THE ECONOMIC JOURNAL

The Economic Journal, **128** (*May*), 1207–1234. Doi: 10.1111/ecoj.12433 © 2016 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

MALARIA AND EARLY AFRICAN DEVELOPMENT: EVIDENCE FROM THE SICKLE CELL TRAIT*

Emilio Depetris-Chauvin and David N. Weil

We examine the effect of malaria on economic development in Africa over the very long run. Using data on the prevalence of the mutation that causes sickle cell disease, we measure the impact of malaria on mortality in Africa prior to the period in which formal data were collected. Our estimate is that in the more afflicted regions, malaria lowered the probability of surviving to adulthood by about ten percentage points, which is twice the current burden of the disease. The reduction in malaria mortality has been roughly equal to the reduction in other causes of mortality. We then ask whether the estimated burden of malaria had an effect on economic development in the period before European contact. Using data at the ethnic group level, we find little evidence of a negative relationship between malaria burden and population density or other measures of development.

It is impossible to understand the pattern of comparative economic development in the world today without understanding comparative development in the past. Consider, for example, a horizon of 500 years. Countries and regions that were highly developed as of the year 1500 are, for the most part, among the most developed today. Exceptions to this regularity, such as China, tend to be growing quickly. Taking into account flows of population over the last half millennium makes this correlation even stronger: countries populated by people whose ancestors lived in the most developed countries are most likely to be rich today. Looking within countries, people who are descended from parts of the world that were highly developed in the year 1500 are on average higher up in the income distribution than people descended from regions that were not developed (Chanda and Putterman, 2007; Putterman and Weil, 2010). Going back further back in time, there is still strong predictive power of past development for present outcomes. Comin *et al.* (2010) show that not only is a country's level of technology from 500 years ago predictive of income today but so is the level of technology 2,000 or 3,000 years ago. Hibbs and Olsson (2004) show that the date at which the transition from hunting and gathering to settled agriculture took place is predictive of a country's income today.

The fact that development in the past is so predictive of development today suggests two possible theories. First, it may be that the same factors that influenced development in previous historical eras are still operative in the present. Examples of such factors are genetic attributes of populations, slowly changing aspects of culture

* Corresponding author: David N. Weil, Department of Economics, Brown University, Box B, Providence, RI 02912, USA. Email: davidweil@brown.edu.

We are grateful to Ronald D. Lee, Andrew Mason, Gordon McCord, Nathan Nunn, Marcella Alsan and Fred Piel for sharing data, to Fatima Aqeel, Federico Droller, Daniel Prinz and Scott Weiner for research assistance, and to Quamrul Ashraf, Adriana Lleras-Muney, Stelios Michalopoulos and seminar participants at Boston University, Harvard University, Harvard Center for Population and Development Studies, Indiana University, New York University, Pennsylvania State University, Princeton University, UCLA, University of Connecticut and University of Pennsylvania for helpful comments

or institutions, and characteristics of geography or climate.[1] Alternatively, it may be that the specific factors that caused past underdevelopment are no longer relevant today but that the fact of past underdevelopment itself is causal of current underdevelopment. For example, it could be that the early development advantage of the Eurasian land mass arose from the historical presence of plentiful species of large seeded grasses and domesticable animals, as argued by Diamond (1997), but that the continuation of the development gap between Eurasia and other regions results from the effect of colonial institutions that Europeans were able to impose on much of the rest of the world as a result of this initial advantage (Acemoglu *et al.*, 2001, Nunn, 2008). A related argument, stressed by Spolaore and Wacziarg (2013) is that past differences in development are causal for current outcomes because of barriers to transmission of productivity-enhancing innovations among populations with different historical roots.

Whichever of these theories is correct (and obviously it is possible for both of them to have some validity), there is clearly much to be learned by looking at the roots of development differences in the past. In this article, we examine the historical impact on development of malaria, which is one of the most significant diseases in the world today in terms of its humanitarian burden. Malaria's control is widely studied by biologists and social scientists. Economists actively debate its role in affecting growth in the modern world.[2] However, as the above discussion makes clear, it would be possible that even if malaria were not important in affecting economic development today directly, it could nonetheless have been an important determinant of development historically and via that channel indirectly affect development today.

In studying the role of malaria in long-run development, we are also inevitably studying the long run development of Africa. Both historically and today, Africa has been the major focus of the disease. Indeed, malaria was not present in the tropical regions of the new world until it was accidentally brought there by Europeans (McNeill, 1977). Historians of Africa attribute a large role to diseases in general, and malaria in particular, in shaping development (Akyeampong, 2006). For example, Webb (2006) describes malaria, along with trypanosomiasis (transmitted by the tsetse fly) as having profoundly influenced African patterns of settlement as well as culture (Alsan, 2015, also finds a large role for trypanosomiasis in shaping population density in Africa). Chiovelli *et al.* (2015) claim that long term malaria exposure affected modern ethnic diversity, in particular contributing to more frequent endogamic marriages and stronger ethnic identities.[3]

[1] See, for example, Maleney *et al.* (2004) and Ashraf and Galor (2013). Spolaore and Wacziarg (2013) discuss the different possible channels by which characteristics that impact economic outcomes might be transmitted intergenerationally.

[2] The widely quoted estimate of Gallup and Sachs (2001) is that malaria reduces growth of GDP per capita by 1.3% per year in the African countries most afflicted. See Weil (2010) for an extensive critique of this literature.

[3] Weil (2014) paints a picture of African development in 1500, both relative to the rest of the world and heterogeneity within the continent itself, using as his indicators population density, urbanisation, technological advancement, and political development. Ignoring North Africa, which was generally part of the Mediterranean world, the highest levels of development by many indicators are found in Ethiopia and in the broad swathe of West African countries running from Cameroon and Nigeria eastward along the coast and the Niger river. In this latter region, the available measures show a level of development just below or sometimes equal to that in the belt of Eurasia running from Japan and China, through South Asia and the Middle East, into Europe. Depending on the index used, West Africa was above or below the level of development in the Northern Andes and Mexico. Much of the rest of Africa was at a significantly lower level of development, although still more advanced than the bulk of the Americas or Australia.

Analysis of the role played by malaria in shaping historical development is severely hampered by a lack of data. Biologists only came to understand the nature of the disease in the late nineteenth century. Accounts from travellers and other historical records provide some evidence of the impact of malaria going back millennia but these are hardly sufficient to draw firm conclusions (Mabogunje and Richards, 1985, Akyeampong, 2006). Even today, trained medical personnel have trouble distinguishing between malaria and other diseases without the use of microscopy or diagnostic tests. As discussed below, there are data (malaria ecology) that measure the extent to which the environment in different geographical regions is supportive of malaria. One can look at the empirical relationship between this malaria ecology measure and economic development, either currently or historically. However, such an approach faces severe limitations. One problem is that the malaria ecology variable is not scaled in a way that allows for easy economic interpretation (e.g. one cannot readily compare the coefficient of malaria ecology in a regression to the output of a quantitative model). Further, it is difficult to know whether one has controlled for correlates of malaria ecology that might independently influence the process of development. These correlates could be other factors that directly influence output today (either other diseases, or the tropical climate, which affects agriculture) or they could be the result of historical processes, for example, institutional quality, which Acemoglu *et al.* (2001) argue was influenced by the disease environment.

In this article, we address the lack of information on malaria's impact historically by using genetic data. In the worst afflicted areas, malaria left an imprint on the human genome that can be read today. Specifically, we look at the prevalence of the gene that causes sickle cell disease. Carrying one copy of this gene provided individuals with a significant level of protection against malaria but people who carried two copies of the gene died before reaching reproductive age (Hedrick, 2011). Thus, the degree of selective pressure exerted by malaria determined the equilibrium prevalence of the gene in the population. By measuring the prevalence of the gene in modern populations, we can thus back out estimates of the severity of malaria historically. We compare these estimates to the burden of malaria today and discuss how the change in malaria mortality has compared to the change in mortality from other sources.[4]

With estimates of the extent of malaria mortality in hand, we then turn to look at the impact of the disease on economic development. We present regressions of a number of measures of development within Africa on a malaria burden measure that we create based on sickle cell prevalence. Of particular note, we apply our analysis to a data set measured at the level of ethnic groups as an alternative to more common country-level analyses. We present simple OLS results, as well as results in which we instrument for malaria burden, using an index of climactic suitability for malaria transmission. The result of this statistical exercise is that we find no evidence of malaria burden negatively affecting historical economic development. We also briefly discuss a more theoretical

[4] As discussed below, data on the prevalence of the sickle cell trait among native populations is also available outside Africa. We do not extend our analysis to other continents because in these regions the prevalence of the trait does not serve as a good indicator of the historical burden of malaria. In many parts of the world that historically had high malaria burdens (including New Guinea, Southeast Asia and parts of China), the mutation is absent. Similarly, in places where malaria itself was absent prior to 1492 but played an important role later on (such as Brazil and central America), the mutation is absent among native peoples.

approach, using a macrodemographic model to quantify the effect of malaria mortality of the magnitude estimated from our genetic data on population density. The model predicts that the effect would be small.

The rest of this article is organised as follows. In Section 1, we discuss the biology of the link between malaria and sickle cell disease. Section 2 presents and applies our model for using the current level of sickle cell prevalence to estimate the historical burden of malaria, including comparisons of malaria's historical burden to its current level. We then turn to the question of how malaria affected development. Section 3 presents a statistical analysis of the relationship between the malaria burden measure we create and a number of measures of development within Africa and also discusses a more model-based approach to the issue. Section 4 concludes.[5]

1. Malaria and Sickle Cell Disease

Malaria is caused by the plasmodium parasite, which is transmitted to humans through the bite of a female anopheles mosquito. Early symptoms of malaria include fever, chills, severe headache and vomiting. In severe cases, these are followed by respiratory distress, severe anaemia or neurological symptoms. Infants are protected from the disease in the first few months of life by a combination of maternal antibodies and characteristics of the structure of foetal haemoglobin. In malaria endemic areas, most children have developed substantial immunity by the age of five.

Africa currently accounts for 82% of world malaria cases and 90% of world malaria deaths (World Health Organization, 2014). The geographical pattern of malaria's severity is largely determined by the climactic conditions that support mosquito breeding as well as by the mix of mosquito species present. There are significant differences in the vectorial capacity among the approximately 20 species of anopheles that transmit malaria, based on factors such as the mosquitoe's preferred targets, biting frequency and lifespan. The most effective vector, Anopheles gambiae, is the principal vector in Africa.

Several mutations have arisen in human populations that provide resistance to malaria. These include the mutation causing thalassemia, which is present in Mediterranean, Arab and Asian populations; the absence of the Duffy blood group in west Africa; haemoglobin E in Southeast Asia; and haemoglobin C in West Africa. The most important such mutation is the one that causes sickle cell disease (Allison, 2002; Nelson and Williams, 2006).

The sickle cell trait is a mutation in the gene the produces haemoglobin, the oxygen-carrying component in red blood cells. Individuals carry two copies of this gene, one received from each parent. Individuals who carry one normal copy of the gene (referred to as A type) and one copy with the sickle cell mutation (S type) are carriers of the disease. In individuals of the AS genotype, a fraction of the haemoglobin in their red blood cells have an abnormal structure. In individuals who have two copies of the sickle cell gene (SS genotype), almost all haemoglobin molecules are of the abnormal type.

[5] Weil (2014) also discusses the link between malaria, the sickle cell trait, and early development, although without the formal model or econometric approach of the current article.

In conditions of inadequate oxygen supply (hypoxia), haemoglobin produced by the S gene becomes rigid, leading to a characteristic sickle shape of red blood cells. Carriers of sickle cell trait generally do not suffer many adverse effects.[6] However, there can be negative consequences from sickling in conditions of low oxygen such as unpressurised airplane flights and extremely rigorous exercise (Motulsky, 1964). In individuals of the SS genotype, such sickling of red blood cells is far more common, leading to acute episodes of disease in which abnormally shaped cells restrict blood flow to organs. Such individuals also suffer from anaemia and reduced resistance to infection. In 1994, life expectancy for SS children in the United States was 42 years for males and 48 years for females. In the absence of modern medical care, individuals of the SS genotype are not able to survive to adulthood.

The sickle cell mutation is relevant to malaria because infection of a red blood cell with the malaria parasite leads to hypoxia. In individuals of the AS genotype such blood cells sickle and are then eliminated by macrophage cells of the body's immune system, lessening the burden of infection (Luzzatto, 2012). Carriers of the sickle cell trait are particularly resistant to severe malarial episodes; they are less resistant to mild cases. The mechanism by which carriers are protected from malaria is different from the acquired immunity that both AA and AS individuals achieve following repeated exposure to the disease.

The benefit that possessing a single copy of the sickle cell gene conveys counterbalances the biological cost incurred when homozygous SS children are stricken with sickle cell disease. An individual of the AS genotype is more likely to reach adulthood than is an individual of the AA genotype but the former is also more likely to see his/her child die of sickle cell disease. This is known as a heterozygote advantage or balanced polymorphism. As shown more formally below, the stronger the pressure of malaria on survival, the more advantaged are individuals who carry the S gene and, in equilibrium, the higher the percentage of the population who will be carriers. Indeed, it was the correlation of high prevalence of the sickle cell gene and the presence of malaria that first led scientists to understand the protective role of the sickle cell mutation (Kwaitkowski, 2005). As will be seen in the next Section, the underlying genetic mechanism by which the sickle cell trait is transmitted provides a means of mapping sickle cell prevalence into an estimate of the mortality burden of malaria.

Piel *et al.* (2010) present a global geo-database of S gene frequency based on comprehensive electronic search of academic publications presenting S gene frequency figures. Each reference included in the data set meets the criterion that the surveyed population was representative of the indigenous population of a particular location. Piel *et al.* assign a geographic coordinate to all samples with the distribution of AS and AA genotypes that meet their strict inclusion criteria. Using a Bayesian model-based geostatistical framework they then create a continuous map of

[6] Williams *et al.* (2005) show the absence of any significant effect of carrier status on a wide range of childhood diseases.

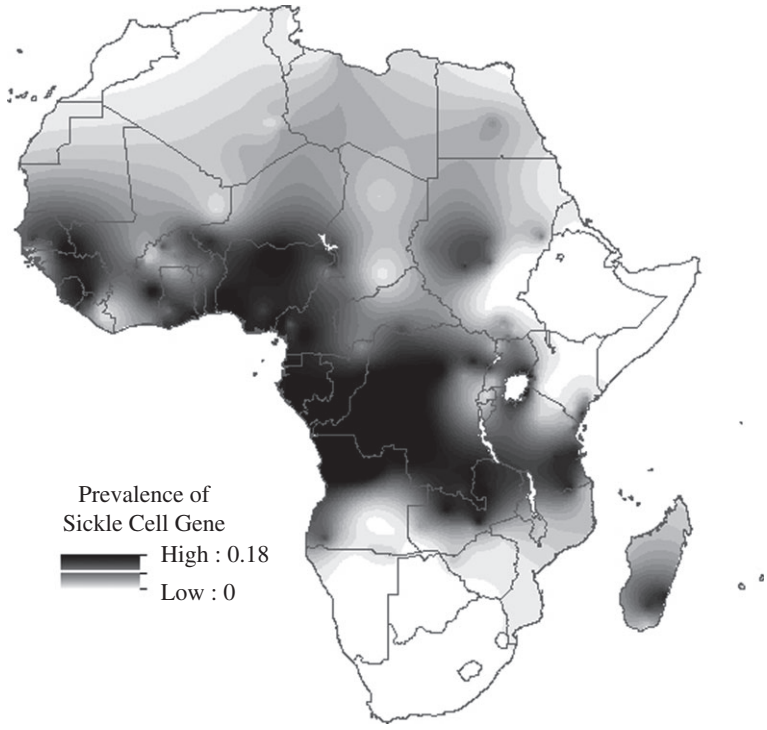

Fig. 1. *Sickle Cell Gene Frequency from Piel et al. (2010)*

the sickle cell gene frequency resulting in 10 km by 10 km resolution global raster grid.[7] It is important to note that throughout our article, we use the terminology sickle cell trait prevalence (the fraction of people who are carriers of the allele S) instead of sickle cell gene (or S allele) frequency as in Piel *et al.* (2010). Since very few S homozygotes survive to adulthood, the sickle cell trait prevalence is very close to twice the sickle cell gene frequency.

Figure 1 presents the geographic distribution of S gene frequency from Piel *et al.* (2010). The maximum levels of the mutation are located in West-Central Africa (Gabon and Congo). The west part of Africa, from Cameroon to Senegal, shows a large heterogeneity in the prevalence of the mutation, ranging from almost complete absence in some places and very high values (above 15%) in others. Medium level frequencies are also located in the proximity of Lake Victoria. Finally, no mutation is documented in the southern part of the continent.[8]

[7] To compute the map of sickle cell distribution, 773 georeferenced independent samples were used worldwide of which 448 were located in Africa.

[8] Genetic evidence supports the hypothesis that the sickle cell trait observed today results from five independent mutations, four in Africa and one in Asia. The four African haplotypes are Benin, Senegal, Cameroon and Bantu (Central Africa). Uncertainty exists regarding the dates of these mutations. Some scholars have argued that they probably arose fewer than 6,000 years ago in response to the spread of agriculture (Livingstone, 1958; Weisenfeld, 1967). Molecular genetic analyses have yielded a broad range of estimates, ranging from between 3,000 and 6,000 generations ago (Kurnit, 1979; Soloman and Bodmer, 1979) to, in the case of the Senegal haplotype, less than 3,000 years (Currat *et al.*, 2002).

2. Measuring the Historical Burden of Malaria Using Data on the Sickle Cell Trait

2.1. *Model*

Our goal is to examine what the prevalence of the sickle cell trait among African populations tells us about the impact of malaria historically. As described above, every adult carries two alleles. These can be either sickle cell (S) or normal (A). A person with the SS genotype will develop sickle cell disease and not survive to adulthood. A person who carries the sickle cell trait (AS) will have a survival advantage against malaria in comparison to someone who does not (AA).[9]

We consider a simple model in which deaths occur due to either malaria or to other causes. Let M be the probability of dying of malaria, conditional on not dying of something else. Similarly, let P be the probability of dying of something else, conditional on not dying from malaria.[10] The deaths that we are concerned with are those between birth and adulthood, which is taken to be the age at which children are produced. The number of adults from a cohort of newborns will be given by:

$$Surviving\ Adults = (1 - M)(1 - P)Newborns \tag{1}$$

Throughout the analysis, we will assume that the probability of dying from non-malaria causes, P, is the same for individuals with the AS and AA genotypes. The probabilities of dying from malaria, differ between AA and AS genotypes. We designate these probabilities M^{AA} and M^{AS}. It is also useful to designate the relative survival rates of these two genotypes, which we call β:

$$\beta = \frac{(1 - M^{AA})(1 - P)}{(1 - M^{AS})(1 - P)} = \frac{(1 - M^{AA})}{(1 - M^{AS})}. \tag{2}$$

β is the probability of a non-carrier living into adulthood relative to the probability of a carrier living into adulthood. The smaller is β, the larger is the advantage of the AS genotype. A value of $\beta = 1$ would indicate that there is no advantage to carrying the sickle cell gene. The values for M^{AA} and M^{AS}, and thus for β, will depend on both the disease environment and the state of medical technology. For example, in a place where there are no malaria mosquitoes, M^{AA} and M^{AS} will both be equal to zero and β will be equal to one. Clearly, the availability of modern medical care should mean that both M^{AA} and M^{AS} are lower today than they were in the past. However, it is not clear *a priori* which mortality rate would be reduced by more.

2.1.1. *Relative survival in modern populations*

Although our main objective is to ask what role malaria played historically, it is of interest to see what information is available about relative survival today. One way to

[9] We ignore other mutations. In the presence of other mutations that also control malaria, the genetic benefit of carrying the S trait is reduced but its cost remains the same. Thus our analysis, based only on the S trait, will understate the full burden of malaria.

[10] For the purposes of this part of the analysis, it does not matter whether deaths due to sickle cell disease are counted as being due to malaria or due to a non-malaria cause. In our analysis below, we include deaths due to sickle cell disease as part of our measure of malaria burden, because in the absence of the disease, such deaths would not occur.

measure relative survival is to compare prevalence among adults to that among newborns. The relevant equation is as follows:

$$\frac{AA\ Adults}{AS\ Adults} = \frac{(1 - M^{AA})(1 - P)AA\ newborns}{(1 - M^{AS})(1 - P)AS\ newborns}. \tag{3}$$

Rearranging the equation:

$$\beta = \frac{(1 - M^{AA})}{(1 - M^{AS})} = \frac{AA\ Adults}{AS\ Adults} \times \frac{AS\ newborns}{AA\ newborns}. \tag{4}$$

Morrow and Moss (2006) report on data from the Garki study, conducted in a region of high malaria transmission in Nigeria in the 1970s. Among adults, 29.0% of the population were AS, while 70.2% were AA. Among newborns, 23.6% were AS and 73.8% AA. Entering these figures in the equation above gives a value of $\beta = 0.775$. More generally, Morrow and Moss report that the prevalence of the sickle cell trait in West Africa trait rises from 20% to 24% among newborns to 26–29% among adults.[11]

Another approach is to look at survival directly. Motulsky (1964, table 2) examined the relative survival of over 15,000 children in the Congo in the 1950s, a time and place where modern treatments for malaria would have been relatively scarce. The study compared families where one parent was AS and one AA, on the one hand, to families in which both parents were AA, on the other. Since half of the children in the former group were carriers, compared to none in the latter, one can back out the relative survival of AS *versus* AA children. The study found mortality from all causes of 24.0% among AS children *versus* 27.4% among AA. The implied value of β is 0.955. Part of the explanation for the different estimates of β in the Congo *versus* West Africa may be a difference in the severity of malaria. In the Congo, the malaria ecology index (discussed below) is 12.1; in West Africa it is generally in the neighbourhood of 20.

2.2. *Measuring Relative Survival in Historical Populations*

We now turn to our main line of inquiry, which is using observed frequency of the sickle cell trait to back out the severity of malaria. Let π_t be the fraction of adults in generation t who are carriers (AS). We assume that no one born with SS lives into adulthood. Thus, the fraction of the adult population who are not carriers is $(1 - \pi_t)$. The fraction of alleles in the adult generation that are S is simply $\pi_t/2$. Assuming that mating between carriers and non-carriers is random, the fractions of children born who are of each type are as follows:

(*i*) *AA*: $(1 - \pi_t/2)^2$;
(*ii*) *AS*: $\pi_t(1 - \pi_t/2)$; and
(*iii*) *SS*: $(\pi_t/2)^2$.

[11] Examining the Baamba of Uganda, who live in an extremely high malaria environment, Lehmann and Raper (1956) report that the fraction of sicklers (AS or SS) rises from 30% of those under five years age to 37% among adults. Using these figures in (4), this would imply a value of $\beta = 0.73$. This calculation understates the size of the mortality differential, however, both because those in the under five age group will have already experienced differential mortality from malaria and because the group of young sicklers presumably includes a larger fraction of SS individuals than does the adult group.

The difference equation for π, which relates prevalence among adults in successive generations, is as follows:

$$\pi_{t+1} = \frac{\pi_t\left(1-\frac{\pi_t}{2}\right)}{\pi_t\left(1-\frac{\pi_t}{2}\right)+\beta\left(1-\frac{\pi_t}{2}\right)^2} = \frac{\pi_t}{\beta+\pi_t\left(1-\frac{\beta}{2}\right)}. \tag{5}$$

We solve for the steady state by setting $\pi_t = \pi_{t+1}$:

$$\pi_{ss} = \frac{1-\beta}{1-\frac{\beta}{2}}. \tag{6}$$

This has the properties we would expect: the smaller is the β, that is the greater is the survival advantage of being a carrier, the larger is the equilibrium fraction of the adult population that will be carriers.

We can turn this equation around to infer the burden of malaria on survival based on the prevalence of the sickle cell trait among adults:

$$\beta = \frac{1-\pi_{ss}}{1-\frac{\pi_{ss}}{2}}. \tag{7}$$

This equation says that if 20% of the adult population are carriers in a steady state, then $\beta = 0.89$, in other words non-carriers are only 89% as likely to live to adulthood as carriers.[12]

2.2.1. *Current versus historical prevalence*

The analysis above considers a population that is in equilibrium in terms of the selective impact of malaria and the prevalence of the sickle cell trait. Such a steady state presumably existed in Africa in the period before European contact. However, the only data on sickle cell prevalence available comes from observations over the last 70 years.[13] One might worry that modern prevalence rates are not the same as those that held historically, because the health environment has been changing over time. To address this question, it is straightforward to use (5) to look at how prevalence π changes in response to a change in relative survival β.

As an example, we consider the case where there is initially a steady state of $\pi = 0.20$, and correspondingly $\beta = 0.89$. In generation 1, the value of β is set to one,

[12] Positive assorting mating (carriers more likely to breed with other carriers) would, for a given prevalence of the S allele in the parent generation, raise the fraction of children who were AA and SS while lowering the fraction who were AS. It can be shown that in this case, the steady-state level of prevalence, for a given value of β, would be lower than in the case of random mating and, similarly, that for a given level of steady-state prevalence, the implied value of β would be lower. Negative assortive mating would shift the results in the opposite direction. Positive assortive mating could occur if there were several ethnic groups with different levels of sickle cell prevalence living in the same area and practicing homogamy. However, in the long run, we would not expect this situation to persist if the groups all faced the same selective pressure from malaria mortality.

[13] Testing predates modern technology for genetic analysis. Carriers of the sickle cell gene can be reliably diagnosed by taking a drop of blood and mixing in a reducing agent to induce hypoxia, then examining with a microscope whether blood cells have sickled.

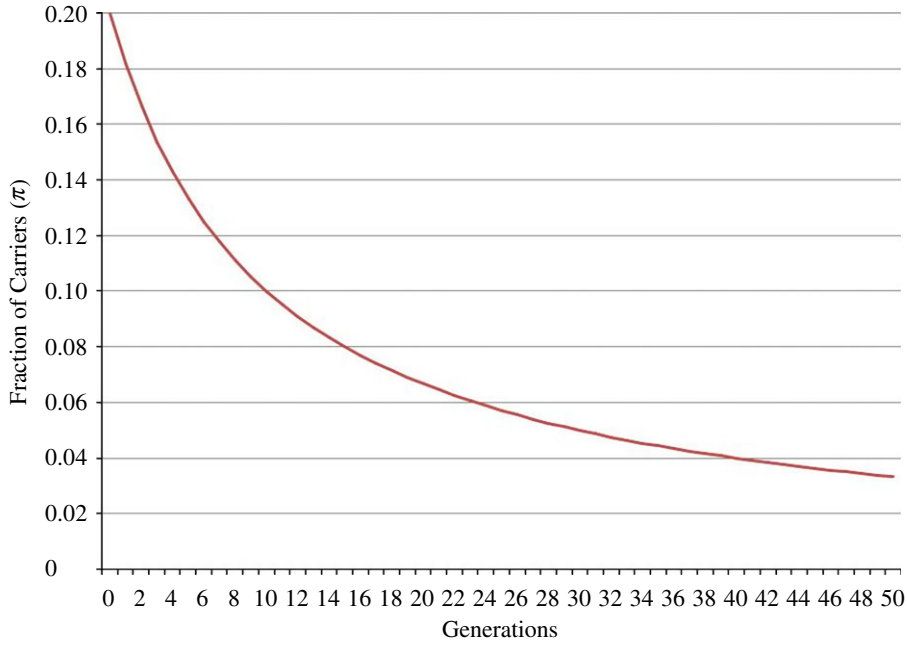

Fig. 2. *Effect of a Hypothetical Malaria Eradication on Evolution of Carriers*
Note. Colour figure can be viewed at wileyonlinelibrary.com.

corresponding to the complete eradication of malaria, or the removal of the population to a place where the disease is not present. Figure 2 shows the fraction of the population that will be carriers of the sickle cell trait after a set number of generations. The Figure shows that the initial decline is very rapid but that there is long tailing off once the prevalence gets sufficiently low.

In principle, one could test this model by looking at how the prevalence of the sickle cell trait changes in a population that was removed from exposure to malaria at a known time. Although the data are in practice not sufficient for a formal test, the case of African Americans provides a rough consistency check. We start by calculating the fraction of adults who are carriers in the countries that were the source of slaves that came to the United States. Specifically, we combined two data sources. We used data from Piel *et al.* (2010) on the prevalence of the sickle cell trait among adults in modern African countries. We combined these with estimates from Putterman and Weil (2010, main appendix, part II.3) on the fractions of the slaves who came to the United States originating in each African country. Multiplying and summing these two series, we estimate that 16.2% of slaves who came to the United States were carriers of the sickle cell trait (calculation available on request).

The other piece of information needed is the date on which the ancestors of today's African Americans left Africa. Obviously, this took place over several centuries, and we do not know of an estimate of the average date of departure. As a rough estimate, we assume that the average ancestor of today's African Americans left Africa around 1750 – a span of roughly 10 generations.

Starting with 16.2% prevalence and applying (5) for ten generations would lead to a prevalence of 9.0% today. However, one has to deal with the issue of admixture with other populations. Putterman and Weil (2010, Appendix B) summarise literature on the fraction of African American heritage that is due to non-Africans, reporting 20% as a consensus figure. That admixture took place at unknown points in time. Mechanically, the later that admixture took place, the lower would be the prevalence of the sickle cell trait among African Americans today. If the entire admixture took place in the last generation, the prevalence today would be 7.2%. By contrast, if the entire admixture took place ten generations ago, our calculation would yield a prevalence of 7.9% today.

In fact, the prevalence of the trait among African Americans today is 8% (National Heart, Lung, and Blood Institute, 1996). Thus, the model slightly under-predicts the prevalence of the sickle cell trait among African Americans. One possible source of this error could be that that African slaves brought to the Americas did not find themselves in a completely malaria-free environment. For example, McGuire and Coelho (2011) stress the immunity of slaves to malaria as one of the reasons that southern planters favoured them over indentured servants.

In terms of our use of current prevalence of the sickle cell trait to measure the historical burden of malaria, it is not clear that this analysis of the dynamics of prevalence matters much, since there is little reason to believe that contact with Europeans did anything to reduce the impact of malaria in Africa until the second half of the twentieth century, which is when most of the measures of sickle cell come from.

2.3. *Measuring the Overall Burden of Malaria*

Knowing the relative survival of carriers and non-carriers does not tell us the overall effect of malaria, for two reasons. First, in addition to deaths from malaria, we must take into account the cost of sickle cell disease itself. This issue is easily addressed by looking at the fraction of adults who are carriers, from which we can derive the fraction of children who will suffer from sickle cell disease. The underlying data on the number of carriers are the same data used to estimate β.

The second reason that knowing β (relative survival) does not tell us the overall burden of malaria is because a given value of β could be consistent with different levels of absolute survival. For example, $\beta = 0.80$ could be consistent with $M^{AA} = 0.20$ and $M^{AS} = 0$ but it could also be consistent with $M^{AA} = 0.60$ and $M^{AS} = 0.50$. Dealing with this issue requires bringing to bear additional data. Specifically, we need an additional piece of information on M^{AA}, M^{AS}, or their ratio. One can look at modern populations for some information, with the caveat that modern data on survival is not necessarily informative about survival in Africa prior to European contact, where both the disease environment and the level of medical care differed from today.

Allison (2002, table 2) reports results from an examination of 104 child malaria deaths from different countries in Africa, in which the weighted average prevalence of the sickle cell trait was 21%. Only one child examined had the sickle cell trait, which would suggest that the trait is almost completely protective against malaria death.

Table 1

Components of the Cost of Malaria

Group	Fraction of births	Death rate from malaria or sickle cell disease	Fraction of all children who die in category
Non-carriers (AA)	$[1 - (\pi/2)]^2$	M^{AA}	$[1 - (\pi/2)]^2 M^{AA}$
Carriers (AS)	$\pi[1 - (\pi/2)]$	M^{AS}	$\pi[1 - (\pi/2)]M^{AS}$
Sickle cell disease (SS)	$(\pi/2)^2$	1	$(\pi/2)^2$

However, a different set of investigations (Allison, 2002, table 1) that looked at severe P. falciparum infections rather than deaths found a relative incidence of infections in AS that was 46% as high as that for AA. Both sets of studies just described were conducted in the 1950s or very early 1960s. A larger and more recent study (Hill *et al.*, 1991) examined children in The Gambia. Children who were severely ill with malaria were compared to a control group. The severely ill children had cerebral malaria or severe malarial anaemia. Without treatment, most of the children in this group would have died. Among the severe malaria group, the frequency of the AS genotype was 1.2%, while among the control group it was 12.9%. This implies that the relative risk of developing severe symptoms (and presumably dying without medical care) in AS as compared to AA is 0.08.[14] A third study (Williams *et al.*, 2005) concluded that AS was 90% protective against severe or complicated malaria. These studies suggest that reasonable bounds on M^{AS} are zero on the low end, and to assume that $M^{AS}/M^{AA} = 0.08$ on the upper end.[15]

With estimates of M^{AS} and M^{AA}, we are in a position to look at the overall costs of malaria. There are three components to this cost: deaths from malaria among children who are carriers of the sickle cell trait, death from malaria among children who are not carriers and deaths from sickle cell disease. Table 1 shows the fractions of births that fall into each category, the death rate for each group and the total fraction of child deaths (from malaria or sickle cell disease) that are due to each category. The overall fraction of children who die due to malaria and sickle cell disease is simply the sum of the three terms in the right hand column.

Figure 3 does a more extensive analysis, considering different values of π, the prevalence of the sickle cell trait among the adult population. We consider values ranging from zero to 40%, which is the highest level observed among specific populations. For each value of π, we calculate the implied value of β, assuming that the prevalence represents a steady state (7). We also show the fraction of newborn children who will die of sickle cell disease and the malaria death rates for non-carriers

[14] Another study (Greenwood *et al.*, 1991) examined children in Kenya, finding that the sickle cell trait was present in only 1.8% of children with severe malaria anaemia but 3.9% of children with uncomplicated malaria. This finding confirms that the trait is more protective against severe malaria than against mild cases. However, because no data are given on the prevalence of the trait in the overall population, one cannot back out the relative risk of AA *versus* AS.

[15] Given a ratio of malaria mortality in the two groups, along with the equation for β, we can solve the for the group rates of malaria mortality. These are $M^{AA} = [(1 - \beta)/(1 - \beta X)]$ and $M^{AS} = [X(1 - \beta)/(1 - \beta X)]$, where X is the ratio of M^{AS} to M^{AA}.

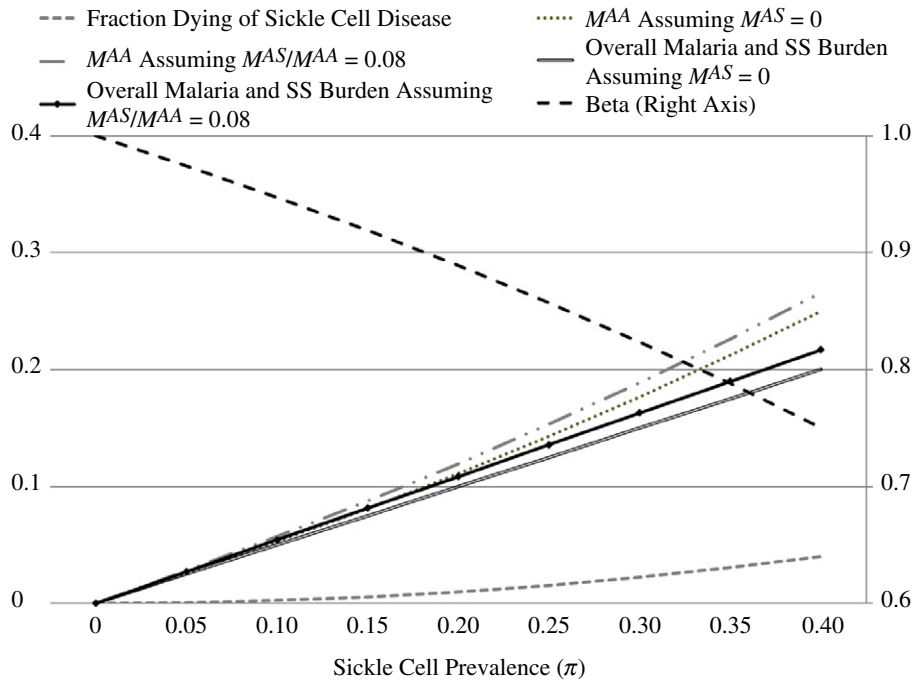

Fig. 3. *Implications of Varying Sickle Cell Prevalence for Malaria Burden*

under the two different assumptions about the death rate for carriers discussed above, specifically that $M^{AS} = 0$ and that $M^{AS}/M^{AA} = 0.08$. Finally, we show the overall burden of malaria and sickle cell disease, once again for the two assumptions about the death rate for carriers discussed above. The Figure shows that, at least within the range of the two estimates we have available, the assumption regarding the degree of protection afforded to carriers of the sickle cell mutation is not very important.

The fraction of the overall burden that takes the form of sickle cell disease rises with prevalence. For example, when $\pi = 0.20$, roughly one-tenth of the overall burden is in the form of deaths from sickle cell disease, with the other nine tenths due to malaria cases. When $\pi = 0.40$, sickle cell deaths account for roughly 20% of the burden. It is also of interest to calculate the net benefit of the sickle cell mutation, that is, the level of mortality in the presence of the mutation relative to the case where it is absent. The level of mortality absent the mutation can simply be read from the line representing M^{AA}, since in this case everyone would have the mortality rate of non-carriers. For example, in the case of 20% prevalence of the sickle cell trait, overall mortality due to malaria and sickle cell disease is 10% but overall mortality due to malaria would be 11.1% if there were no sickle cell mutation. In areas of high malaria pressure, the value of the sickle cell mutation was higher. In the worst afflicted areas, the sickle cell trait reduced malaria related mortality by 20%.

We take sum of malaria and sickle cell deaths as our measure of the overall mortality impact of the disease, which we call malaria burden.[16] In Section 4 we explore the empirical relationship between the overall malaria burden and different measures of economic development at the ethnic group level.

2.4. *Comparison of Malaria Burden to Modern Malaria Mortality Rates*

For the WHO AFRO region, the under-five death rate from malaria is 0.59% per year. Multiplying this number by five gives an approximation to the probability of dying from malaria in the first five years of life, which is very close to the probability of dying from malaria before reproductive age.

The fraction of children who die of malaria is not exactly comparable to what we estimate in the historical data, because our measure M is the probability of dying of malaria conditional on not dying of something else. Such a measure is not usually examined in modern data but it can be constructed relatively easily. Consider the case of Nigeria, which is a very heavily afflicted country. The life table for Nigeria for 2006 shows that the probability of woman surviving to age 25 is approximately 0.75. Taking this age to be the equivalent of adulthood in our historical data, we have the following:

$$0.75 = (1 - M)(1 - P). \tag{8}$$

Annual malaria deaths for children under five in Nigeria are estimated to be 8.8 per thousand, or 0.88%. This implies that roughly 4.4% of children will die of malaria before their fifth birthday. We assume that there are no further malaria deaths beyond this age. To incorporate these data into an estimate, we need to deal with the ambiguity of timing in the simple demographic model with which we started. Specifically, the model implies that a fraction $(1 - M)(1 - P)$ of children will survive both malaria and

[16] Our historical malaria burden measure is a potential alternative to the widely-used malaria ecology index of Kiszewski *et al.* (2004). That index takes into account both climatic factors and the dominant vector species to give an overall measure of how congenial the environment is to the spread of malaria. The index is calculated for grid squares of 0.5 degree longitude by 0.5 degree latitude. With the exception of New Guinea and some areas of southeast Asian, Africa is the only part of the world in which the index reaches its highest levels.

We compared the predictive power of these two measures for two different indicators of the malaria intensity in Africa (outside of Africa we do not view our measure as a viable alternative to the malaria ecology index because the sickle cell mutation is not present among many indigenous populations). The indicators were a dummy variable for having a highly malarious environment c.1900, based on Lysenko and Semashko's (1968) map and average plasmodium falciparum transmission intensity in the year 2007 using data from Hay *et al.* (2009). We used data at two different levels of aggregation: first, grid squares of 0.5 degree longitude by 0.5 degree latitude, and second, ethnic groups mapped by Murdock (1959). In all cases, the malaria burden measure had significantly greater explanatory power, as measured by R^2 by standardised regression coefficients, although when both measures are entered into a regression, they both remain signficant (results available upon request).

An additional advantage of the malaria burden measure that is worth mentioning is that its quantitative interpretation is very simple: it represents the fraction of children expected to die from malaria or sickle cell disease, conditional on not dying of something else. By contrast, malaria ecology is an index of the stability of transmission of the malaria parasite and does not have a simple interpretation in terms of the human effect of the disease. Using it to assess the importance of malaria for health or economic outcomes requires scaling the malaria ecology index in a regression framework, as in Sachs (2003).

other conditions but it is less clear about what those who do not survive die of.[17] If malaria mortality comes before that from other conditions, then a fraction M will die of malaria and a fraction $(1 - M)P$ will die of other causes; if other conditions come first, then a fraction P will die of other causes and $(1-P)M$ will die of malaria. The truth is obviously somewhere in the middle and, for lack of any firm data, we simply assume that deaths due to malaria and 'other' had equal time profiles, which implies:

$$0.044 = M(1 - P) + \frac{PM^2}{M + P}. \tag{9}$$

Solving for (8) and (9) yields $M = 0.054$ and $P = 0.207$. To compare the malaria burden, we can look at the fraction of the Nigerian population who are carriers of the sickle cell trait today. In the data of Piel *et al.* (2010), 9.8% of adult alleles are of the S type (this is the average over the area Nigeria, weighted by current population density), implying that roughly 19.6% of adults are carriers.[18] This in turn implies that slightly less than 10% of children would have died of malaria before reproductive age, conditional on not dying of something else. Thus the burden of malaria has fallen by a little less than half.[19]

3. Assessing the Importance of Malaria to Early African Development

We turn now to an statistical analysis on the importance of malaria to early African development. Our data are at the level of ethnic groups. Analysing data in this form, rather than countries or sub-national units, has two advantages. First, we expect ethnic groups to be less heterogeneous in terms of their levels of early development than political units, which are usually larger and were often created with little attention to realities on the ground. Second, the ethnic group data that we have allows us to get a picture of development further back in time than available data from political units.

In our first econometric exercise we focus on population density for 398 ethnic groups from Alsan (2015).[20] We also exploit a variety of information from Murdock's (1967)

[17] The fraction who die is $M + P - MP$. This can be rewritten as $M(1 - P) + P(1 - M) + MP$, where the first term is children who died of malaria but would not have died otherwise, the second term is children who died of something else but would not have died of malaria and the third term is children who died of one but would have died of the other.

[18] Luzzatto (2012) presents very similar data for Nigeria, also derived ultimately from Piel *et al.* (2010).

[19] To compare the decline in malaria mortality (M) to the decline in mortality from other causes (P) requires additional information on life expectancy. Weil (2014) examines evidence on historical survival in Africa and concludes that 35 years is a reasonable estimate. The United Nations, Department of Economic and Social Affairs, Population Division (1982) model life table with life expectancy of 35 (general model, for females) implies 57.2% of girls will survive to age 25, which we can take as an average for childbearing. Thus, $0.572 = (1 - M)(1 - P)$. Using a value of $M = 0.10$ (consistent with the prevalence of the S allele in Nigeria, as discussed above) implies a value of $P = 0.364$. Thus our estimates of the decline in malaria mortality (from $M = 0.10$ to $M = 0.054$, a total of 44%) and non-malaria mortality (from $P = 0.364$ to $P = 0.207$, a total of 43%) are almost exactly the same. However, this calculation is very sensitive to the choice of life expectancy in the historical period.

[20] We thank Marcella Alsan for providing population density data. Alsan (2015) exploits population figures and total land area for each ethnic group, as implied by the shape-file from the digitalised version of the ethnicity map in Murdock (1959).

ethnographic atlas, a database on 1,167 ethnic groups of six different regions of the world, including Sub-Saharan Africa. This database includes information from the level of subsistence economy to the degree of political integration of each ethnic group.[21] In order to account for the potential confounding effects of a range of geographical and climatical factors, we also construct different geographic and climatic variables by combining georeferenced data with a digitalised version of Murdock's (1959) ethnic maps.[22]

In order to assess the effect of malaria environment on our measures of early development we estimate different specifications of the following equation:[23]

$$y_i = \alpha + \beta M_i + \boldsymbol{\gamma}' \boldsymbol{G}_i + \boldsymbol{\delta}' \boldsymbol{W}_i + \boldsymbol{\rho}' \boldsymbol{T}_i + \boldsymbol{\lambda}' \boldsymbol{X}_i + \mu_C + \epsilon_i. \quad (10)$$

Where the subscript i denotes ethnic group, y is a measure of early development, and M is our measure of historical malaria burden.[24] The vectors $\boldsymbol{G}$, $\boldsymbol{W}$, $\boldsymbol{T}$, and $\boldsymbol{X}$ include geographic, access to waterways, climatic and other plausibly exogenous controls, respectively. μ_C is a collection of 41 ethnographic cultural cluster dummies from Murdock (1967). We describe these controls in detail below. Finally, ε is the error term that is allowed to be heteroscedastic.

3.1. *OLS Estimates*

3.1.1. *The effect of malaria on population density circa 1900*

Table 2 presents the first statistical results where population density is our dependent variable. We focus on population density following the usual Malthusian reasoning that, holding land quality constant, steady state population density is a good indicator of the state of development. In order to ease the interpretation of the results and compare them across the different specifications, we report standardised coefficients.[25]

In column 1 we only control for a set of geographic controls including land quality, elevation and ecological diversity. Land quality is the mean value of the index of land suitability for agriculture from Ramankutty *et al.* (2002) and represents the probability that a particular grid cell (of the size of about 50 × 50 kilometres) may be cultivated.

[21] There is not a perfect match between ethnicity names in Murdock's (1967) Ethnographic atlas, from which we take data on population and other characteristics, and Murdock's (1959) map, which gives outlines of ethnic homelands. We follow Fenske (2013) to match a total of 520 ethnic groups. This leaves more than 300 groups that appear on the map unmatched, potentially leading to bias in our estimates. Although we cannot check if there are systematic differences between matched and unmatched groups in terms of population and other characteristics, we can do so in the case of malaria. Table A1 presents means and standard deviations for our measure of historical malaria burden and malaria ecology across four different samples: the two samples exploited in our analysis that follows (Tables 2 and 3) and the sets of homelands that appear on the map but are not in these samples. As the Table shows, the groups look very similar, with the excluded groups having slightly lower malaria burden.

[22] The main datasets and program that allow replication of all the results in the article are available in the Journal's website.

[23] We use our malaria burden measure for the case in which AS relative malaria mortality of 0.08 in all the empirical analysis in this article. Using a different assumptions regarding the relative protection of AS genotypes does not affect our statistical results.

[24] Because we are using relatively recent data on the prevalence of the sickle cell trait to estimate the historical burden of malaria, there is a danger that migration will introduce measurement error in this variable. Below we use an IV approach to address this and other forms of measurement error.

[25] The standardised coefficients report the number of standard deviation changes in the dependent variable for a one-standard deviation change in the independent variable.

Table 2

OLS Estimates: Historical Malaria Burden and Population Density

	Dependent variable: log of population density (Alsan, 2015)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Malaria burden	0.236***	0.247***	0.228***	0.182***	0.179***	0.185***	
	(0.053)	(0.049)	(0.051)	(0.069)	(0.069)	(0.07)	
Highly malarious environment c.1900							0.329***
							(0.082)
Geographic controls	Y	Y	Y	Y	Y	Y	Y
Waterways controls	N	Y	Y	Y	Y	Y	Y
Climate controls	N	N	Y	Y	Y	Y	Y
Cultural cluster dummies	N	N	N	Y	Y	Y	Y
Suitability slash and burn agriculture	N	N	N	N	Y	Y	Y
Tsetse suitability	N	N	N	N	N	Y	Y
Observations	398	398	398	398	398	398	398
R^2	0.168	0.234	0.238	0.499	0.499	0.500	0.522

Notes. Standardised beta coefficients. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Geographic controls are mean value of the index of land suitability for agriculture (Ramankutty *et al.*, 2002), ecological diversity (Fenske, 2014) and mean elevation of terrain. Waterways controls are the log total area of water bodies accessible for the ethnic and the log of the geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline. Climate controls are mean and standard deviation of temperature calculated for each ethnic group homeland over the time period 1901–2012 (using CRU TS3.21 data set). Cultural cluster dummies account for 41 cultural provinces in the ethnographic atlas. Suitability for Slash and Burn Agriculture is the mean value of the suitability index for cultivation of yams and bananas (FAO's Global Agro-Ecological Zones project -GAEZ- v3.0). Tsetes suitability is an index from Alsan (2015).

To account to the fact that malaria is more prevalent in less ecologically diverse areas which also tend to be less densely populated, we include a measure of ecological diversity from Fenske (2014).[26] Finally, we also include mean elevation of terrain to account for the fact that elevation has a strong negative effect on malaria prevalence and may also have an independent effect on development.[27] Interestingly, the coefficient estimate for β suggests a strong positive statistically significant relationship between our measure of malaria burden and population density. This striking positive association is also documented in other works based on modern data at the country level (such as in Gallup *et al.*, 1999).

The presence of water bodies may positively correlate with both population density and malaria prevalence (such as around Lake Victoria). In column 2, we add the vector *W*, which includes the log total area of water bodies accessible for the ethnic group (note also that humidity tends to be high near rivers, facilitating mosquito breeding) and the log of the geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline (the greater the distance to the coast, the

[26] This measure of ecological diversity consists in a Herfindhal index based on the territorial shares of different ecological types within the boundaries of each ethnic group's homeland.

[27] For instance, Gallup *et al.* (1999) show that population density is greater (lower) at high altitudes in the tropics (temperate zones).

lower the population density). Nonetheless, the point estimate remains virtually unaltered (column (2)).

We next include a set of climatic controls (column (3)). We pay special attention to temperature which plays a key role in determining the suitability for the transmission of malaria (Gething *et al.*, 2011). Using monthly temperature data from the CRU TS3.21 data set (National Center for Atmospheric Research Staff, 2014), we calculate mean temperature over the time period 1901–2012 for each high-resolution grid (0.5 $\times$ 0.5 degree). We then compute mean and standard deviation for each ethnic group homeland in Murdock (1959). The introduction of these controls does not affect the previous results. This is not surprising since malaria's temperature dependence is highly non-linear. Therefore, including temperature in levels is likely to be only indirectly linked to the intensity of malaria transmission (Gething *et al.*, 2011). In column (4) we add a set of 41 cultural province dummies. Although the estimated coefficient on malaria is reduced somewhat, it remains positive and strongly statistically significant.

Another potential bias in the estimation of β results from the omission of ethnic-group agricultural practices. Using data for 60 communities in East and West Africa, Weisenfeld (1967) shows a positive correlation between the prevalence of the sickle cell trait and the cultivation of crops associated with swidden agriculture. Swidden agriculture (clearing small areas of forest with slash and burn practices) multiplies the number of breeding places for Anopheles gambiae. Roots and tree fruits are the main crops associated with this agricultural method and the most important ones for the African case are yams and bananas. To account for this potential omitted variable, we add the mean suitability for cultivation of yams and bananas in column (5).[28] There is little effect on our estimate of β. To account for the potential confounding effect of the tsetse fly, we also include Alsan's (2015) tsetse suitability index in column (6). The point estimate remains unaltered. In sum, after controlling for a battery of geographic, waterways, climates, and regional controls, we find that one standard deviation increase in our malaria burden measure is statistically associated with an increase of almost one-fifth of a standard deviation of population density.

Finally, we check that this striking positive association documented so far is not driven by our measure of historical malaria burden. In column (7), we replace our measure by an alternative proxy for the historical severity of malaria. This alternative measure is constructed based on Lysenko and Semashko's (1968) map categorising the world at the beginning of the twentieth century into six classes of malaria transmission intensity in order of severity: no malaria, epidemic, hypoendemic, mesoendemic, hyperendemic and holoendemic. We focus in the two worst malaria environments as highly malarious, and construct a measure which is the fraction of the territory of our unit of analysis in which malaria fell in this range. We again find a positive and strongly significant association between malaria severity and population density.[29]

[28] We use suitability for cultivation of yams and bananas rather than data on actual cultivation of these crops (which is available in the ethnographic atlas) because the latter is likely to be endogenous. Specifically, Weisenfeld (1967) argues that the sickle cell trait could reduced the environmental limitation of malaria, allowing populations to embrace slash-and-burn practices that have high yields per unit labour.

[29] We also experimented with other specifications including region dummies (i.e: West, East, Central, North, and South) and country dummies. The results are not substantially affected.

Table 3

OLS Estimates: Historical Malaria Burden and Other Measures of Prosperity

	Dependent variable			
	> 20 k City in 1850 (1)	Settlement complexity (2)	Centralisation of power (3)	Intensive agriculture (4)
Malaria burden	0.049	0.147*	0.102	−0.079
	(0.054)	(0.077)	(0.073)	(0.067)
Observations	520	482	520	520
R^2	0.190	0.359	0.216	0.352

Notes. Standardised beta coefficients. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All specifications include the full set of controls from Table 2.

3.1.2. *Other measures of development*

Before turning to an instrumental variable approach we repeat the previous exercise to examine the relationship between malaria burden on other measures of ethnic development. Following Nunn and Wanchekon (2011), we exploit a rich source of ethnic information provided by the Murdock (1967) and estimate a new version of (10) where our dependent variable is now a given measure of ethnic development. Table 3 presents the main results. All the specifications include the full set of controls as in column (6) of Table 2. Since we no longer rely on the availability of population figures, the sample size for each specification will depend only on the availability of the ethnographic variable for ethnic groups that can be located in the Murdock's map (our largest sample consists of 520 ethnic groups).

All the dependent variables in Table 3 are binary, with 1 indicating a higher prosperity level (see note in Table) and come from the ethnographic atlas, except for the variable in column (1) which is constructed using Bairoch (1988).[30] We first study the statistical relationship between malaria burden and the probability of at least one city with population above 20,000 people located in the homeland of the ethnic group by the year 1850. We find no statistically significant association between this measure of urbanisation and historical malaria burden. In column (2), we find a statistically significant and positive association between our measure of malaria burden and the complexity of settlement pattern of the ethnic group (i.e: whether the ethnic group settlement pattern is either compact and relatively permanent or complex, as opposed of being fully nomadic or semi sedentary). These results suggest that malaria did not impose important impediments to the creation of complex settlement and urbanisation.

We next look at the statistical association between our malaria burden and pre-colonial political institutions. Following Gennaioli and Rainer (2007), we construct a measure of centralisation of political power that takes the value 1 if the ethnic group has any jurisdictional level transcending the local community. In line with previous

[30] All the results in Table 3 are OLS coefficients. We also estimated probit and ordered probit (for the original categorical variables) and reached qualitatively similar results.

statistical associations, point estimate from column (3) suggests that ethnic groups located in areas with worst malaria burden do not tend to have less sophisticated political institutions beyond the local communities. We next investigate whether malaria burden negatively impacts the ability of an ethnic groups to generate agricultural surpluses. From column (4) in Table 3, malaria burden is not statistically associated with intensive agriculture being the main contributor of the subsistence economy (as opposed to gathering, fishing, hunting, pastoralism or casual and extensive agriculture). In sum, all the standardised coefficients reported in Table 3 suggest an statistically weak and economically small association between our malaria burden measure and different measures of pre-colonial prosperity.

3.2. *Instrumental Variable Approach*

We view the previous analysis as exploratory. In the absence of a source of exogenous variation on historical malaria burden, it is difficult to identify the precise impact of malaria on development in the past with any reasonable degree of certainty. Although, we attempt to account for several confounding factors in our analysis, bias due to measurement error in our malaria burden variable and reverse causality is likely to be present in the point estimates we report above. In this Section, we exploit variation in an temperature-dependent index of suitability for transmission of malaria to instrument for our historical malaria measure. Temperature affects the life cycles of both the plasmodium parasites and mosquito vector and thus it shapes the landscape of malaria.

Gething *et al.* (2011) present a dynamic biological model incorporating the key mechanisms through which temperature affects malaria transmission to create high resolution maps of temperature suitability for malaria transmission. Unlike previous attempts to map malaria environmental suitability, their measure incorporates the effects of continuously changing temperature regimes on vector and parasite populations. It is also important to note that this measure does not depend on the actual distribution of mosquito vectors, which can be influenced by human economic activity (McCann, 2011). In this way, it avoids a potential endogeneity problem with Kiszewski *et al.* (2004) malaria ecology measure.

3.2.1. *First stage and comparison to malaria ecology*

Gething *et al.* (2011) provide two suitability indices, one for plasmodium falciparum and another for plasmodium vivax. The two are highly correlated, so we simply take their average, which we refer to as plasmodium suitability. In Table 4, we look at the ability of plasmodium suitability to predict malaria burden. All the specifications account for the potential direct impact of altitude, proximity to water, agricultural suitability, ecological diversity and cultural cluster dummies.

We first show that plasmodium suitability is a strong linear predictor of malaria burden even when the first and second moments of historical temperatures are accounted for. This is an important check, since our proposed instrument is constructed from temperature data and temperature can indirectly impact development through agricultural productivity or other diseases. Therefore, after controlling for the mean and standard deviation of temperature, the variation in the instrument

Table 4

First-stage: Plasmodium Suitabiliy Versus Malaria Ecology

	Dependent variable: historical malaria burden					
	(1)	(2)	(3)	(4)	(5)	(6)
Plasmodium suitability	0.48***	0.41***			0.46***	0.40***
	(0.075)	(0.079)			(0.076)	(0.078)
Malaria ecology			0.18***	0.07	0.08	0.04
			(0.064)	(0.053)	(0.053)	(0.059)
Climate controls	N	Y	N	Y	N	Y
Other controls in Table 2	Y	Y	Y	Y	Y	Y
Observations	520	520	520	520	520	520
R^2	0.744	0.753	0.716	0.735	0.745	0.754

Notes. Standardised beta coefficients. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All specifications include the full set of controls from Table 2.

that we are actually exploiting comes from non-linear effects of the time pattern of temperature variation on mosquito survival. The highly statistically significant point estimate in column (1) suggests that one standard deviation increase in the plasmodium suitability index predicts almost half of a standard deviation increase malaria burden. As expected, accounting for the climate controls in column (2) reduces the size of the point estimate but the standardised beta coefficient of 0.41 remains highly significant.

We next run the same exercise, using malaria ecology as main explanatory variable for malaria burden. When none of the climate controls is added (column (3) of Table 4), malaria ecology is a strong linear predictor of malaria burden. However, its predictive power as measured by the standardised coefficient is one fifth the size of coefficient for the plasmodium suitability index. When the climate controls are added to the specification in column (4), we find no statistically significant association between malaria ecology and malaria burden.

Finally, we run horse races between our proposed instrument and malaria ecology and find that, regardless of the addition of climate controls, the plasmodium suitability index outperforms malaria ecology as a predictor for malaria burden.[31]

3.2.2. *Instrumental variable estimates*

Table 5 presents IV estimates for our preferred specifications from Tables 2 and 3 (that is, including all the aforementioned controls). For the specifications where the outcome variables are either population density or indicators for the existence of a large city, centralised political power, and intensity of agriculture, the signs of the point estimates are reversed but they remain statistically insignificant. Therefore, even after

[31] To test whether this same pattern hold using a more recent measure of malaria's impact, we ran similar regressions using the average plasmodium falciporium transmission intensity (parasite rate) in 2007 from Hay *et al.* (2009) as a dependent variable. The results were very similar to those in Table 4: the plasmodium suitability index outperforms malaria ecology as a linear predictor of malaria endemicity, and the predictive power of malaria ecology is statistically weaker when controlling for the mean and standard deviation of temperature.

Table 5

IV Estimates: Historical Malaria Burden and Pre-colonial Development

	Dependent variable				
	Log of population density (1)	>20 k City in 1850 (2)	Settlement complexity (3)	Centralisation of power (4)	Intensive agriculture (5)
Malaria burden	−0.059	−0.597	0.208	−0.180	−0.368
	(0.354)	(0.377)	(0.259)	(0.283)	(0.278)
First Stage F-Statistic	21.809	26.893	24.372	26.893	26.893
Observations	398	520	482	520	520

Notes. Standardised beta coefficients. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All specifications include the full set of controls from Table 2.

instrumenting our measure of historical malaria burden with plausible exogenous variation in malaria suitability we do not find strong evidence that the impact of malaria would have been very significant on early African development. Taking the case of population density as an example, the IV point estimate – albeit highly statistically insignificant – suggests that one standard deviation increase in malaria burden is associated with a small effect on population density: a reduction of only 6% of its standard deviation.

In Table 6, we provide additional robustness checks. Each cell in Table 6 reports the coefficient on the historical malaria burden from a separate IV regression depending on the outcome and the specification we ran. Below the standard errors, the first stage F-statistics are reported in brackets. The outcome variable of interest is listed at the left, while the information on the econometric specification is noted at the bottom.

Some of our early development indicators may be influenced by external factors. If some ethnic groups located in places where the disease environment and other geographic factors handicapped development were also the most affected by the colonial rule, then not taking into account cross-ethnicity variation in colonial power might lead to an important bias in the estimates of the effect of malaria on development. For the specifications in column (1), we add three variables accounting for external influence in the development of the ethnic group. In particular, we account for the importance of slave trades for the period 1400–1900 (i.e. log of 1 + total slaves exports normalised by area of homeland) and two dummy variables indicating if European explorers travelled (between 1768 and 1894), or a historical trade route passed through, the homeland of the ethnic group. Although, we acknowledge a potential endogeneity of these controls, their addition does not seem to affect previous results.

Studies in molecular genetics found that the S gene is associated with at least four different beta-globin haplotypes in Africa, which differ in at least three different genetic markers, providing evidence of the multicentric origin of the mutation. We construct the distance from each ethnic homeland to the closest hypothesised origin of the mutation. Although this measure is strongly correlated with malaria burden, its inclusion as a control in column (2) does not affect the previous results. In column (3),

Table 6

Further Robustness Checks (IV Estimates)

Dependent variable	(1)	(2)	(3)	(4)	(5)
Log of population density	−0.083	−0.067	−0.006	0.001	−0.361
	(0.362)	(0.479)	(0.369)	(0.515)	(0.472)
	[18.9]	[11.04]	[20.23]	[9.1]	[13.72]
>20 k City in 1850	−0.519	−0.960	−0.558	−0.756	−0.859*
	(0.399)	(0.595)	(0.394)	(0.608)	(0.468)
	[23]	[14.9]	[25.5]	[12.83]	[17.83]
Settlement complexity	0.134	0.289	0.193	0.157	0.263
	(0.282)	(0.375)	(0.269)	(0.403)	(0.298)
	[20.5]	[15.6]	[22.64]	[13.46]	[17.13]
Centralisation of power	−0.130	−0.155	−0.144	−0.029	−0.169
	(0.307)	(0.433)	(0.289)	(0.464)	(0.325)
	[23.04]	[14.9]	[25.5]	[12.83]	[17.83]
Intensive agriculture	−0.439	−0.508	−0.359	−0.584	−0.307
	(0.302)	(5.038)	(0.287)	(0.471)	(0.319)
	[23.04]	[14.9]	[25.5]	[12.83]	[17.83]
Robustness	External influence controls	Shortest distance to origin of mutation	HbC mutation	All previous controls	Only tropical sample

Notes. Standardised beta coefficients. Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. First stage F-statistics in brackets. External influence controls are the log of (1 + total slaves exports normalised by area of homeland) from Nunn and Wanchekon (2011) and two dummy variables indicating if European explorers travelled (between 1768 and 1894), or a historical trade route passed through, the homeland of the ethnic group. Shortest distance to origin of mutation captures the shortest distance from each ethnic homeland to the closest hypothetised origin of the mutation. HbC Mutation is the average prevalence of the Haemoglobin C mutation (from the Malaria Atlas Project). Tropical sample includes only ethnic group with homelands within the tropics ($N = 483$).

we include the prevalence of the Haemoglobin C mutation as a control and we still find that malaria burden does not impose a negative effect on early development. For the specifications in column (4) we add all the previous controls together and we find no change in our previous results.

For the specifications in column (5), we only consider ethnic groups located in tropical Africa. Again we do not find statistical evidence that malaria burden negatively affected population density, settlement complexity, centralisation of power, or the intensity of agriculture. We do find, however, that for ethnic groups located in tropical Africa, one standard deviation increase in our measure of historical malaria burden is related to a 12% reduction in the probability of having a large city in 1850 (significant at the 10% level).

To summarise, exploiting plausible exogenous variation from a temperature-dependent index of suitability for transmission of malaria we do not find evidence that the impact of malaria would have been very significant on early African development. We do not find any negative statistically significant association between our measure of malaria burden and population density when we focus on ethnic groups. We do not find consistent evidence of any negative effect on other measures of ethnic development and prosperity such as settlement patterns, the degree of subsistence economy and one measure of sophistication of the political institutions.

3.3. *Model-based Estimates of the Economic Burden of Malaria*

In this subsection, we ask how large an effect economic theory suggests we should see. Weil (2014) analyses this issue using a simple demographic model which takes as its starting point life-cycle patterns of consumption and labour input. These are estimated from relatively undeveloped agricultural societies. Weil uses this model to analyse the effect of a change in malaria mortality on the standard of living, which results from a change in population age structure.[32] More specifically, because the disease generally kills people before they have begun working, malaria mortality raises the ratio of effective consumers to producers, thus lowering consumption per capita (for a given birth rate). However, because malaria mortality is concentrated at such young ages, and because the consumption of young is low relative to that of adults, the quantitative effect on needs-adjusted consumption per capita is small. For example, in the baseline calculation, the effect of going from zero to moderately high malaria prevalence (consistent with $\pi = 0.20$), which would lower life expectancy at birth by five years, is a reduction in consumption per capita by only 0.63%.

In ongoing work (available upon request), we have extended this approach by explicitly incorporating a Malthusian mechanism, in which output per capita is affected by population density and fertility is a function of consumption. In this setting, it is not consumption that is affected by malaria in steady state but rather population density (which is also the object that we examine in the regressions above). Once again, the quantitative effect is small: going from zero malaria to a moderately high level would reduce density by 2.8%. This is obviously an extremely small effect for such a large change in mortality.[33]

We can put further this number in context in two different ways. First, Alsan (2015) also examining African data, finds that a one standard deviation change in an index of tsetse fly impact leads to a 45% change in population density in the period prior to European settlement. This effect is much larger than ours. Second, we can compare the effect of malaria to variation in land quality. Ashraf and Galor (2011), using data for 1500, estimate the coefficient in a regression of log population density on the log of a land quality index as 0.587.[34] A reduction in density of the same magnitude as the one we attribute to malaria could thus be induced by a reduction of 0.048 in the log of land quality, which is less than one twenty-fifth of the standard deviation of that measure, once again showing that the effect of malaria on density that we estimate is small.

The above analysis shows that, at least in terms of the extra costs associated with raising children who were subsequently going to die of malaria, the effect of the disease on population density in Africa should have been relatively low.

[32] Weil (2014) uses as its baseline the UN model life table with life expectancy at birth of 35. In the ongoing work discussed below, we extend the analysis to life tables with lower life expectancy.

[33] In a regression of log population density on the level of malaria burden, this variation would produce a regression coefficient of −0.28. This is very close to zero in comparison to the coefficient on malaria burden in regression with log population density as the dependent variable (the first column of Table 5 but in non-standardised form), which is −2.38. Among the robustness checks, we apply to this calculation are a consideration of the costs of gestation and maternal mortality, as well as variations in baseline life expectancy.

[34] Table 2, column (2); the t-statistic is 8.3.

Another possibility is that malaria affected development in Africa through morbidity (ill health) rather than mortality. This in turn could be via two channels: individuals suffering episodes of malaria infection were less productive, or episodes of malaria infection early in life produced lasting decrements to individual productivity (Table A1).[35] Ashraf *et al.* (2008) evaluate both of these channels in the context of modern Africa, finding their economic impact to be small. For example, they calculate that the long-run effect of a complete eradication of malaria in Zambia, which has a relatively typical burden, would result in labour input per worker rising by 1.5% in the long run (this includes the effect of induced increases in schooling). Part of the reason for this relatively small effect is that malaria morbidity is simply not as important as morbidity from other diseases, and this in turn is due to the short average duration of malaria episodes among adult. Malaria accounts for only 5% of total life-years lost to disability, with most of this resulting from cases during the first five years of life (Murray and Lopez, 1996). Unfortunately, no data are available regarding how malaria morbidity historically compared to today and there is no genetic evidence that can be brought to bear as is the case with malaria mortality.

4. Conclusion

Maleney *et al.* (2004, p. 141) write, 'how better to evince the power of the parasite than with a potentially lethal modification of the genetic code as a desperate Darwinian defence against the even more deadly ravages of malaria? Accordingly, it may be expected that a force strong enough to rewrite our DNA will rewrite many of the lives and economies that it touches'. In this article, we try to address this issue directly. That is, we use the extent to which malaria left its mark on the human genome to back out the severity of the disease's impact and then, in turn, we try to assess how large the economic impact of that disease would have been.

In areas of high malaria transmission, 20% of the population carry the sickle cell trait. Our estimate is that this implies that historically between 10% and 11% of children died from malaria or sickle cell disease before reaching adulthood. Such a death rate is roughly twice the current burden of malaria in such regions. Comparing the most affected to least affected areas, malaria may have been responsible for a ten percentage point difference in the probability of surviving to adulthood. In areas of high malaria transmission, our estimate is that life expectancy at birth was reduced by approximately five years. In terms of its burden relative to other causes of mortality, malaria appears to have been perhaps about as important historically as it is today.

Thus, malaria imposed a heavy mortality burden. Did it hold back economic development? We find little reason to believe that it did. Examining the economic burden of malaria mortality in a simple life cycle model suggests that the disease was

[35] In the context of a modern economy, a potentially important channel by which malaria affects economic outcomes is through schooling. Bleakley (2010), Cutler *et al.* (2010) and Lucas (2010) all use national anti-malaria campaigns in the middle of the twentieth century as quasi-experiments in order to study the effect of childhood exposure to the disease on human capital accumulation and adult economic outcomes. Their findings are highly variable, with Cutler *et al.* estimating a small effect and Bleakley a very large one. Lucas, whose findings fall between the other two, estimates that a 10% reduction in malaria incidence raises completed schooling by 0.1 years.

not very important, primarily because the vast majority of deaths that it caused were among the very young, in whom society had invested few resources. This model-based finding corroborates the findings of our statistical examination. Within Africa, areas with higher malaria burden, as evidenced by the prevalence of the sickle cell trait, do not show lower levels of economic development or population density in the colonial era data that we examine.

Appendix A.

Table A1
Malaria Indicators By Sample Used in Regressions

	Population density sample		Ethnographic atlas sample	
	Included ($N = 398$)	Excluded ($N = 481$)	Included ($N = 520$)	Excluded ($N = 356$)
Historical malaria burden	0.056 (0.04)	0.052 (0.04)	0.056 (0.04)	0.05 (0.04)
Malaria ecology index	13.85 (9.8)	12.4 (9.17)	13.47 (9.6)	12.5 (9.34)

Note. Each malaria indicator represents its mean value for the ethnic homeland (SD in parenthesis).

Instituto de Economía e Instituto de Ciencia Política, Pontificia Universidad Católica de Chile
Brown University and NBER

Accepted: 28 July 2016

Additional Supporting Information may be found in the online version of this article:

Data S1.

References

Acemoglu, D., Johnson, S. and Robinson, J. (2001). 'The colonial origins of comparative development: an empirical investigation', *American Economic Review*, vol. 91(5), pp. 1369–401.
Akyeampong, E.K. (2006). 'Disease in West African history', in (E.K. Akyeampong, ed.), *Themes in West Africa's History*, pp. 186–207, Athens, OH: Ohio University Press.
Allison, A.C. (2002). 'The discovery of resistance to malaria of sickle-cell heterozygotes', *Biochemistry and Molecular Biology Education*, vol. 30(5), pp. 279–87.
Alsan, M. (2015). 'The effect of the tsetse fly on African development', *American Economic Review*, vol. 105(1), pp. 382–410.
Ashraf, Q. and Galor, O. (2011). 'Dynamics and stagnation in the Malthusian epoch', *American Economic Review*, vol. 101(5), pp. 2003–41.
Ashraf, Q. and Galor, O. (2013). 'The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development', *American Economic Review*, vol. 103(1), pp. 1–46.
Ashraf, Q., Lester, A. and Weil, D.N. (2008). 'When does improving health raise GDP?', in *NBER Macroeconomics. Volume 23*. Cambridge, MA: NBER.

Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*, Chicago, IL: University of Chicago Press.
Bleakley, H. (2010). 'Malaria eradication in the Americas: a retrospective analysis of childhood exposure', *American Economic Journal: Applied Economics*, vol. 2(2), pp. 1–45.
Chanda, A. and Putterman, L. (2007). 'Early starts, reversals and catch-up in the process of economic development', *Scandinavian Journal of Economics*, vol. 109(2), pp. 387–413.
Chiovelli, G., Cervellati, M. and Esposito, E. (2015) 'Bite and divide: malaria and ethnic diversity', Working Paper, University of Bologna.
Comin, D., Easterly, W. and Gong, E. (2010). 'Was the wealth of nations determined in 1000 BC?', *American Economic Journal Macroeconomics*, vol. 2(3), pp. 65–97.
Cutler, D., Fung, W., Kremer, M., Singhal, M. and Vogl, T. (2010). 'Early-life malaria exposure and adult outcomes: evidence from malaria eradication in India', *American Economic Journal: Applied Economics*, vol. 2(2), pp. 72–94.
Currat, M., Trabucher, G., Rees, D., Perrin, P., Harding, R.M., Clegg, J.B., Langaney, A. and Excoffier, L. (2002). 'Molecular analysis of the β-globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the β Senegal mutation', *American Journal of Human Genetics*, vol. 70, pp. 207–23.
Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies*, New York: W. W. Norton and Co.
Fenske, J. (2013). 'Does land abundance explain African institutions?', Economic Journal, vol. 123(573), pp. 1363–90.
Fenske, J. (2014). 'Ecology, trade, and states in pre-colonial Africa', *Journal of the European Economic Association*, vol. 12(3), pp. 612–40.
Gallup, J.L. and Sachs, J.D. (2001). 'The economic burden of malaria', *Amercian Journal of Tropical Medicine and Hygeine*, vol. 64 (1, Supplement), pp. 85–96.
Gallup, J.L., Sachs, J.D. and Mellinger, A. (1999). 'Geography and economic development', *International Regional Science Review*, vol. 22(2), pp. 179–232.
Gennaioli, N. and Rainer, I. (2007). 'The modern impact of precolonial centralization in Africa', *Journal of Economic Growth*, vol. 12, pp. 185–234.
Gething, P., Van Boeckel, T., Smith, D., Guerra, C., Patil, A., Snow, R. and Hay, S. (2011) 'Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax', *Parasites and Vectors*, vol. 4(1), pp. 92.
Greenwood, B., Marsh, K. and Snow, R. (1991). 'Why do some African children develop severe malaria?', *Parasitology Today*, vol. 7(10), pp. 277–81.
Hay, S., Guerra, C., Gething, P., Patil, A., Tatem, A., Noor, A., Kabaria, C., Manh, B., Elyazar, I., Brooker, S., Smith, D., Moyeed, R. and Snow, R. (2009) 'A world malaria map: Plasmodium falciparum endemicity in 2007', *PLoS Med*, vol. 6(10), e1000048.
Hedrick, P.W. (2011). 'Population genetics of malaria resistance in humans', *Heredity*, vol. 107, pp. 283–304.
Hill, A. Allsopp, C.E.M., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J. and Greenwood, B.M. (1991). 'Common West Africa HLA antigens are associated with protection from severe malaria', *Nature*, vol. 352(6336), pp. 595–600.
Hibbs, D.A. and Olsson, O. (2004). 'Geography, biogeography, and why some countries are rich and others are poor', *Proceedings of the National Academy of Sciences*, vol. 101(10), pp. 3715–20.
Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, J. and Sachs, S.E. (2004). 'A global index of the stability of malaria transmission', *American Journal of Tropical Medicine and Hygiene*, vol. 70(5), pp. 486–98.
Kurnit, D.M. (1979). 'Evolution of sickle variant gene', *Lancet*, vol. 313(8107), pp. 104.
Kwaitkowski, D.P. (2005). 'How malaria has affected the human genome and what human genetics can teach us about malaria', *American Journal of Human Genetics*, vol. 77(2), pp. 171–92.
Lehmann, H. and Raper, A.B. (1956). 'Maintenance of high sickling in an African community', *British Medical Journal*, vol. 2(4988), pp. 333–6.
Livingstone, F.B. (1958). 'Anthropological implications of sickle cell gene distribution in West Africa', *American Anthropologist*, vol. 60, pp. 533–62.
Lucas, A.M. (2010). 'Malaria eradication and educational attainment: evidence from Paraguay and Sri Lanka', *American Economic Journal: Applied Economics*, vol. 2(2), pp. 46–71.
Luzzatto, L. (2012). 'Sickle cell anaemia and malaria', *Mediterranean Journal of Hematology and Infectious Diseases*, vol. 4(1), e2012065.
Lysenko, A.J., Semashko, I.N. (1968). 'Geography of malaria. A medico-geographic profile of an ancient disease', in (Lebedew A.W., ed.) *Itogi Nauki: Medicinskaja Geografija*. Moscow: Academy of Sciences, USSR, pp. 25–146.
Mabogunje, A.L. and Richards, P. (1985). 'The land and peoples of West Africa', in (Ajayi, J.F.A. and M. Crowder, eds.), *History of West Africa*, vol. 1, 3rd edn, pp. 5–47, New York, NY: Longman.
Maleney, P., Spielman, A. and Sachs, J.D. (2004), 'The malaria gap', *American Journal of Tropical Hygiene and Medicine*, vol. 71(2 Suppl), pp. 141–6.
McCann, J. (2011) 'Deposing the malevolent spirit: a historical cultural ecology of malaria in northwest Ethiopia', PSAE research series Issue 9. African Studies Center, Boston University.

McGuire, R.A. and Coelho, P.R.P. (2011). *Parasites, Pathogens, and Progress: Diseases and Economic Development*, Cambridge, MA: MIT Press.
McNeill, W. (1977). *Plagues and Peoples*, New York: Anchor Books.
Morrow, R.H. and Moss, W.J. (2006). 'The epidemiology and control of malaria', in (N. Kennard, and C.M. Williams, eds.), *Infectious Disease Epidemiology: Theory and Practice,* 2nd edn, London: Jones and Bartlett Publishers, pp. 1087–138.
Motulsky, A. (1964). 'Hereditary red cell traits and malaria', *American Journal of Tropical Medicine and Hygiene*, vol. 13(Part 1), pp. 147–58.
Murdock, G.R. (1959). *Africa: Its People and Their History*, NY: McGraw-Hil.
Murdock, G.R. (1967). *Ethnographic Atlas*, Pittsburgh, PA: University of Pittsburgh Press.
Murray, C.J.D. and Lopez, A.D. (1996). *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020*, Cambridge, MA: Harvard University Press.
National Center for Atmospheric Research Staff (eds) (2014). 'The Climate Data Guide: CRU TS3.21 Gridded precipitation and other meteorological variables since 1901.'
National Heart, Lung, and Blood Institute. (1996). 'Facts about sickle cell anemia', NIH Publication No. 96-4057.
Nelson, K. and Williams, C. (2006). *Infectious Disease Epidemiology: Theory and Practice*. 2nd edn, Burlington, MA: Jones and Bartlett Publishers.
Nunn, N. (2008). 'The long-term effects of Africa's slave trades', *Quarterly Journal of Economics*, vol. 123(1), pp. 139–76.
Nunn, N. and Wanchekon, L. (2011). 'The slave trade and the origins of mistrust in Africa: dataset', *American Economic Review*, vol. 101(7), pp. 3221–52.
Piel, F.B., Patil, A.P., Howes, R.E., Nyangiri, O.A., Gething, P.W., Williams, T.N., Weatherall, D.J. and Hay, S.I. (2010). 'Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis', *Nature Communications*, vol. 1, (November 2), p. 104. doi: 10.1038/ncomms1104.
Putterman, L. and Weil, D.N. (2010). 'Post-1500 population flows and the long-run determinants of economic growth and inequality', *The Quarterly Journal of Economics*, vol. 125(4), pp. 1627–82.
Ramankutty, N., Foley, J.A., Norman, J. and McSweeney, K. (2002). 'The global distribution of cultivable lands: current patterns and sensitivity to possible climate change', *Global Ecology and Biogeography*, vol. 11, pp. 377–92.
Sachs, J.D. (2003). 'Institutions don't rule: direct effects of geography on per capita income', NBER Working Paper 9490, National Bureau of Economic Research Inc.
Soloman E. and Bodmer, W.F. (1979). 'Evolution of sickle variant gene', *Lancet*, vol. 1(8122), p. 923.
Spolaore, E. and Wacziarg, R. (2013). 'How deep are the roots of economic development?', *Journal of Economic Literature, American Economic Association*, vol. 51(2), pp. 325–69.
United Nations, Department of Economic and Social Affairs, Population Division (1982). 'Model Life Tables for Developing Countries', Geneva: United Nations.
Webb, J.L.A. (2006). 'Ecology and culture in West Africa', in (E.K. Akyeampong, ed.), *Themes in West Africa's History*, pp. 33–51, Athens, OH: Ohio University Press.
Weil, D.N. (2010). 'Endemic diseases and African economic growth: challenges and policy responses', *Journal of African Economies*, vol. 19(Suppl 3), pp. iii81–109.
Weil, D.N. (2014). 'The impact of malaria on African development over the longue duree', in (E. Akyeampong, R. Bates, N. Nunn, and J.A. Robinson eds.), *Africa's Development in Historical Perspective*, pp. 89–130, Cambridge: Cambridge University Press.
Weisenfeld, S.L. (1967). 'Sickle cell trait in human biological and cultural evolution', *Science*, vol. 157(3793), pp. 1135–40.
Williams, T.N., Mwangi, T.W., Wambua, S., Alexander, N.D., Kortok, M., Snow, R.W. and Marsh, K. (2005). 'Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases', *Journal of Infectious Diseases*, vol. 192(1), pp. 178–86.
World Health Organization (2014). World Malaria Report, Geneva: WHO.