements will bias coefficients toward the null. The advantage of the IV estimate is that that it addresses the measurement error and reverse causality problems. To the extent that the identifying assumption and the exclusion restriction hold, the coefficient γ_1 identifies the causal effect of malaria on outcome y. Controlling for the main effect of elevation, the instruments assigns to malaria the differential response of the outcome variable to elevation across slopes.

Given this, consider the meaning of the identifying assumption and exclusion restriction. Since education decisions are taken in the past, the first requires that *present* malaria incidence, as predicted by elevation and slope, is correlated with unobserved *past* malaria risk, as predicted by elevation and slope. This assumption is violated if malaria is a very recent phenomenon, or if malaria transmission is sufficiently unstable such that topography generally does not predict malaria risk. In Ethiopia, the association between the natural environment and malaria risk has been present for hundreds if not thousands of years. In his book on the history of malaria in Ethiopia, James McCann (2015) describes malaria presence in the Blue Nile region: "malaria has long lived there" (p. 63). Travellers as far back as the 1800s contrasted "the wholesomeness of the "Abyssinian Alps" that looked… "salubrious" compared to the feverish lowlands." (p. 15). Highlighting the fact that incidence may be unstable but the *risk* stable, McCann states, "The patterns are predictably periodic, episodic, and maddingly erratic." (p.58). Thus, the historical record indicates significant variation in the severity of malaria over time, it also suggests that the predictive power of elevation (and possibly slope) was also true in the past.

The exclusion restriction requires that the interaction between elevation and slope does not affect the outcome variable through some channel other than malaria. There are many omitted variables that are likely to affect development outcomes and that are correlated with elevation and slope, including land productivity, crop choice, agricultural shocks, distances to transportation hubs, population density, and (more broadly) investments in public goods. The exclusion restriction requires that these omitted variables are controlled directly or indirectly by the covariate matrix

X, the district fixed effects, and the direct measure of elevation and slope. It thus assumes away the situation where these covariates have a high correlation with elevation in flatlands, but correlate less and less with elevation at higher slopes. The empirical section tests for the presence of additional confounders. Unfortunately, due to data limitations it is not possible to test every possible covariate, and an omitted variable problem remains a possibility. A second problem with the IV strategy is that the instruments will be somewhat weak. This is likely due to the fact that malaria incidence was measured at the end of the malaria season, and thus the first stage does not identify the topographical effects as precisely as one wishes. It is a well-known fact that IV estimated with weak instruments lead to estimates that are biased towards the OLS (Angrist and Pischke, 2009). In all cases provided, because of endogeneity and measurement error, these OLS estimates are close to zero while the IV estimates move the estimates in the expected direction and provide estimates that are larger in magnitude and significance. I also report IV estimates from a limited maximum likelihood estimation (LIML), which reduces any biases originating from having many instruments or weak instruments. To the extent that LIML estimates are similar to the IV estimates in magnitude and significance, the IV bias is likely to be small (Angrist and Pischke, 2008).

Omitted variables and weak instruments notwithstanding, the IV estimate γ_1 corrects measurement error and endogeneity biases in the correlation between malaria incidence and development outcomes. It moves OLS estimates towards the true causal estimates, and provides at least a lower bound of the true estimate.

5. Local conditions and topography

The disease environment

To show how topography influences health outcomes through malaria, figure 5 shows a local linear estimate of the relationship between self reported malaria incidence (as described in section 3) and village elevation. As altitudes increase, the proportion of the villagers reporting having suffered from malaria decreases. The proportion of villages with malaria remains flat and close to zero for elevations above 2,500 meters, as expected. Using the same procedure, the relationship between slope and malaria for villages located below 2,500 meters is shown in figure 6. Malaria rates are high in flatlands, with the estimated rate falling with the slope.

To formalize the results from the two figures, table 2 provides an initial overview of the relationship between health and topography in Ethiopia through a regression of village disease incidence on village elevation, slope, and a full set of covariates.[14] The odd columns of the table report coefficient estimates for villages located below 2,500 meters of elevation. Even columns also include villages above the threshold; the column reports the main estimated effect for villages below the threshold as well as the estimate from an "*above 2,500 meters*" dummy interacted with elevation.

Columns 1 and 2 consider overall health incidents. It confirms that higher elevation villages are healthier: on average, each 100 meters of elevation gain reduces sickness by 0.8 percentage points, or approximately 3.3 per cent of the average level of village sickness. Moreover, this correlation disappears above the 2,500-meter line (column 2), something that would be expected if the correlation were driven by malaria. Land slope is negatively correlated with sickness, but the coefficient is insignificant. In columns 3 and 4, I consider self-reported malaria. The coefficient on altitude indicates that just 100 meters gain in elevation is sufficient to reduce malaria by 0.6 percentage points, which is 10.5 per cent of mean malaria---a very high rate. Slope is now strongly significant, indicating that more sloped villages have lower reporting of malaria. One way to consider the magnitude of the effect is to consider that moving from an area with zero slope to one with average slope (6.7%) reduces malaria incidence by 2 percentage points-- over a third of mean malaria. Above 2,500 meters (column 4) the negative correlation between malaria and elevation disappears. Columns 5 and 6 show the relationship between other types of sickness (all causes excluding malaria) and elevation and slope. The estimated effects are small and statistically insignificant. This assures us that the linkage between health and elevation is driven by malaria.[15] Finally, the last two columns of the table show results for the measure of mortality present in the survey (number of deaths in the household in the past five years). This measure is noisy, as it includes all mortality for all age groups and for any reason. While the estimates move in the same direction as those found in the malaria regressions, the estimated coefficients are all insignificant.

Estimates of equation (1) are presented in table 3. The table shows whether the interaction between elevation and slope can be used to predict adverse health reporting. Column 1 shows that this is not the case for overall sickness. While sickness decreases with elevation, this relationship is found across slopes. When focusing on malaria only (column 2), I find that elevation predicts malaria in villages located in the bottom four quintiles of land slope---but not in the steepest villages, which have low incidence rates everywhere. The point estimates suggest that the impact of elevation on malaria declines the steeper the village. Moreover, the interaction coefficients are jointly significant: the p-value on the F-statistic of joint significance is 0.07. Column 3 looks at all-cause sickness excluding malaria. I find that sickness and elevation are not related in flatlands, but overall other cause health problems decline at higher slopes. Taken individually, diseases such as diarrhoea, tuberculosis, ear, nose and throat problems, and other types of injuries are sometimes correlated with elevation at higher slopes only (not shown). Thus, the pattern of high correlation with elevation in flatlands is specific to malaria. Finally, mortality (column 4) again follows a clear pattern consistent with malaria, but the coefficients are not statistically significant.

Other factors and topography

A key issue is that topography could be correlated with a host of other factors affecting local development other than malaria. For instance, elevation alone could affect farm productivity and exposure to weather shocks. Higher elevation villages might also have closer access to schools, services, or labour markets where returns to education are higher. In addition, the highlands in Ethiopia are known to be densely populated, and this could affect the level of wealth or schooling in the community. To the extent that these factors are correlated with the interaction between elevation and slope, they represent alternative pathways through which topography shapes local development.

Table 4 explores a number of these potential correlates[16]. First, column 1 through 4 report estimates from the regression (1) on a number of different types of agricultural shocks--namely, the number of instances the household suffered floods, droughts, livestock losses, a price shock or some other (non-health) shock in the prior five years. Droughts are indeed correlated with the interaction between slope and elevation, with the incidence of droughts declining with elevation in flatlands but not in steep lands. Shocks to livestock are significantly correlated with elevation in quintile 2 and 5 of slope, while floods and other shocks cannot be explained by topographic factors. Next, I make use of household-specific distances to schools and clinics to test whether location to public services differs by elevation and slope (columns 5 and 6).[17] School distance does not decrease with elevation in low-sloped areas, but it does for highly sloped areas--the opposite pattern observed for disease. Third, I would like to check how topography correlates with population density. As a noisy proxy of population pressure, I consider in column 7 the average size of land holdings. This measure seems to vary somewhat with topography; in particular, land sizes decrease with elevation, especially in the steepest areas.

Overall, the table demonstrates the presence of some topography covariates that may be important co-determinants of development outcomes. Most of the positive evidence in the table points to an elevation gradient in steep lands; much less so for flatlands. In all regressions, I control for land sizes and droughts, even though the estimates are not sensitive to their inclusion.

6. Malaria environment and individual development outcomes

Schooling

Reduced form results

Having established that recently reported malaria is strongly correlated to topography through the interaction between village elevation and slope, I now show how topography correlates with other development outcomes in a similar way. In table 5, I turn to the sample of individuals and regress years of schooling for children (age 7-19) and adults (age 20+) on elevation, slope, and individual covariates, first focusing on elevations below the 2,500-meter threshold (columns 1 and 2) and then on all villages (columns 3 and 4).[18] Average schooling increases with elevation for both adults and children. Conversely, the higher the slope of the village, the lower the educational attainment. In columns 5 and 6 I interact elevation with slope quintiles. Both regressions show that there are gains in education from elevation in flat areas, but the higher the slope, the lower the gain; in addition, there are no gains for those living in the steepest areas. Coefficients are more precisely estimated for the sample of children.

Instrumental variable results

I turn to instrumental variable estimates of equation (2) in table 6. Here, I use village topographical characteristics as instruments for malaria incidence, and look directly at the impact of recent self-reported village malaria on average years of schooling for children and adults, with results disaggregated by gender. In the first column, I do not focus on topography and report simple OLS results. In column 2, I use a 2SLS model where malaria is instrumented by elevation and the slope/elevation instruments are not included in either first or second stage. In column 3 I report IV estimates using the first stage reported from equation (1), which controls for elevation and slope quintiles and uses the interaction between elevation and slope as instruments. In column 4, I include all observations from all elevations, where the instrument set includes the interaction between quintiles of slope, elevation, and whether a village is located above 2,500 meters. In column 5, I report LIML estimates.

Looking at the first column of table 6, the coefficients on malaria are small and statistically insignificant for all OLS regressions, consistent with significant measurement error in the malaria measure. When turning to the IV specification in column 2, however, the coefficients on malaria become all negative, large in magnitude, and statistically significant for children, and marginally insignificant for adults (p-value for the all-adult regression: 0.13). Coefficients for women are slightly more negative but within the standard error of those for men for both adults and children. Column 3 uses the interaction of elevation and slope as instruments. Instrumented malaria remains negatively associated with lower schooling levels, but the results are not statistically significant for children education. Column 4 reports the results using the expanded set of instruments and villages. The coefficients estimates remain very similar to the previous column but turn statistically significant for children, while they become smaller but remain significant for adults (although only for the male adult sample). Point estimates are more negative for males than females on both child and adult samples, which is consistent with epidemiological evidence showing that Ethiopian males are more exposed to the disease (Yeshiwondim et al, 2009).[19] LIML estimates (column 5) are very similar to IV estimates. This suggests that our weak instruments (P-value of F-test is 3.4) are identifying the correlation of interest. Overall, considering the coefficients from panel A and D, the regressions imply a difference in education between no malaria villages and average malaria villages (where the average is 5.7 per cent for all villages) of 0.36 fewer years of school for children, and 0.25 fewer years for adults.

Other outcomes

Table 7 considers the effect of topography on a number of other outcomes that might be affected by malaria incidence. I start with child labour, measured as the proportion of children aged 10-17 who reported working in the prior seven days, in column 1 through 4. One could expect child labour to be more prevalent in areas with more malaria if malaria drives up the demand for child labour, possibly because children may be sent to work in the fields to substitute for sick parents. Second, frequent malaria spells might reduce the earning ability of a household. For agricultural households, this might translate into more food insecurity in the short run. To verify this, in column 5-8 I consider self reported instances of food insecurity (measured as the number of times villagers reported suffering a food shortage in the prior five years). Finally, over the long run, a more adverse environment should translate into lower levels of asset accumulation. To capture this, I consider the average value of the asset index in the remaining columns. For all the outcome variables, I first report the (biased) OLS estimates, and then consider progressively less biased 2SLS specifications from the last three columns of table 6.

The IV estimates of child labour effects (columns 3 and 4) are significant at the 10 per cent level and indicate that there is 0.057 X 1.685=9.6 per cent more child labour in areas with average malaria than in no malaria areas. The estimated correlation of malaria and food insecurity is not as precisely estimated, as one would have liked. However, it is notable that the estimates become larger and larger in magnitude as we move from the more biased OLS estimates to the less biased LIML estimates. The IV estimates are statistically insignificant, so a relevant conclusion cannot be drawn, although the coefficients move in the expected direction. The LIML estimates implies 0.057 X 3.15=0.18 more instances of food shortage periods. Finally, the estimate on the asset index regression is small and has a very large standard error, so no positive conclusion can be drawn.

7. Climate vs. topographical predictors

Since the justification for the use of topography is that transmission of the malaria parasite declines with temperature, a natural alternative instrument is the village climate. One strategy is to replace elevation with average village temperature. This strategy yields similar results with somewhat weaker instruments.[20] An alternative approach is to use the malaria index developed by Craig, Snow and le Sueur (1999); this index is the basis for the widely used maps of malaria transmission across sub-Saharan Africa created under the MARA/ARMA project (see figures 1 and 2). The climate index assigns a value of 0 to areas that are not suitable to malaria transmission, a value of 1 to areas that are suitable for transmission, and a value between 0 and 1 for areas that are partially suitable. As the discussion below will make clear, this measure captures well malaria *presence*, but it is a poor proxy for malaria *incidence*.

To show this, I first constructed the MARA index for the villages in the sample.[21] The MARA index is highly correlated with elevation (correlation coefficient -0.84); the merits of the MARA index relative to an elevation-only method can be seen in table 8. The first column shows estimates from regressions of recent village malaria on the MARA index on all villages. The index strongly predicts village malaria. In the second column, I set up a "horse race" between the topographical and climate instrument by including elevation interacted with the above 2,500 meter dummy. When topography is included, the MARA index ceases to be significant. As a measure of malaria *intensity,* the MARA model is not as useful. Columns 3 and 4 report the same regressions, where the dependent variable is now an indicator variable equal to one if at least one person in a village reported having malaria. As can be readily seen, the MARA model strongly predicts malaria presence, and it wins the horse race against elevation. Since that the focus of the index is in measuring the suitability of transmission rather than the intensity of transmission, it is unsurprisingly much better at predicting malaria presence rather than incidence.

Using the MARA model as the IV for malaria, the main results of the paper are replicated in table 9. Estimated malaria effects on schooling are somewhat more negative than those in table 6 (columns 1 and 2); however, standard errors are large and the instrument set appears overall weaker. For comparison, in column 3 and 4, I replace malaria intensity with malaria presence. The F-statistic on the instrument becomes large, and a (negative) correlation is now evident among children (but not adults). Estimates do not survive more sophisticated estimation methods (such as fuzzy-RD designs).

8. Conclusion

This paper presents correlates of malaria and development outcomes including schooling, mortality and asset accumulation using data from Ethiopia. The paper first provides evidence that malaria is strongly associated with topographical features. In particular, it makes use of the fact that malaria prevalence is sensitive to air temperature and to the presence of mosquito breeding grounds. Since temperature declines with elevation, and mosquito breeding grounds are less likely to form in sloped areas, it follows that elevation, slope, and the interaction between elevation and slope are potential time invariant predictors of malaria risk. The paper then uses these topographical features as instruments for malaria to reduce measurement error and reverse causality in a regression of malaria on local outcome variables from a large-scale survey of Ethiopian villages. I find that Ethiopian villages that are less prone to malaria have higher schooling rates for both adults and children. I also find some suggestive (but ultimately inconclusive) evidence that malaria incidence may be an important driver of food security.

The estimates presented in this paper may be considered a lower bound on the (local) effect of malaria on long run outcomes. Because the methodology adopted here relies on topographical differences across villages, one should be careful to compare them from estimates obtained through other methodologies, such as the ones adopted by Bleakley (2010), Lucas (2010). In particular, estimates from papers that rely on eradication identify the effects of eradication on a specific age group (that is, children exposed in utero or in early life), and by necessity capture partial equilibrium effects. To the extent that the exclusion restriction in this paper applies, my estimates are capturing an effect that incorporates both partial and general equilibrium. While this approach has a particular limitation in that the exclusion restriction is difficult to verify, it may be the best one can do in contexts where there are no exogenous changes in malaria prevalence induced by eradication or other malaria control measures.

---

[1] For instance, the Ethiopia National Malaria Indicator Survey of 2007 states that ITN coverage was 3.5 per cent in 2005, increasing to 53.3 per cent by 2007.

[2] Depetris-Chauvin and Weil (2013) study the effect of malaria on early African development (mortality and economic development) by exploiting variation in the sickle cell trait across African ethnic groups. Barofsky, Anekwe and Chase (2015) study a failed eradication campaign in Uganda and find that cohorts born in the eradication period increased their schooling by 0.5 years.

[3] The Mapping the Malaria Risk in Africa collaboration integrates rainfall and temperature information to generate a model of malaria prevalence for sub-Saharan Africa.
[4] In this paper, I focus on elevation as opposed to temperature because the entomological literature is unclear about how to take into consideration daily and seasonal variations in temperature. Nonetheless, given the strong linear relationship between the two measures, all results in this paper are very similar when replacing elevation with average daily temperature, with results available upon request.
[5] Aseffa [1993] reports no cases in Ethiopia from 1966, and the WHO in 1995 reported no cases between 2000 and 2004. In 2004, there were a total of 128 suspected yellow fever cases in sub-Saharan Africa.
[6] Meningococcal bacteria is spread person to person through cough droplets, so there is not a direct connection with topographical factors. The most recent epidemic prior to the study period was in 2001-2002, where a total of 1,332 cases and 85 fatalities were identified (WHO, 2002).
[7] The second malaria season in Ethiopia is September to November, after the long rains (*meher*). The timing of rains varies from region to region.
[8] Malaria in the highlands fluctuates from year to year and seasonally. Fortunately, the survey was taken between June 24 and July 3, a period that generally coincides with the peak of the minor transmission season. Thus the survey covers a period of malaria transmission. See online appendix for a description of malaria transmission patterns in Ethiopia.
[9] See the online appendix for detailed information on the matching procedure and comparison between matched and unmatched villages.
[10] Note that the average of the topographic measures over the entire surface of the administrative unit is, by necessity, an imperfect measure of the true topographic characteristic of the village. This is because housing units are not uniformly spread around the administrative unit, and are in fact often concentrated in a smaller area. See data appendix table for a greater description.
[11] The first quintile of slope goes from 0 to 2.5 per cent; the second quintile from 2.5 to 4.5 per cent; the third from4.5 per cent to 6.8 per cent; the fourth from 6.8 per cent to 10.2 per cent, and the last from 10.2 per cent to 26 per cent.
[12] While some of the controls in X_{iv}could represent possible pathways for malaria to affect outcomes, in practice their inclusion is not important—estimates of γ_1 obtained from regressions excluding all household controls are slightly noisier but similar in magnitude (see tables in online appendix).
[13] The main types of investments available are (relatively expensive) environmental management to reduce the productivity of mosquito breeding sites. At the time of the survey, mosquito nets were quite rare in rural Ethiopia.
[14] For simplicity, tables 2, 3 and 4 employ the child regression controls averaged at the village level. Alternative specifications (without controls, or using the individual as unit of observation) lead to very similar estimates and are available in the online appendix.
[15]Note that, due to the structure of the questionnaire, there should be a negative correlation between reporting of malaria and reporting of other sickness types at the individual level (because answers are mutually exclusive). If respondents report other sickness ahead of malaria, there could be an attenuation bias in the malaria estimates. Since the fraction of malaria or other sickness reports is generally small, the bias is unlikely to be large.
[16] Table shows village-level regressions. Regressions at the individual level are very similar and are available in the online appendix.
[17] School and clinic distance is a convenient proxy to distance to other services, since distances to other services (like markets, roads, sources of farm goods, and so on) behave in a similar pattern.
[18] Results are robust to changes in the covariates. While coefficients on covariates are not shown, they move in the expected direction: parental wealth and education increase children's schooling and distance to schools decreases it. Livestock ownership is associated with more schooling, suggesting that the income effect from livestock trump substitution effects arising from the fact that animal husbandry is an important labor activity of Ethiopian children. Graphs of local linear estimates of schooling and topography available in the online appendix.

[19] Boys and adult men spend more time in the fields than girls, who spend a larger fraction of their time at home. Since sleeping outside to protect the field is quite common, farm workers are more exposed to mosquito bites.

[20] Results available upon request.

[21] Due to slight differences in data availability, my MARA measures may be marginally different from those observed in figure 1. See online appendix for a discussion of how the MARA index was constructed for the WMS villages in my sample.

Bibliography

Angrist, J.D., & Pischke, J.S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Aseffa, A. (1993). Viral Diseases in Ethiopia: A Review. *East African Medical Journal*, 70(10): 624-625.

Balls MJ, Bødker R, Thomas CJ, Kisinza W, Msangeni HA, & Lindsay SW. (2004). Effect of topography on the risk of malaria infection in the Usambara Mountains, Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 98(7), 400-408.

Barofsky, J., Anekwe, T., & Chase, C. (2014). Malaria Eradication and Economic Outcomes in Sub-Saharan Africa: Evidence from Uganda. Mimeo.

Barreca, A. (2010). The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria. *Journal of Human Resources*, 45(4), 865-892.

Bleakley. H. (2010). Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics,* 2(2), 1-45.

Cook, G.C., & Zumla, A.I. (2008). *Manson's Tropical Diseases*, 22nd Edition.

Craig, H. Snow, R.W., & le Sueur, D. (1999). A Climate-based Distribution Model of Malaria Transmission in Sub-Saharan Africa. *Parasitology Today*, 15(3), 105-111.

Cuevas, L.E., Jeanne, I., Molesworth, A., Bell, M., Savory, E.C., Connor S.J. & Thomson M.C. (2007). Risk Mapping and Early Warning Systems for the Control of Meningitis in Africa. *Vaccine*, 25(S), A12-A17.

Cutler, D., Fung, W., Kremer, M. Singhal, M. & Vogl, T. (2010). Early-life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India. *American Economic Journal*, 2(2), 72-94.

Daily, J. (1998). Malaria. In S. L. Katz, A.A. Gershon, & P.J. Hotez (Eds), *Krugman's Infectious Diseases of Children*, 10^{th} Edition.

Depetris-Chauvin, E. & Weil, D.N. (2013). Malaria and Early African Development: Evidence from the Sickle-Cell Trait. NBER Working paper 19603.

Ethiopian Ministry of Health. *National five-year strategic plan for malaria prevention and control in Ethiopia, 2006-2010*. Addis Ababa: Federal Democratic Republic of Ethiopia, Ministry of Health.

Fevre, E.M., Picozzi, K., Jannin, J. Welburn, S.C. & Maulin, I. (2006). Human African Trypanosomiasis: Epidemiology and Control. *Advances in Parasitology*, 61, 167-221.

Gallup, J. L. & Sachs, J. D. (2001). The Economic Burden of Malaria, *American Journal of Tropical Medicine and Hygiene,* 64((1,2)S), 85--96

Korenromp, E. (2005). Malaria Incidence Estimates at Country Level for the Year 2004. Proposed Estimates and Draft Report. Mimeo.

Lucas, A.M. (2010). Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka. *American Economic Journal: Applied Economics*, 2(2), 46-71.

McCann, J. (In press). *The Historical Ecology of Malaria in Ethiopia.* Athens, OH: Ohio University Press.

MARA, Retrieved from http://www.mara.org.za/ on October 30, 2011.

Percoco, M. (2013). The Fight against Disease: Malaria Eradication and Regional Development in Italy", *Economic Geography*, 89(2):105-125

Spielman, A. & D'Antonio, M. (2001). *Mosquito: A Natural History of Our Most Persistent and Deadly Foe*. New York: Hyperion.

Teklehaimanot, H. D., Lipsitch, M., Teklehaimanot, A. & Schwartz, J. (2004). Weather-based Prediction of Plasmodium Falciparum Malaria in Epidemic-Prone Regions of Ethiopia I. Patterns of Lagged Weather Effects Reflect Biological Mechanisms. *Malaria Journal*, 3(41).

WHO. (2006). Economic Costs of Epidemic Malaria in Ethiopia, Final Report.

WHO. (2009). 10 Facts on Malaria. Retrieved from http://www.who.int/features/factfiles/malaria/en/index.html on November 1, 2011.

WHO. (2010). Rift Valley Fever. Fact Sheet n. 207.

Yeshiwondim, A.K., Gopal, S. Hailemariam, A.T., Dengela, D.O. & Patel, H.P. (2009). Spatial Analysis of Malaria Incidence at the Village Level with Unstable Transmission in Ethiopia. *International Journal of Health Geographics*, 8(1).

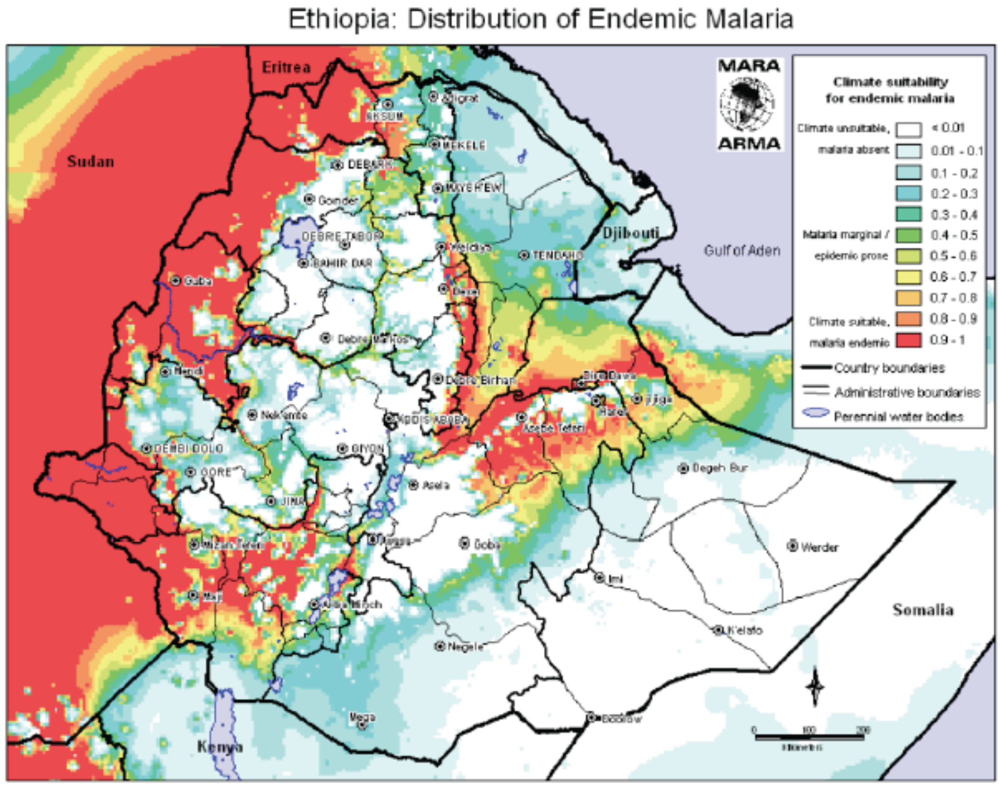

Figure 1: Predicted malaria incidence in Ethiopia. Source: MARA

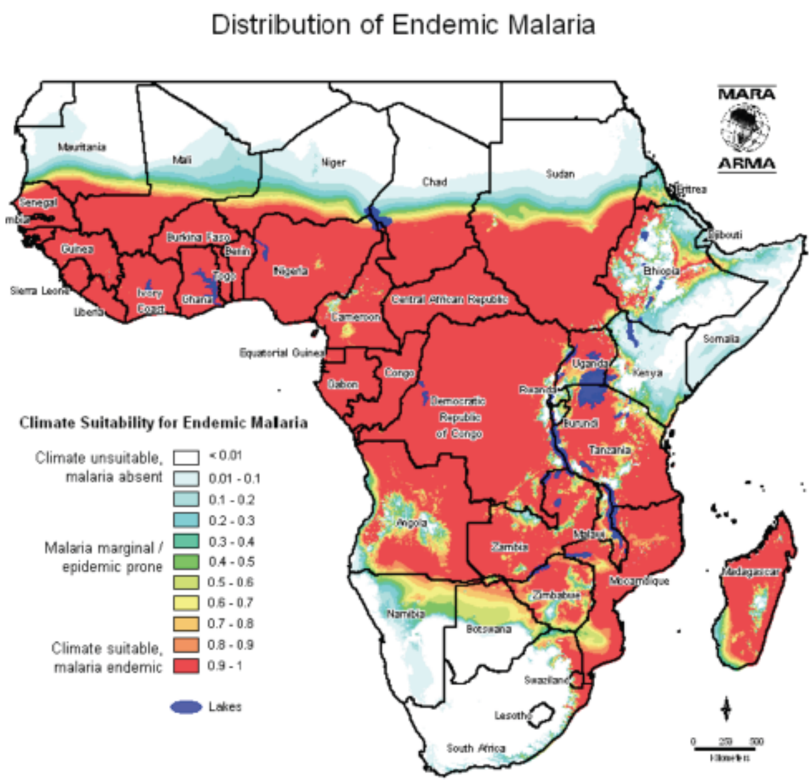

Figure 2: Predicted malaria incidence in Africa. Source: MARA

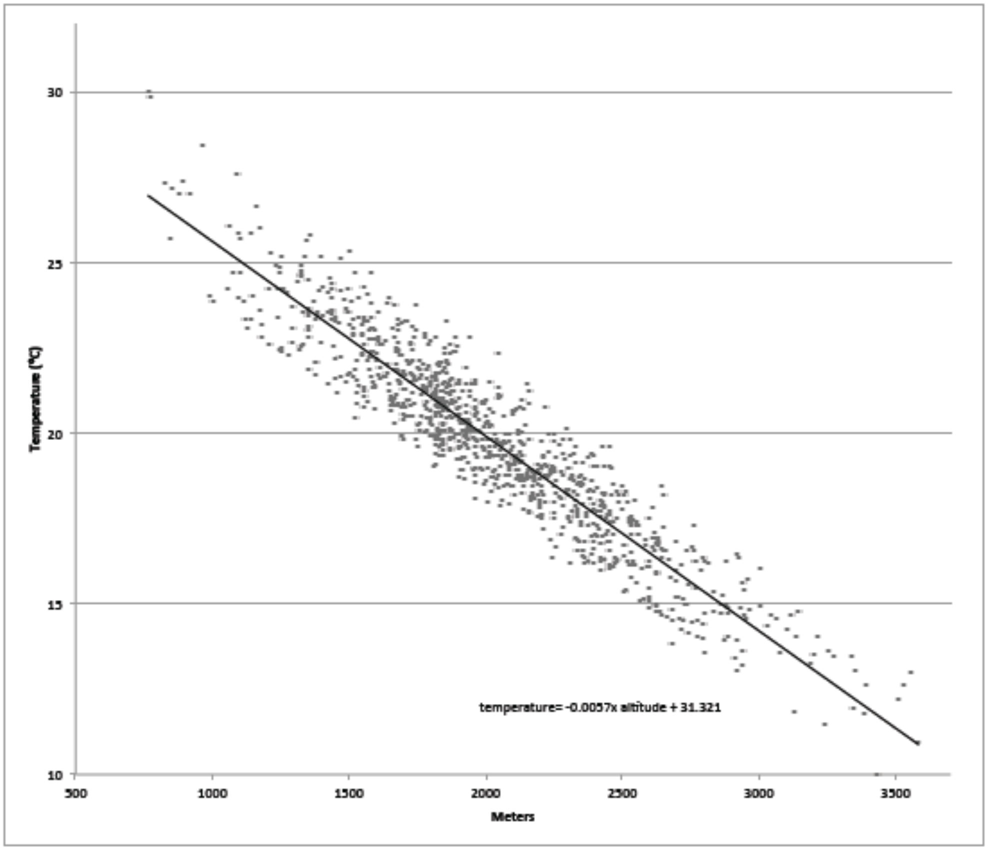

Figure 3: Temperature and elevation (all villages)

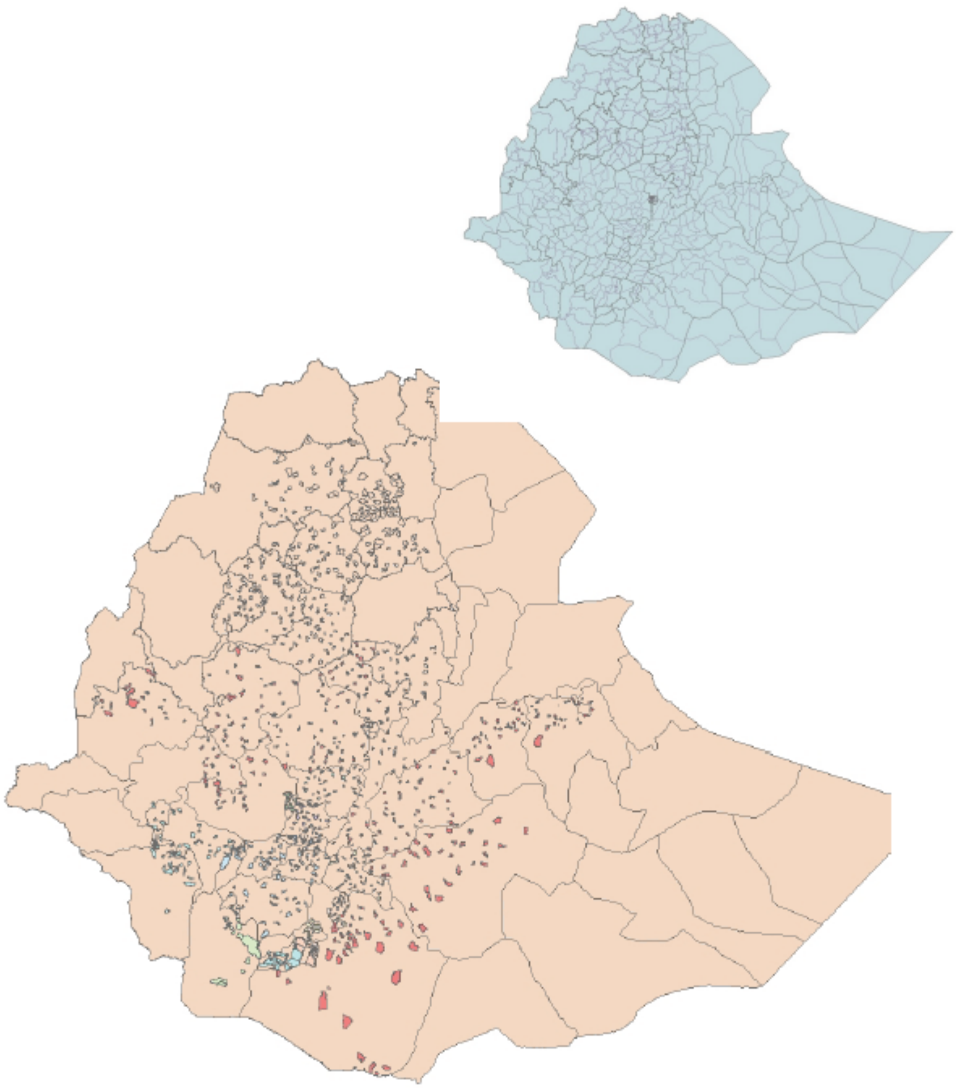

Figure 4: Distribution of administrative areas and location of study villages in Ethiopia. Bottom figure: study villages are colored. Administrative lines describe Administration zones. Top figure: distribution of provinces within zones.

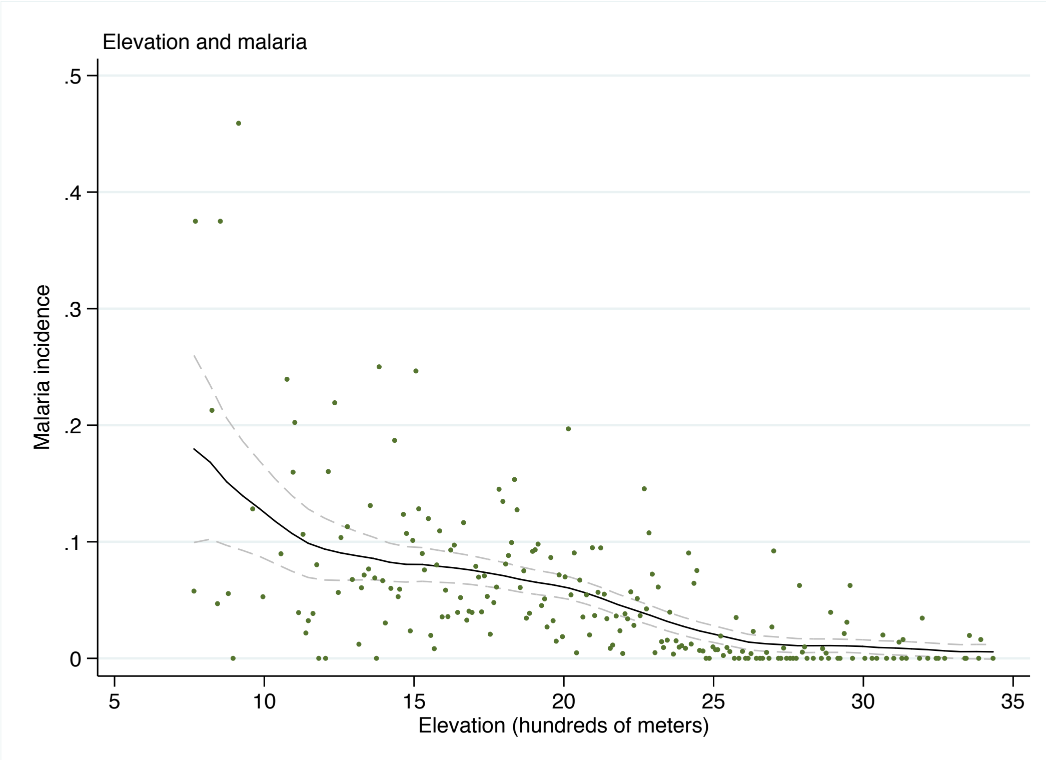

Fig 5: Local linear estimate of malaria and elevation. Each dot plots the average malaria for villages located within 10 meters of elevation.

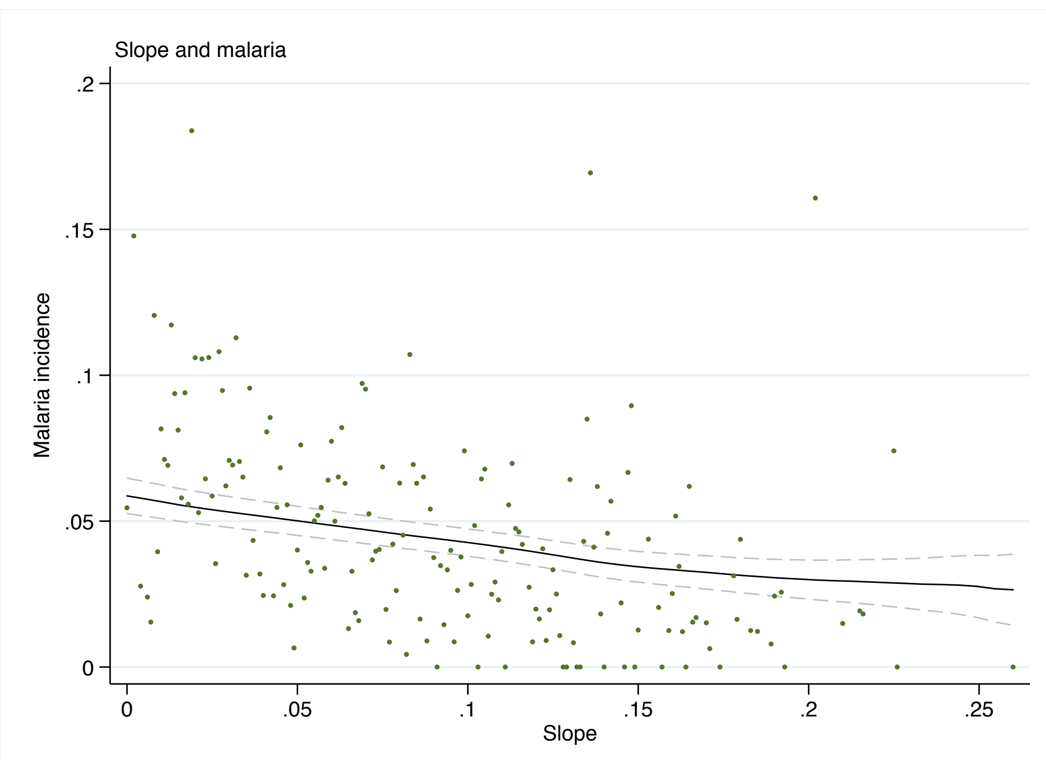

Fig 6: Local linear estimate of malaria and slope. Each dot plots the average malaria for villages located within 0.1% of slope.

Table 1: Summary statistics

	All villages		Villages below 2,500 meters	
	Mean	St. Dev.	Mean	St. Dev.
Topographic characteristics				
Elevation ('00 meters)	20.36	4.64	18.99	3.47
Slope (percentage)	0.0675	0.0475	0.0667	0.0470
Rainfall	12.17	3.17	12.16	3.31
Recent reported village health problems:				
All sickness fraction of village)	0.244	0.128	0.251	0.129
Malaria fraction of village)	0.057	0.095	0.066	0.100
Number of deaths in household (past 5 years)	0.287	0.425	0.292	0.446
Years of schooling:				
Children (age 7-19)	1.073	0.762	1.079	0.786
Adults (age >20)	1.058	0.931	1.078	0.955
Other outcomes				
Asset index	-0.975	0.600	-0.977	0.618
Labor past 7 days (age 10-17)	0.539	0.276	0.535	0.277
Food insecurity (past 5 years)	0.782	0.917	0.8082	0.9371
Household characteristics				
Fraction farming households	0.922	0.110	0.921	0.111
Fraction landless	0.041	0.066	0.042	0.066
Tenure (years)	10.069	5.608	9.712	5.229
Fraction female head	0.229	0.131	0.228	0.132
Number of livestock	4.676	4.097	4.258	3.916
Number of oxen	0.927	0.657	0.913	0.595
Number of shocks (past 5 years)				
Floods	0.133	0.393	0.136	0.406
Droughts	0.417	0.798	0.434	0.816
Loss of livestock	0.386	0.618	0.387	0.619
Price/other shocks	0.048	0.234	0.049	0.250
Distances to facilities (in hrs)				
School	0.710	0.599	0.715	0.626
Health facilities	5.069	7.185	4.864	6.951
Observations				
villages	1000		844	
provinces	296		276	

Table 2: Relationship between elevation, slope and reported health incidents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All sickness types		Malaria		Other health problem		Death	
Elevation ('00s meters)	-0.008***	-0.008***	-0.006***	-0.006***	-0.003	-0.002	-0.006	-0.006
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.011)	(0.010)
Elevation x Above 2,500 m		0.007		0.009***		-0.002		0.020
		(0.005)		(0.002)		(0.005)		(0.016)
Slope	-0.199	-0.152	-0.312***	-0.261***	0.111	0.107	-0.405	-0.276
	(0.151)	(0.120)	(0.081)	(0.060)	(0.123)	(0.104)	(0.351)	(0.297)
Sample of villages	<2,500m	All	<2,500m	All	<2,500m	All	<2,500m	All
Observations	844	1,000	844	1,000	844	1,000	844	1,000
R-squared	0.073	0.071	0.142	0.148	0.033	0.029	0.095	0.077
Province f.e.	YES	YES	YES	YES	YES	YES	YES	YES
P-value of F-test: Elevation + Elevation X Above = 0								
		0.927		0.1996		0.495		0.2635

OLS regressions at the village level. Odd columns include only villages with elevation <2,500 meters. Even columns include all villages in the sample. Dependent variable is the fraction of respondents reporting having some health problem in the prior 2 months. Death is the number of deaths in the family in the prior 5 years. Even columns include an above 2,500 meters dummy. Controls included: village average of rainfall, fraction agricultural households, fraction female headed household, average household size, number of livestock, oxen, wealth, land sizes, schooling of adult males and females, child age, and distance to schools and health clinics, number of droughts in the past 5 years. Errors clustered at the province level in parenthesis.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Elevation, slope, and reported health incidents

		(1)	(2)	(3)	(4)
		All sickness types	Malaria	Other health problem	Death
Elevation		-0.006	-0.009**	0.003	-0.011
		(0.005)	(0.004)	(0.004)	(0.017)
Elevation X quintile:					
	Slope 2	-0.001	-0.000	-0.001	0.006
		(0.005)	(0.004)	(0.004)	(0.015)
	Slope 3	-0.004	-0.000	-0.004	-0.010
		(0.005)	(0.004)	(0.004)	(0.013)
	Slope 4	-0.002	0.003	-0.005	0.004
		(0.005)	(0.004)	(0.004)	(0.017)
	Slope 5	-0.001	0.007*	-0.008*	0.014
		(0.006)	(0.004)	(0.004)	(0.018)
Observations		844	844	844	844
R-squared		0.080	0.162	0.045	0.104
Province f.e.		YES	YES	YES	YES
P-values of F-test: Elevation+ Elevation X quintile=0:					
Elev. 2nd quintile		0.105	0.004	0.678	0.679
Elev. 3rd quintile		0.001	0.000	0.551	0.043
Elev. 4th quintile		0.033	0.032	0.370	0.684
Elev. 5th quintile		0.056	0.507	0.065	0.831

Regressions on villages <2,500 only. Regression specification (1) is used in all columns. Controls include slope quintiles and other controls as in table 2. Outcome variables described in table 2. Errors clustered at theprovince level in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Topography correlates with other factors

Column title is dependent variable	(1) Floods	(2) Droughts	(3) Livestock losses	(4) Price shocks/ Other shocks	(5) School distance	(6) Health facility distance	(7) Land sizes
Elevation	0.003	-0.073**	0.025	0.005	-0.020	-0.081	-0.001
	(0.016)	(0.033)	(0.030)	(0.014)	(0.021)	(0.341)	(0.036)
Elevation x:							
Slope 2	-0.010	0.053*	0.028	0.005	0.015	0.376	-0.066*
	(0.014)	(0.028)	(0.025)	(0.017)	(0.020)	(0.326)	(0.037)
Slope 3	-0.006	0.048	0.002	-0.002	0.015	0.122	-0.047
	(0.016)	(0.029)	(0.029)	(0.013)	(0.018)	(0.369)	(0.031)
Slope 4	0.001	0.063*	0.014	0.002	-0.037	0.024	-0.037
	(0.016)	(0.035)	(0.030)	(0.013)	(0.033)	(0.342)	(0.033)
Slope 5	-0.015	0.056*	-0.049	-0.012	-0.030	-0.061	-0.077**
	(0.018)	(0.033)	(0.037)	(0.013)	(0.031)	(0.336)	(0.033)
Observations	844	844	844	844	844	844	844
R-squared	0.016	0.049	0.049	0.022	0.139	0.059	0.153
Province f.e.	YES	YES	YES	YES	YES	YES	YES
Mean dep. Var	0.132	0.414	0.386	0.048	0.712	4.83	2.88
P-value of F-test: Elevation + Elevation X Quintile = 0							
Elev., 2nd quintile=0	0.635	0.411	0.021	0.231	0.797	0.154	0.024
Elev., 3rd quintile=0	0.815	0.308	0.258	0.606	0.766	0.824	0.037
Elev., 4th quintile=0	0.720	0.732	0.079	0.316	0.071	0.739	0.134
Elev., 5th quintile=0	0.441	0.494	0.280	0.187	0.056	0.321	0.000

Sample of villages <2,500 meters. Shocks are the village average number of occurrences in the household in the prior 5 years. Tenure length is the number of years in the residence. Walking distance to facilities measured in hours. Land size is measured as quintiles of land size. Controls include slope quintiles and other controls as in table 2. Errors clustered at the province level.

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Relationship between village topography and schooling

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:	Villages <2,500 m		All villages		Villages < 2,500 m	
Average yrs. schooling	Children	Adults	Children	Adults	Children	Adults
Elevation	0.058***	0.041*	0.053***	0.033*	0.109***	0.064
	(0.013)	(0.023)	(0.013)	(0.020)	(0.029)	(0.044)
Slope	-2.112**	-3.232**	-2.117***	-3.220***		
	(0.815)	(1.286)	(0.694)	(0.986)		
Elevation X:						
Above 2,500 meters			-0.066**	-0.017		
			(0.027)	(0.024)		
Slope 2					-0.028	0.017
					(0.026)	(0.036)
Slope 3					-0.050*	0.008
					(0.029)	(0.038)
Slope 4					-0.057*	-0.038
					(0.031)	(0.042)
Slope 5					-0.068**	-0.052
					(0.030)	(0.044)
Observations	17,178	19,957	20,457	23,672	17,178	19,957
R-squared	0.214	0.067	0.213	0.066	0.215	0.069
Province f.e.	YES	YES	YES	YES	YES	YES
P-Value of F-test: Elevation + Elevation X Quintile = 0						
Elev. 2nd quintile=0					0.002	0.010
Elev. 3rd quintile=0					0.084	0.011
Elev. 4th quintile=0					0.020	0.407
Elev. 5th quintile=0					0.706	0.352

OLS regressions at the individual level. Children are aged 7-19, adults are aged 20+. Column 3 and 4 include all villages, remaining columns exclude villages located above 2,500 meters.

Controls for regressions on children schooling include: average village rainfall, farming household dummy, female headed household dummy, household size, number of livestock, oxen, household wealth index, land size, schooling of adult males and females in the household, child age, distance to schools and health clinics, number of droughts in the past 5 years.

Controls for regressions on adult schooling include rainfall, age, land size, distance to schools and health clinics, and number of droughts in the past 5 years. Columns 3 and 4 include an "above 2,500 meters" dummy. Columns 5 and 6 include slope quintiles. Errors clustered at the province level in parenthesis.

*** p<0.01, ** p<0.05, * p<0.1

Table 6: IV estimates of malaria intensity on schooling outcomes

	(1)	(2)	(3)	(4)	(5)
	OLS	IV	IV	IV	IV-LIML
Instruments		Elevation	Elevation	Elevation X	
Dep Var:		only	X Slope	Slope quintiles	
Years of schooling			quintiles	X Above 2500m	
A. All children aged 7-19					
Malaria intensity	0.294	-9.447***	-5.668	-6.332**	-6.779**
	(0.364)	(3.492)	(3.658)	(2.649)	(2.910)
Observations	17,178	17,178	17,178	20,457	20,457
B. Boys aged 7-19					
Malaria intensity	0.468	-9.352**	-5.850	-5.931**	-6.653*
	(0.536)	(3.843)	(4.561)	(2.970)	(3.442)
Observations	8,886	8,886	8,886	10,559	10,559
C. Girls aged 7-19					
Malaria intensity	0.259	-8.350**	-2.239	-5.437**	-5.833*
	(0.350)	(3.511)	(3.198)	(2.742)	(2.987)
Observations	8,292	8,292	8,292	9,898	9,898
D. All adults					
Malaria intensity	-0.064	-7.147	-8.554**	-4.412*	-4.550*
	(0.358)	(4.796)	(4.255)	(2.266)	(2.347)
Observations	19,957	19,957	19,957	23,672	23,672
E. Male adults					
Malaria intensity	-0.361	-8.141	-12.919**	-7.194**	-7.527**
	(0.515)	(5.732)	(6.067)	(3.083)	(3.268)
Observations	9,585	9,585	9,585	11,342	11,342
F. Female adults					
Malaria intensity	0.131	-6.883	-5.073	-2.356	-2.448
	(0.301)	(4.414)	(3.524)	(2.161)	(2.240)
Observations	10,372	10,372	10,372	12,330	12,330
Province f.e.	YES	YES	YES	YES	YES
Sample	<2,500m	<2,500m	<2,500m	All	All
F-test of excluded instruments		9.427	1.515	3.426	3.426

Table reports coefficients on instrumented village malaria.
Child controls included in panel A-C. Adult controls used in panel D-F.
Controls listed in table 5. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Estimates of malaria on other outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Excluded instruments	None	Elevation X Slope quintiles	Elevation X Slope X Above 2,500m		None	Elevation X Slope quintiles	Elevation X Slope X Above 2,500m		None	Elevation X Slope quintiles	Elevation X Slope X Above 2,500m	
	OLS	IV	IV	IV-LIML	OLS	IV	IV	IV-LIML	OLS	IV	IV	IV-LIML
Column title is dependent variable	Child labor (children >10 y.o. only)				Food insecurity (All adults)				Asset index (All adults)			
Village malaria	0.321**	-0.160	1.524*	1.685*	1.235***	3.125	2.916	3.149	-0.100	-0.510	-0.643	-0.697
	(0.143)	(1.608)	(0.865)	(0.990)	(0.382)	(2.869)	(2.270)	(2.594)	(0.282)	(3.139)	(2.057)	(2.260)
Sample of villages	<2,500	<2,500	All	All	<2,500	<2,500	All	All	<2,500	<2,500	All	All
Observations	11,563	11,563	13,780	13,780	19,957	19,957	23,672	23,672	19,955	19,955	23,670	23,670
R-squared	0.005	0.142	0.113	0.108	0.379	0.373	0.350	0.348	0.263	0.263	0.234	0.234
Province f.e.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
F-test of excluded instrum.		1.33	3.75	3.75		1.87	3.43	3.43		1.87	3.43	3.43

Columns 1-4 include child regression controls. Remaining columns include adult regression controls (see table 5). Child labor is thefaction of children 10-17 who reported working in the prior 7 days. Food security refers to the average number of times households faced food shortages in the prior 5 years. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Climate vs. Topographical estimates of malaria

	(1)	(2)	(3)	(4)
VARIABLES	Village malaria		Dummy: Village has some malaria	
MARA model	0.030**	-0.012	0.328***	0.226**
	(0.012)	(0.025)	(0.066)	(0.112)
Elevation		-0.007**		-0.018
		(0.004)		(0.016)
Elevation X Above 2,500m		0.010***		0.029
		(0.004)		(0.022)
Observations	1,000	1,000	1,000	1,000
R-squared	0.134	0.149	0.125	0.138
Province f.e.	YES	YES	YES	YES
P-value of joint significance of elevation and elevation X Above		0.017		0.55

Village regressions on malaria. See table 2 for list of village controls. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Climate IV estimates of malaria effects on schooling outcomes MARA model predictions

Dep Var:	(1)	(2)	(3)	(4)
Years of schooling	IV	IV	IV	IV
Instrumented variable	Malaria intensity		Malaria presence	
A. Education children aged 7-19				
Malaria measure	-9.715**	-16.188	-0.864***	-1.233**
	(4.428)	(12.619)	(0.304)	(0.596)
Observations	20,457	20,457	20,457	20,457
B. Education adults				
Malaria measure	-5.786	-7.635	-0.509	-0.606
	(5.325)	(9.617)	(0.393)	(0.634)
Observations	23,672	23,672	23,672	23,672
C. Child Labor				
Malaria measure	2.140*	3.704	0.192*	0.277
	(1.234)	(3.200)	(0.111)	(0.211)
	13,209	13,209	13,209	13,209
Province f.e.	YES	YES	YES	YES
Elevation controls	None	Linear	None	Linear
Sample	All	All	All	All
F-test of excluded instruments	8.765	2.307	36.34	10.96

IV regressions using MARA model as instrument for malaria. Individual controls for panel A and C are children controls, for panel B are adult controls. See table 5 for list of controls. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1