International Center for Public Policy
Working Paper 21-01
January 2021

The Role of Historical Malaria in Institutions and Contemporary Economic Development

Elizabeth Gooch
Jorge Martinez-Vazquez
Bauyrzhan Yedgenov

Georgia State University | Andrew Young School of Policy Studies

International Center for Public Policy
Working Paper 21-01

The Role of Historical Malaria in Institutions and Contemporary Economic Development

Elizabeth Gooch
Jorge Martinez-Vazquez
Bauyrzhan Yedgenov
January
2021

International Center for Public Policy
Andrew Young School of Policy Studies
Georgia State University
Atlanta, Georgia 30303
United States of America

Phone: (404) 413-0235
Fax: (404) 651-4449
Email: paulbenson@gsu.edu
Website: http://icepp.gsu.edu/

International Center for Public Policy
Andrew Young School of Policy Studies

The Andrew Young School of Policy Studies was established at Georgia State University with the objective of promoting excellence in the design, implementation, and evaluation of public policy. In addition to two academic departments (economics and public administration), the Andrew Young School houses seven leading research centers and policy programs, including the International Center for Public Policy.

The mission of the International Center for Public Policy (ICePP) at the Andrew Young School of Policy Studies is to provide academic and professional training, applied research, and technical assistance in support of sound public policy and sustainable economic growth in developing and transitional economies.

ICePP is recognized worldwide for its efforts in support of economic and public policy reforms through technical assistance and training around the world. This reputation has been built serving a diverse client base, including the World Bank, the U.S. Agency for International Development (USAID), the United Nations Development Programme (UNDP), finance ministries, government organizations, legislative bodies, and private sector institutions.

The success of ICePP reflects the breadth and depth of its in-house technical expertise. The Andrew Young School's faculty are leading experts in economics and public policy and have authored books, published in major academic and technical journals, and have extensive experience in designing and implementing technical assistance and training programs. Andrew Young School faculty have been active in policy reform in over 40 countries around the world. Our technical assistance strategy is not merely to provide technical prescriptions for policy reform, but to engage in a collaborative effort with host governments and donor agencies to identify and analyze the issues at hand, arrive at policy solutions, and implement reforms.

ICePP specializes in four broad policy areas:

- Fiscal policy (e.g., tax reforms, public expenditure reviews)
- Fiscal decentralization (e.g., reform, intergovernmental transfer systems, urban finance)
- Budgeting and fiscal management (e.g., local, performance-based, capital, and multiyear budgeting)
- Economic analysis and revenue forecasting (e.g., micro-simulation, time series forecasting)

For more information about our technical assistance activities and training programs, please visit our website at http://icepp.gsu.edu or contact us by email at paulbenson@gsu.edu.

The Role of Historical Malaria in Institutions and Contemporary Economic Development

Elizabeth Gooch[1], Jorge Martinez-Vazquez[2], and Bauyrzhan Yedgenov[3]

January 2021

Abstract

This research examines the causal impact of institutional quality on economic development from a novel perspective. At the country level, we exploit variation in the malaria prevalence in 1900, just before vector-control methods were developed, to instrument for institutional quality using a two-stage least squares instrumental variables framework. Our instrument is a population-weighted average of malaria endemicity estimates for the year 1900 developed by the WHO in the 1960s. We argue that this measure of historical malaria offers more expansive geographic information about the disease environment than other metrics, and our baseline IV estimates reveal that greater institutional quality causes greater contemporaneous economic growth. Next, we investigate the robustness of these baseline results to alternative explanations, including the role of geography and early colonizers' experiences, as the causal link between the early disease environmental, institutional quality and contemporary growth. As an additional test of the explanatory power of malaria endemicity, we replace our instrument for settler mortality and replicate the core results from the seminal study on the colonial origins of comparative development by Acemoglu et al. (2001). In summary, we propose that malaria endemicity, estimated for 1900, holistically explains the legacy of early disease on institutional quality development and contemporary economic development.

Keywords: Malaria Endemicity, Institutions, Economic Development, European Colonization
JEL: O11, O43, N10, O57

[1] Gooch: Naval Postgraduate School, 699 Dyer Road, Room M8, Monterey, CA 93943 (elizabeth.gooch@nps.edu, phone: 831-656-3669)
[2] Martinez-Vazquez: International Center for Public Policy (ICePP), Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992 (jorgemartinez@gsu.edu, fax: 404-413-0244)
[3] Yedgenov (corresponding author): International Center for Public Policy (ICePP), Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992 (byedgenov1@gsu.edu, fax: 404-413-0238)

1. Introduction

There has been much debate over the determinants of modern economic development, especially between the competing hypotheses emphasizing the role of institutions versus the role of geography. Even though this debate has currently subsided, there are still many important questions that remain unanswered beyond the respective roles of institutions versus geography, especially those concerning the role played by early disease environments on early institutions' development and the validity of the measures used as a proxy for institutional quality.

In this paper, we contribute to ongoing debate by: (1) introducing a new country-level measure of the early disease environment; (2) explaining holistically the impact of early disease on institutional quality and economic development; and (3) estimating the causal impact of institutional quality on economic development driven by variation in early disease. Specifically, we use malaria endemicity in the year 1900 (henceforth endemicity 1900) as an alternative instrumental variable for institutional development in all countries, including European colonies, non-European colonies and countries that were never colonies. Endemicity 1900 offers geographically expansive information about the disease environment in 1900 by measuring the prevalence of malaria before the discovery that mosquitoes were the malaria transmission vector and, therefore, before the malaria eradication efforts that started in the early twentieth century.

In our analysis, we first examine the relationship between endemicity 1900 and regulatory quality in 2000, considering population and cropland in 1900. In the first stage, we find evidence of a statistically significant relationship between endemicity 1900 and regulatory quality. In the second stage, according to our local average treatment effect (LATE) estimate, regulatory quality has a substantial impact on income in countries where the mosquito-borne disease environment affected regulatory quality. Specifically, we find that a one-standard-deviation increase in regulatory quality more than doubles GDP per capita for a nation in our

sample with mean GDP per capita. Further, we find that, even after accounting for countries' geographic attributes such as temperature, precipitation, latitude, coastline, tropical climate and malaria ecology, endemicity 1900 is a significant determinant of institutional development. Lastly, we find evidence that the average treatment effect in former colonies for which the malaria environment influenced regulatory quality is twice as large as for non-colonies. Economic development in countries that were never colonized is more robust, although it is still significantly linked to the formation of institutional quality determined by early malaria. We also replicate the evidence in Acemoglu, Johnson and Robinson (2001), henceforth AJR (2001), by substituting endemicity 1900 for settler mortality and find evidence that European settlements and mortality were shaped by historical malaria endemicity. While AJR (2001) settler mortality strictly confines the analysis to former European colonies, our new instrumental variable expands the country selection, allowing us to analyze expanded European colonies not in AJR and further extend the analysis to non-colonies. In fact, we find that the formation of institutional quality determined by early malaria is *more* robust for non-colonies.

This paper makes several important contributions to the existing literature. First, we introduce an exogenous measure of historical disease environment—endemicity 1900, which measures the prevalence of malaria in countries prior to the discovery that mosquitoes were the transmission vector for malaria and, therefore, before the eradication efforts of the early twentieth century. Second, by using coverage and variability of endemicity 1900, we can unify under a single holistic framework the two different explanations (institutions versus geography) that have prevailed in the previous literature for how early institutions, and therefore modern economies, developed. We show how using endemicity 1900 in AJR (2001) framework improves on their results. We also show that geography is a key determinant, but not the only one, of

endemicity 1900. Thus, the previous findings in the literature regarding the role of geography can now be linked to the main cause determining the disease environment of countries and why early institution developed. This provides an answer to the question of what mechanism has been at work behind the "geography" explanation.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature. Section 3 presents our instrumental variable, malaria endemicity in the year 1900. In section 4, we present the empirical approach and results, while in section 5 we replicate the core results from AJR using endemicity 1900. Finally, section 6 concludes.

2. Literature Review

In the literature on the determinants of modern economic development, two hypotheses have been dominant over the last few decades. In one strand of the literature, the chief contribution has been by AJR (2001) who use estimates of European settlers' mortality as an instrument for early institutions development in European colonies. While the AJR (2001) approach has been extended by others—for example Auer (2013) added additional covariates such as distance from the equator, average temperature, rainfall, drought, humidity and dummies for climate—the validity of the AJR (2001) instrument has been widely challenged (Albouy, 2008; Albouy, 2012; Sachs, 2012).

In the second strand of the literature, Gallup et al. (1999) and others (Gallup and Sachs, 2001; Sachs and Malaney, 2002; Cartensen and Gundlach, 2006) use geography, in several dimensions, as the main driver of early institutional and economic development, with one of those dimensions being "malaria ecology." However, there has been a lot of discontent over their measure of malaria prevalence, because malaria ecology is not exogenous to modern development, nor does it provide enough variation. In the following sections, we delve into the comparison of

the various measures of malaria prevalence, and we argue the advantages of endemicity 1900 over other measures, such as "malaria ecology," in terms of exogeneity, variation and coverage.

The central argument within the literature about the relationship between disease environments and economic growth is whether the effects of the former on the latter are ongoing and direct or historical and indirect. AJR (2001) argued for an indirect impact of malaria on current economic growth, claiming that the prevalence of malaria is highly endogenous and the contemporary persistence of malaria stems from the poor institutions of some low-income countries that could not eradicate malaria. In addition, AJR (2001) express skepticism over malaria's direct effect on economic performance—as proposed by Gallup et al. (1999)—which they expected it to work through poor health and high mortality rates. AJR (2001) note that most people living in high malaria areas have developed some immunity to the disease, such that if they survive to the age of five and afterwards get sick, most probably it won't be fatal. Therefore, they argue, the effect of malaria has been mainly an indirect one through its effect on settler mortality and the type of institutions established by the settlers, which defined the long-term economic development of countries, including their current performance. In a later paper, AJR (2002) further develop the indirect channel argument for the effect of malaria on economic growth through the type of institutions that were established. The authors argue that, since developed areas before colonization were those that were more urbanized and more densely populated and malaria was more endemic in such areas because of more frequent contacts, Europeans preferred to settle in less dense, and hence less endemic, areas where they established inclusive institutions. AJR (2001; 2002) found that the malaria incidence 1994 variable used by Gallup et al. (1999) was mostly statistically insignificant by itself as an additional control variable. This opened an additional debate over the proper measure of malaria prevalence.

Later, following the criticism of AJR (2001; 2002), Gallup and Sachs (2001) and Sachs and Malaney (2002) used a malaria risk index, which is based on the 1994 world malaria prevalence map by the World Health Organization (WHO). Their main finding was that even after controlling for institutions, a higher risk of malaria negatively affects current income per capita, thus supporting the argument of a direct link. Gallup and Sachs (2001) and Sachs and Malaney (2002) also added that the reason AJR (2001) didn't find a direct effect of malaria is because they restricted their data sample to former colonies, which are mainly in the tropics, therefore leading to low variability of the malaria environments. Similar results to those in Gallup and Sachs (2001) and Sachs and Malaney (2002) were reported by Cartensen and Gundlach (2006). These authors argue that even though population in malaria endemic areas develop immunity through sickle cells, these cells affect the health and human capital of the population through sickle cell anemia, and so they also find an independent effect of malaria on GDP per capita after controlling for institutions. However, the criticism over the exogeneity and variability of their malaria measure stands.

Although the focus of the literature has largely been on European colonies, others have attempted to conduct some comparative analyses of colonies versus non-colonies. Two studies apply to our work. First, Rigobon and Rodrik (2005) study the interrelationships between the rule of law, democracy, and openness with income by splitting a sample of countries into colonies and non-colonies. For the specific case of the impact of rule of law and democracy on income, they do not find a significant difference between the two samples. Second, Auer (2013) similarly compared colonies and non-colonies to estimate the direct versus indirect effect of disease environment on development by replicating AJR results. In particular, he substitutes the settler mortality measure with disease environment, measured using several variables such as malaria ecology from Kiszewski et al. (2004). However, as Auer (2013) notes, the malaria ecology variable is not reliable

because of very low variation among non-colonies. This leads Auer (2013) to use alternative proxies that are correlated with settler mortality such as distance from the equator, average temperature, rainfall, drought, humidity and dummies for climate; however, these proxies are also highly correlated with variables capturing non-disease channels impacting development, and therefore cannot be considered exogenous. Overall, Auer (2013) finds that one quarter of the correlation between the disease environment and income can be attributed to the direct effect of the disease rather than the indirect effect of settler mortality rates on colonization policies, meaning therefore that AJR's results were inflated. He also concludes that the early disease environment is strongly correlated with development in the former colonies, while it is not the case in the sample of non-colonies.

As we see in our empirical analysis below, our findings using endemicity 1900 (and not "malaria ecology") as the instrumental variable is in open contradiction with Auer's (2013) finding regarding non-colonies and non-European colonies. Endemicity 1900 may have affected what European colonizers actually did in the colonies, but we find that endemicity 1900 also explains early institutional development and modern economic development in non-colonies and non-European colonies, which clearly signals a direct role for the disease environment, perhaps working through human capital formation (Bloom et al., 2004; Glaeser et al., 2004) and direct costs, such as forgone income (Sachs and Malaney, 2002). But, at the end, we will also see geography play a role as a key determinant, but not the only one, of endemicity 1900.

3. Malaria Endemicity in the Year 1900

3.1 Description of the variable

Lysenko and Semashko (1968) first published malaria endemicity estimated for the year 1900 as part of a 1968 World Health Organization (WHO) report. Those researchers captured the distribution of malaria in 1900, specifically, because that year coincides with both the initiation

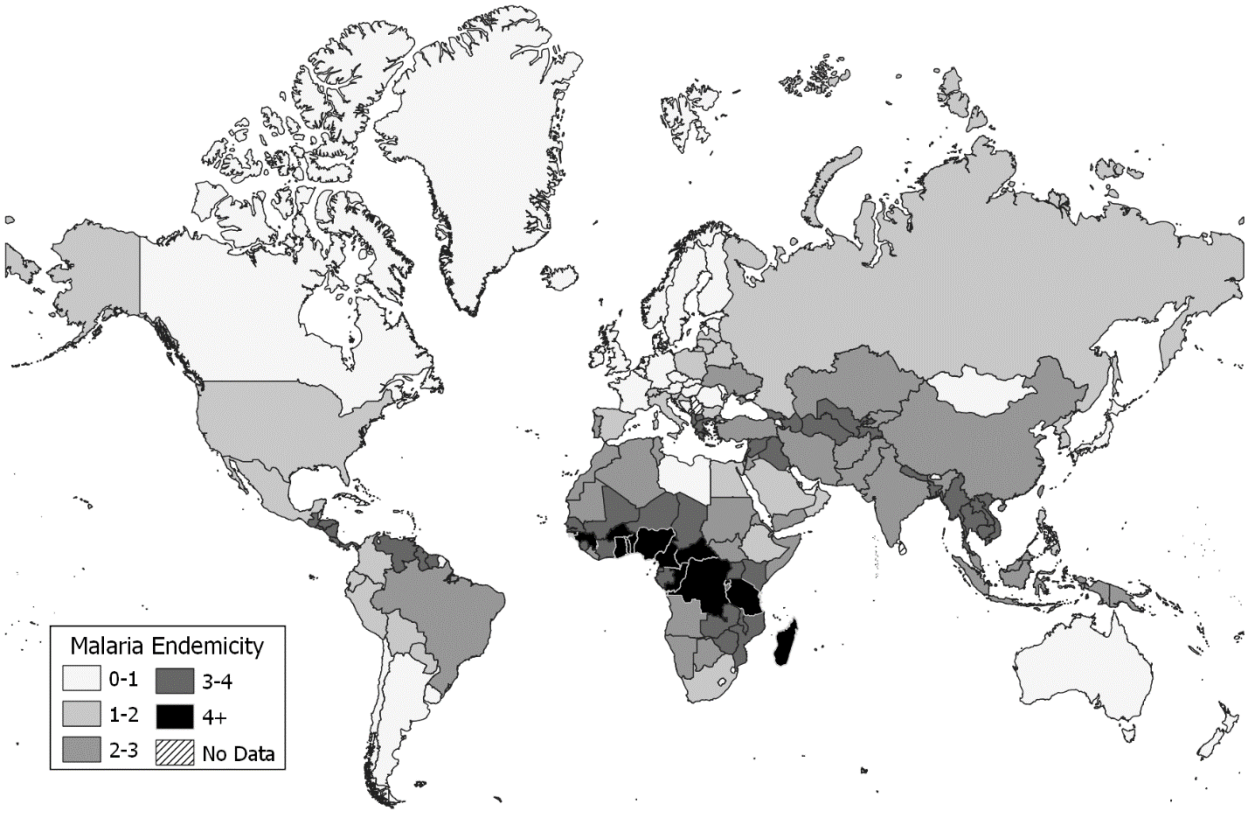
of vector control policies to fight malaria across the world and the largest global population before the vector control campaigns. Lysenko and Semashko (1968) constructed the parasite rate from the interpolation of data from records of disease and vector presence (e.g., spleen rates, parasite rates, sickle cell incidence, sporozoite rates and biting rates) and mapped malaria at the peak of its assumed historical distribution, using a combination of expert opinion and climatic measures such as temperature and rainfall isohyets.

Endemicity 1900 is an ordered variable, delineated by differences in the parasite rate for the 2- to 10-year-old age cohort. The highest endemicity level is holoendemic with $PR > 0.75$; the remaining regions, from high to low, are classified as hyperendemic with $PR \in (0.5,0.75]$, mesoendemic $PR \in (0.1,0.5]$, hypoendemic $PR \leq 0.1$, and epidemic regions, which include places where some malaria existed and malaria-free areas. Contemporary malariologists have revived the index to characterize historical malaria geography and prevalence across countries (Hay et al., 2004).

In this paper, we define endemicity 1900 as the mean malaria endemicity for each country weighted by the estimated population in 1900. Gooch (2017) converted the endemicity map into geospatial data. The GIS dataset is made up of grid cells taking the Harvest Choice Grid Database at the one-degree resolution which is then used to calculate the population-weighted mean endemicity for each country. Figure 1 shows a map of country-level malaria endemicity in the year 1900 (Guo, 2015).[1]

[1] Note that for estimation purposes we remove islands and small countries from the analysis because of the likelihood of inaccuracy due to the size of the raster data and spillovers with the ocean.

Figure 1: Malaria Endemicity Levels and Geographic Extent



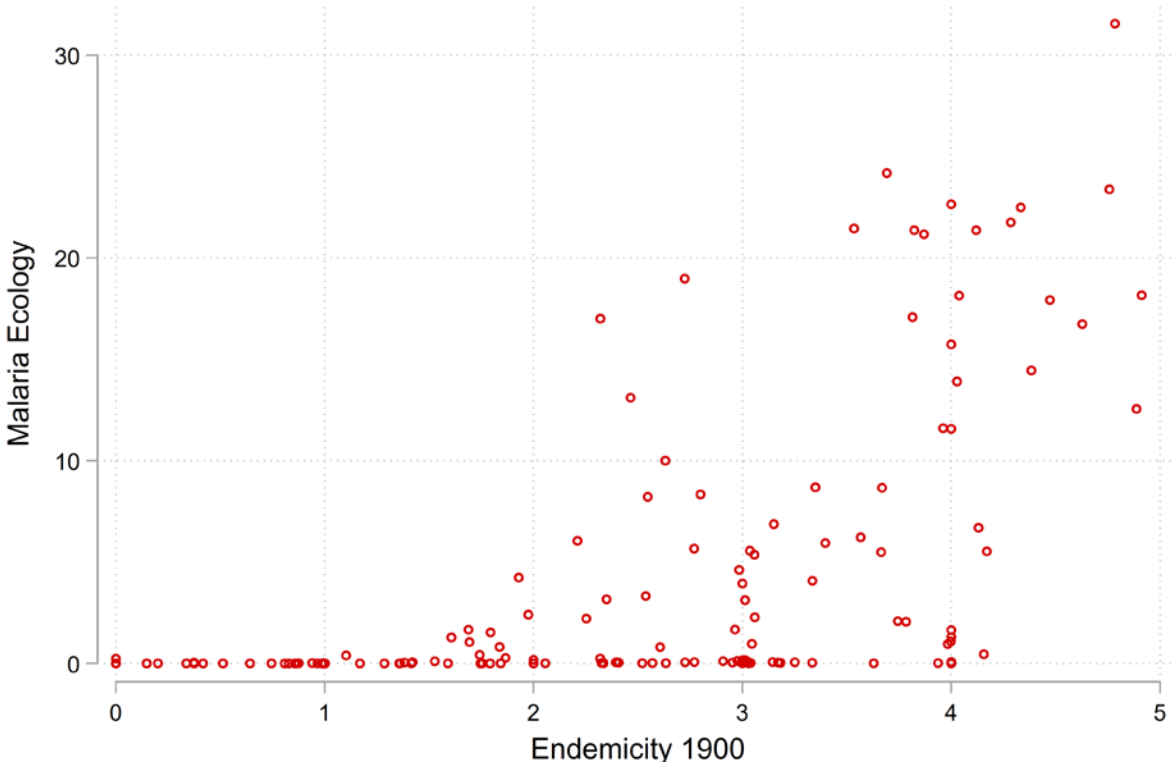
Source: Gooch (2017) and Lysenko et al. (1968)

3.2 Relationship with Other Commonly Used Measures of Malaria Prevalence

Two malaria measures have been prominently used in the literature: the time-invariant malaria ecology index created by Kiszewksi, et al. (2005) and a malaria index for the year 1994 used in Sachs et al. (2001). In the panels of Figure 2, we graph the pair-wise correlation between malaria endemicity 1900 on the x-axis and the y-axis, malaria ecology (Panel A) and malaria index for 1994 (Panel B). A striking feature of the spread of countries across each panel is the incongruence between historical malaria risk (as in endemicity 1900) and what we can observe today about malaria risk using these other common measures. According to endemicity 1900, the prevalence of malaria was widespread across countries and continents, while both malaria ecology and malaria index 1994 mainly fixate on African countries. The stability of malaria in the central part of the African continent partially explains the focus on the contemporary African experience with the disease.

Figure 2: Correlation between Endemicity 1900 and other Measures of Mosquito-borne Disease Environments

(A) Malaria Risk, 1994 (Source: Sach, et al., 2001)



(B) Malaria Ecology (Source: Kiszewski, et al., 2005)

Malaria ecology represents the stability of malaria transmission. Malaria ecology is available at the 0.5-degree grid cell and is based on characteristics of the regionally dominant mosquito species such as a portion of blood meals taken from human hosts, daily survival of the vector, duration of the transmission season and incubation. The stability of malaria and the prevalence of malaria are related but characterize malaria differently. Malaria is stable if it is transmitted throughout the year by long-lived mosquitoes. Conceptually, if malaria is stable, suppression efforts can work, but virulence of the vector can quickly erase the progress. However, if there is a spike in the prevalence of malaria in an unstable environment, the malaria epidemic will be short-lived, and suppression efforts will probably be very effective.

The malaria index for the year 1994 quantifies the risk of infection by one strain of *Plasmodium*, *Plasmodium Falciparum*, which predominates in Africa and contributes to most morbidity and mortality because of malaria. This malaria index is constructed using the fraction

of the population at risk of malaria in 1990, multiplied by the fraction of cases of malaria that were *P. Falciparum* malaria in 1994. Severe malaria is defined as having a malaria index greater than 0.5. The 1994 distribution of malaria reflects the fact that "... poverty and inequality as much as geography and climate make malaria a tropical disease" (Packard, 2007, pp. xi).

Endemicity 1900 does not reflect communities' access to vector control campaigns of the 20th-century and is available for most of the countries in the world; it is better at capturing the variation of mosquito-borne disease (MBD) risk than measures previously used in the empirical literature. Gallup and Sachs (2001) describe the relevance of a historical measure of malaria prevalence regarding the experiences of historical communities compared to the malaria stability captured by the malaria ecology index and contemporary prevalence captured by the risk of infection by *P. Falciparum* in 1994.

> "The long-standing problem of malaria in Italy contributed to the major role of Italians in early malaria research. Just before the control campaign [circa early 1930s], Italy had > 300,000 cases of malaria per year, with ~20,000 deaths. The Pontine Marshes south of Rome were rendered uninhabitable by the disease. *Plasmodium falciparum* was eliminated by the end of the 1940s, with *P. vivax* and *P. malariae* disappearing more slowly. Spain reported 400,000 cases of malaria with 1,700 deaths in 1943, but it had effectively controlled the disease by the end of the 1940s" (Gallup and Sachs, 2001).

3.3 Conceptual Argument for an Exogeneous Relationship with 20th-century Institutions and Development

An important contribution of this paper is to propose endemicity in 1900 as an instrumental variable for institutional quality at the turn of the 20th century. We argue that endemicity 1900's exogeneity rests on its archival nature which accurately captures the incidence of malaria in the period just preceding 1900, the environment of which afterward became radically altered during the first half of the 20th century through local and global mosquito-eradication efforts. In Packard's 2007 book *The Making of a Tropical Disease*, part of Johns Hopkins' Biographies of Disease series, he explains the uniqueness of disease history:

> "Environment, demography, ideas, and applied medical knowledge all interact to create particular ecologies of disease at particular moments in time. Disease is thus historically as much as biologically specific.… Biography implies a chronology and narrative—a movement in and through time.… No disease illustrates the complex interdependencies that shape disease incidence and experience better than malaria.… Malaria is multifactorial, exquisitely sensitive to particular environmental circumstances and social and economic relationships" (Packard, 2007, pp. viii-ix).

We take advantage of the historical mosquito-borne disease (MBD) snapshot in colonies and non-colonies. In particular, there have been cases of significant decreases in malaria incidence in some developed countries before the 1900s, such as in the UK due to decreasing acreages of marsh wetlands in the mid-19th century (Kuhn et al. 2003). However, despite this potential endogeneity associated with overall disease environment, we argue that endemicity 1990 is an exogenous variable regarding economic development in the 20th-century when crucial characteristics of a country are considered. Packard (2007, pp. ix) continues,

> "Geography is not malaria's one-dimensional destiny; in the past it has flourished in both old England and New England.… Malaria is both actor and acted upon in agricultural history, in the distribution of political and economic power, in imperial relationships, and in the movement of populations. Although a significant factor, it has never been an independent variable."

As addressed in the quote above, a potential concern may be the relationship between the geography and society of 1900 and the malaria endemicity at that moment in time. To address the possibility that historical agricultural development and population density could have affected variation in malaria endemicity in 1900, in our empirical analysis we include proxy measures for these characteristics of each country in our sample. Specifically, we use population and cropland data estimated by the Historic Database of the Global Environment (HYDE) (Klein Goldewijk, 2001; Klein Goldewijk et al., 2010).

4. Empirical approach and results

Using an instrumental variable (IV) approach, we start by investigating the explanatory power of endemicity 1900 for endogenous institutional quality in the first-stage estimation. Endemicity 1900 offers geographically expansive information about the disease environment in 1900. We use the regulatory quality indicator reported by the World Governance Indicators

dataset as a proxy for institutional quality. And we measure regulatory quality and GDP in the year 2000 because that year allows the largest set of countries to be considered. In later years, a few countries, such as Syria and Afghanistan, do not report GDP data.

We model the relationship between malaria endemicity 1900 E_i and regulatory quality in 2000 R_i as follows:

$$R_i = C + \gamma E_i + \omega X'_i + \upsilon_i \quad (1)$$

where $\hat{X}'_i$ is a vector of covariates that arguably affect the distribution of endemicity, the natural log of population density in 1900 and the natural log of cropland as a portion of the country in 1900. The random error term is υ_i.

In the second stage, we test the relationship between the instrumented institutional quality R_i and current economic performance y_i, as defined by the following equation:

$$y_i = \mu + \alpha \hat{R}_i + \omega X'_i + \epsilon_i \quad (2)$$

where y_i is the nature log of GDP per capita 2010 in country i, $\hat{R}_i$ is the regulatory quality instrumented by malaria endemicity 1900, $\hat{X}'_i$ is a vector of other covariates, and ϵ_i is a random error term. The coefficient of interest is α, the effect of institutions on income per capita. The summary statistics for all the variables used in the estimation of equations (1) and (2) are shown in Table 1.

Table 1: Summary Statistics

Variable	Observations	Mean	Standard Deviation
GDP per capita 2000	148	8.38	1.186
Endemicity 1900	148	2.432	1.326
Regulatory Quality 2000	148	46.649	28.102
Population Density 1900	148	3.688	1.6
Cropland 1900	148	0.216	1.859
Absolute Latitude	148	2.982	0.971
Temperature	148	3.615	0.257

Precipitation	148	4.145	0.904

4.1 Baseline Results

We present our baseline empirical results in Table 2. The first three columns correspond with the first stage of the IV approach and the next three columns present the second stage IV results. In column 3, the most conservative specification of the relationship between endemicity 1900 and regulatory quality in 2000 is estimated, controlling for population and cropland in 1900. Evidence for a non-weak relationship between endemicity 1900 and regulatory quality is apparent: endemicity 1900 has a statistically significant relationship with regulatory quality, and the F-statistics reported in column 3 is F=24.95.

Table 2: Estimating the Impact of Regulatory Quality on Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First-stage results			Second-stage results			OLS
Dependent variable:	Regulatory Quality, 2000			GDP per capita, 2000			
Endemicity 1900	-12.207***	-12.475***	-12.296***				
	(1.429)	(1.446)	(1.470)				
Regulatory Quality 2000				0.047***	0.046***	0.046***	0.031***
				(0.005)	(0.005)	(0.005)	(0.002)
Population Density 1900	2.120*		1.605	-0.064		-0.014	0.060
	(1.184)		(2.238)	(0.048)		(0.090)	(0.077
Cropland 1900		1.711*	0.529		-0.060	-0.050	-0.097
		(1.032)	(1.946)		(0.040)	(0.076)	(0.066)
Adjusted R^2	0.333	0.331	0.328	0.403	0.418	0.412	0.545
F	37.620	37.285	24.945	50.820	51.999	34.469	59.720
Observations	148	148	148	148	148	148	148

Notes: Coefficients are reported with classic standard errors in brackets. * p < 0.10, ** p < 0.05, *** p < 0.01

For purely investigatory purposes, according to the estimate in column 3, for a country initially with mean regulatory quality of 46.65 (which is measured on an index of 0 to 100), a one-standard-deviation decrease in endemicity 1900 would have raised the regulatory quality index to 62.95, which represents a 34.94 percent increase in regulatory quality.

Our first stage estimates support a strong relationship between endemicity 1900 and

regulatory quality. This empirical evidence taken together with the analytical arguments for the exogeneity of endemicity 1900 in section 3.3 mean that the subsequent IV estimates, reported in columns 4-6 of Table 2, should be interpreted as a portion of the causal relationship between institutional quality and income. The IV coefficients in column 6 represent the most conservative estimated impact because that specification includes the full set of covariates, including population density and cropland in 1900.

Our IV estimate in column 6 captures a local average treatment effect (LATE). The LATE is the effect of regulatory quality on income for those countries that were induced to alter their regulatory quality by the instrument, endemicity 1900. The LATE is the average treatment effect on the effectively treated, or those marginal countries for which the disease environment mattered in establishing regulatory quality.

Our baseline estimate of the impact of regulatory quality in 2000 on GDP per capita in 2000 can be interpreted as: for a country initially with mean income of $8,097.52, a one-standard-deviation increase in regulatory quality would have raised income to $17,654.19, which represents a 118.02 percent increase in income.

According to our LATE estimate, regulatory quality has a substantial impact on income in countries where the mosquito-borne disease environment affected regulatory quality. However, it is important to note that estimating a LATE has two drawbacks: (1) identifying this subset of countries, known as compliers from the treatment effect literature, is not possible and (2) the LATE may have poor external validity.

4.2 Endemicity 1900 versus Geography

The geographic distribution of malaria endemicity in 1900 is related to geographic determinants of institutional quality and economic development. Through the specification that

we present in Table 3, we explore the explanatory power of our instrument, endemicity 1900, while considering country geographic characteristics such as temperature, precipitation, latitude, coastline and tropical climate that could be associated with mosquito-borne disease rates and other channels linking malaria and underdevelopment. We include malaria ecology as a geographic characteristic to be controlled for in our IV approach. Our aim with including malaria ecology as an additional covariate is to highlight the differences between endemicity 1900 and malaria ecology, the latter being a time-invariant characteristic of a region.

The first six columns of Table 3 present again the first-stage relationship between endemicity in 1900 and regulatory quality. In columns 1 through 5, we include another geographic characteristic to the baseline specification. The only covariate with a strong relationship with regulatory quality, besides endemicity 1900, is the percentage of a country's landmass that is within 100 kilometers of a coastline.

Table 3: Estimating the Impact of Regulatory Quality on Income while Accounting for Geography

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	First-stage results						Second-stage results					
Dependent variable:	Regulatory Quality, 2000						GDP per capita, 2000					
Endemicity 1900	-11.749***	-11.456***	-13.957***	-12.410***	-12.053***	-10.912***						
	(1.834)	(1.830)	(1.821)	(1.829)	(1.404)	(2.070)						
Regulatory Quality 2000							0.041***	0.040***	0.039***	0.033***	0.046***	0.031***
							(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
Pop. Density 1900	1.611	1.413	1.776	1.607	-0.337	-0.559	-0.008	-0.021	-0.011	0.005	-0.019	-0.036
	(2.182)	(2.255)	(2.230)	(2.246)	(2.194)	(2.176)	(0.081)	(0.082)	(0.079)	(0.072)	(0.091)	(0.069)
Cropland 1900	-0.533	0.432	0.677	0.540	2.098	1.047	-0.027	-0.055	-0.055	-0.059	-0.045	0.003
	(1.937)	(1.953)	(1.939)	(1.956)	(1.901)	(1.946)	(0.072)	(0.070)	(0.067)	(0.061)	(0.077)	(0.061)
Temperature	-11.417					-17.172	-0.253					-0.386
	(9.304)					(11.306)	(0.383)					(0.394)
Precipitation	6.196***					4.104	-0.252***					-0.275**
	(2.159)					(3.393)	(0.082)					(0.115)
Absolute Latitude		1.935				3.417		0.172*				-0.011
		(2.507)				(3.324)		(0.097)				(0.109)
% Tropical Climate			0.086			0.034			-0.005***			0.002
			(0.056)			(0.098)			(0.002)			(0.003)
Malaria Ecology				0.036		0.373				-0.050***		-0.042***
				(0.338)		(0.335)				(0.011)		(0.010)
Percent Near					0.207***	0.201***					0.001	0.005**

Coast					(0.053)	(0.060)					(0.002)	(0.002)
Adjusted R^2	0.362	0.326	0.335	0.324	0.388	0.409	0.521	0.512	0.542	0.626	0.409	0.662
F	17.657	18.805	19.472	18.583	24.307	12.286	26.305	33.058	33.366	45.862	31.062	27.283
Observations	148	148	148	148	148	148	148	148	148	148	148	148

Notes: Coefficients are reported with classic standard errors in brackets. * p < 0.10, ** p < 0.05, *** p < 0.01

The coefficient for our variable of interest, endemicity 1900, in column 6 of Table 3 is not statistically different from our baseline estimate from column 3 of Table 2. However, the F-statistic in column 6 drops in magnitude considerably compared to the baseline. When we include the geographic characteristics, the F-statistic of the first-stage regression is F=12.29. In columns 7-12 of Table 3, we estimate the IV regressions while considering the geographic characteristics of a country. Column 12 contains the most conservative coefficients. We find that the impact of regulatory quality on income remains economically and statistically significant. However, the magnitude of the coefficient in column 12 of Table 3 is nearly one-third lower in magnitude than our baseline line estimate in Table 2. The covariates, precipitation, malaria ecology and land near the coast all have significant relationships with income. The comparison of the IV estimate of regulatory quality on income in Table 2 relative to Table 3 illuminates the role of geography in early disease environment. Beyond country geographic attributes, endemicity 1900 contributes to the development of institutions for the set of countries on the margin, the treatment compliers.

4.3 Examining Colonizer History

The role of colonists and their susceptibility to malaria has been put forth, most prominently by AJR (2001), as a mechanism through which the early disease environment determined institutional quality and influenced economic development. In this section, we parse our set of countries into two types: colonies and non-colonies. Our coefficient of interest, the impact of regulatory quality on income, will probably differ between the two samples as the countries on the margin that determine the LATE effect are different. We present the estimates in Table 4. We specify the colonies and non-colonies below:

Colonized by Western European Countries (Colonies): Algeria, Angola, Argentina, Australia, Bangladesh, Belize, Bolivia, Botswana, Benin, Brazil, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, Colombia, Congo Republic, Costa Rica, Cote d'Ivoire, Djibouti, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Gabon, The Gambia, Ghana, Guatemala, Guinea, Guinea-Bissau, Guyana, Honduras, India, Indonesia, Iraq, Israel, Jordan, Kenya, Kuwait, Laos, Lebanon, Libya, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mexico, Mozambique, Myanmar, Namibia, Nepal, New Zealand, Nicaragua, Niger, Nigeria, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Qatar, Rwanda, Saudi Arabia, Senegal, Sierra Leone, South Africa, Sudan (including South Sudan), Suriname, Tanzania, Togo, Tunisia, Uganda, United Arab Emirates, United States, Uruguay, Venezuela, Vietnam, Yemen, Zambia and Zimbabwe.

Not Colonized by Western European Countries (Non-Colonies): Albania, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, China, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Greece, Germany, Hungary, Iceland, Ireland, Iran, Italy, Japan, Kazakhstan, Kyrgyzstan, Liberia, Lithuania, Latvia, Moldova, Mongolia, North Macedonia, Netherlands, Poland, Portugal, Republic of Korea, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Thailand, Tajikistan, Turkey, Turkmenistan, Ukraine, United Kingdom and Uzbekistan.

Table 4: Accounting for the Colonial Past When Estimating the Impact of Regulatory Quality on Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First-stage results			Second-stage results			OLS		
Dependent variable:	Regulatory Quality, 2000			GDP per capita, 2000					
Endemicity 1900	-12.296***	-6.786***	-15.115***						
	(1.470)	(2.070)	(3.361)						
Regulatory Quality 2000				0.046***	0.054***	0.027***	0.031***	0.026***	0.025***
				(0.005)	(0.015)	(0.005)	(0.002)	(0.004)	(0.003)
Population Density 1900	1.605	-3.616	9.239*	-0.014	0.099	-0.032	0.060	-0.041	0.007
	(2.238)	(2.462)	(5.049)	(0.090)	(0.137)	(0.153)	(0.077)	(0.093)	(0.126)
Cropland 1900	0.529	0.646	-3.178	-0.050	-0.151	0.039	-0.097	-0.185**	0.016
	(1.946)	(2.074)	(4.337)	(0.076)	(0.098)	(0.115)	(0.066)	(0.075)	(0.103)
Adjusted R^2	0.328	0.203	0.402	0.412	0.184	0.620	0.545	0.508	0.624
F	24.945	9.044	12.409	34.469	15.140	14.067	59.720	33.741	29.189
Sample	Global	Colony	Non-Colony	Global	Colony	Non-Colony	Global	Colony	Non-Colony
Observations	148	96	52	148	96	52	148	96	52

Notes: Coefficients are reported with classic standard errors in brackets. * p < 0.10, ** p < 0.05, *** p < 0.01

In Table 4, the first three columns present the first-stage results. Within the bottom lines of the table, we specify the sample of underlying data: the baseline global sample, colonies, or non-colonies. We use a Wald test to establish whether the coefficient of endemicity 1900 differs between the sample of colonies and non-colonies (columns 2 and 3, respectively) and find that

that we can reject the null hypothesis. The relationship between endemicity 1900 and institutional quality is significantly stronger for countries that were never colonized. The F-statistics in columns 2 and 3 for the smaller samples of colonies and non-colonies is much smaller than the global sample in column 1. Taken together, there is evidence that endemicity 1900 is a non-weak instrument for both the sub-samples.

The second stage IV results for the sub-sample of colonies and non-colonies are presented in columns 5 and 6 of table 4. The estimate for colonies is nearly double the magnitude of non-colonies. In LATE terminology, the average treatment effect in particular former colonies in which the malaria environment influenced regulatory quality is twice as large as that found in non-colonies in which endemicity 1900 influences regulatory quality.

Do our results suggest any nuanced legacy of colonization? Perhaps. For a former colony with a mean income of $5,564.28, a one-standard-deviation increase in regulatory quality would have raised income to $14,889.62, which represents a 167.59 percent increase in income. In contrast, in a country that was never colonized with a mean income of $12,774.26, a one-standard deviation increase in regulatory quality would have raised income to $18,745.82, representing just a 46.75 percent increase in income. Economic performance in former colonies is more sensitive to differences in institutional quality, as determined by early disease environments. Another way to look at this is that economic development in countries that were never colonized is more robust, although it is still linked to the formation of institutional quality determined by early malaria endemicity.

5. Replication of Acemoglu, Johnson, and Robinson (2001) Core Results

The main finding in AJR's (2001) seminal work is that differences in colonial history for a country could cause differences in early institutions, contemporary institutions, and therefore,

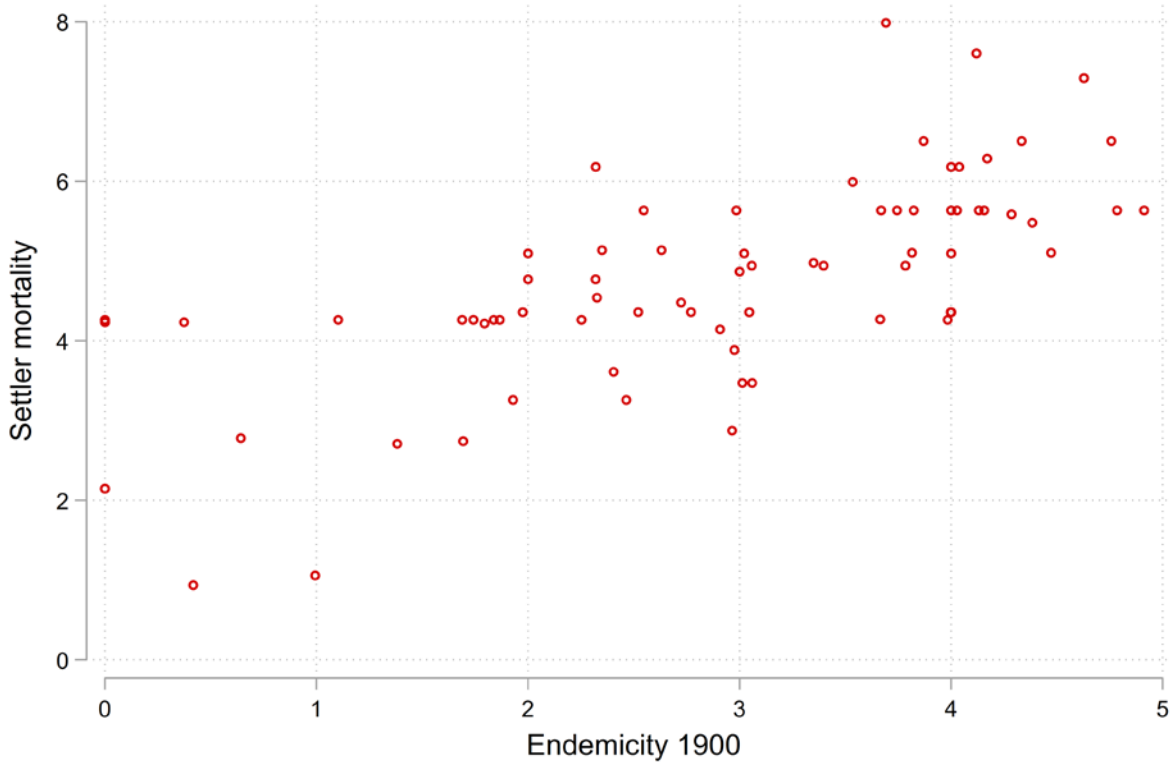
in the levels of contemporary development. The idea of the "European colonial experience" is the central feature of the AJR (2001) causal mechanism, which postulates that European colonizers' decisions on what form of colony to establish—inclusive or extractive—influenced the subsequent development trajectories of many non-European nations around the world.

Mirroring the estimating equation (1) from the empirical strategy section of this paper, AJR (2001) use settler mortality to explain early institutions. These latter have a significant relationship with contemporary institutional quality, which AJR proxy for with protection against expropriation (R_i). In the second stage, we test the relationship between instrumented institutions (R_i) and current performance (y_i), reproducing equation (2) presented earlier in this paper

In this section, we aim to compare settler mortality with malaria endemicity 1900 by replicating the core analysis in AJR (2001) using both the variables and sample of countries from their study. Then, by exploiting the global nature of our malaria endemicity measure, we compare the relationship between the early disease environment, institutions, and development for countries with and without colonial histories. We establish that the transmission mechanism between early disease environment and economic development is not dependent on the role of European colonizers.

To begin, we graph the pair-wise correlation between settler mortality and endemicity 1900 in Figure 3. The two variables are closely related, especially for African countries. In South America, we can observe minor variation in settler's mortality across countries, while endemicity 1900 variation is considerably higher. Because of significant differences in the disease environment across the South American continent, we believe endemicity 1900 is a better approximation of the actual environment of settlers' decisions.

Figure 3: Correlation between Endemicity 1900 and Settler Mortality (AJR 2001)



Despite the wide acceptance of their mechanism, AJR's (2001) contribution had its critics. One serious limitation of their approach was the measurement error involved with settler mortality, the key variable in AJR (2001)'s the causal mechanism. Settler mortality was constructed from a combination of death records ranging from European soldiers to Catholic bishops during diverse times of peace and military campaigns, and in addition, 36 of the 64 country-level observations in AJR's (2001) sample were assigned mortality rates from other countries, often based on what would appear to be mistaken or conflicting evidence (Albouy, 2008; Albouy, 2012; Sachs, 2012).

Other authors have emphasized the role European colonizers played on human capital accumulation in their colonies. Glaeser et al. (2004) postulated that Europeans brought human capital to the colonies, which formed the basis for good institutions, rather than bringing ready institutions themselves. Putterman and Weil (2010) emphasized the role of migration during the colonization era, focusing on people rather than places. Using a similar idea, Chanda et al. (2014)

argued for the persistence of fortune, finding that human capital accumulation and the technology brought by ancestor populations determine current economic prosperity. Those colonies with a higher share of technologically advanced European settlers define higher prosperity today.

Spolaore and Wacziarg (2009) stressed the role of genetic distance between populations as a determinant for the diffusion of technology, which then explains why countries with higher European population and closer genetic distance could advance in economic development by transferring the technology from the European mainland. Encompassing these views, Easterly and Levine (2016) constructed a variable that approximates the share of Europeans in former colonies and suggests that the higher share of Europeans in the local population during colonization significantly improved the economic prosperity of those countries today. However, in colonized countries with a low share of Europeans (4.8% or lower), the effect of European presence was negative for current development compared to countries with a "zero" Europeans' share in the population.

We estimate the relationship in (2) using the original sample in AJR (2001) comprising 62 countries and present the results in Table 5. The results show that early institutions used by AJR (2001), European settlements, and European settlers' mortality were shaped by historical malaria endemicity. Columns 1-4 relate historical malaria endemicity to early institutions. Comparing the explanatory power of endemicity 1900 to settler mortality in table 5, both variables explain 31 to 37 percent of the variation in early institutions. Columns 5-6 show that endemicity 1990 explains around 60 percent of the variation in European share of population in 1900. The degree to which endemicity 1900 and settler mortality explain settlement pattern is similar.

Table 5: Determinants of institutions, Replication of AJR (2001) Table 3 Panel B

	Constraint on executive 1900		Democracy in 1900		European share of population 1900	
	(1)	(2)	(3)	(4)	(5)	(6)
	-0.750***		-1.005***		-0.000***	

Settler Mortality	(0.205)		(0.278)		(0.000)	
Endemicity 1900		-0.885***		-1.018***		0.000***
		(0.230)		(0.326)		(0.000)
Absolute Latitude	3.848*	1.439	7.552***	5.483*	0.000***	0.000***
	(1.928)	(2.264)	(2.614)	(3.192)	(0.000)	(0.000)
Adjusted R^2	0.313	0.329	0.378	0.345	0.458	0.612
F	14.238	14.225	18.304	14.980	26.762	44.414
Observations	59	55	58	54	62	56

Notes: Coefficients are reported with classic standard errors in brackets. * p < 0.10, ** p < 0.05, *** p < 0.01

We present the two-stage least-squares estimation and its accompanying first-stage regression results in Table 6. The baseline results are presented in columns 1 to 4, where the first- and second-stage results using settler mortality are reported in columns 1 and 3, respectively, and the first- and second-stage results using endemicity 1900 are reported in columns 2 and 4, respectively. In comparing the first-stage results, the effect of both settler mortality and endemicity 1900 on the average expropriation risk are not statistically different from one another (z = -0.21) (Clogg et al., 1995). Likewise, the causal relationship between institutional quality and income estimated using the second-stage regression for settler mortality in column 3 and then for endemicity 1900 in column 4 are each statistically different from zero but not statistically different from one another (z = -0.04).

Table 6: Replication of Core Results for AJR (2001)

	Colonies within AJR (2001) sample				Base sample excluding Neo-Europes				Base sample excluding Neo-Europes & African nations			
	1st stage: Expr. Risk		2nd stage: GDP 1995		1st stage: Expr. Risk		2nd stage: GDP 1995		1st stage: Expr. Risk		2nd stage: GDP 1995	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Expropriation Risk			0.997***	1.012***			1.226***	1.210***			0.259	0.806**
			(0.219)	(0.291)			(0.352)	(0.449)			(0.356)	(0.297)
Settler Mortality	-0.509***				-0.389***				-0.666			
	(0.138)				(0.136)				(0.432)			
Endemicity 1900		-0.465***				-0.363**				-0.506**		
		(0.160)				(0.151)				(0.187)		
Absolute Latitude	1.625	1.581	-0.365	-0.767	-0.711	-1.421	1.536	1.347	-0.798	-3.341	-0.047	0.299
	(1.327)	(1.577)	(1.270)	(1.656)	(1.459)	(1.674)	(1.412)	(1.597)	(2.054)	(2.170)	(1.166)	(1.589)
Adjusted R^2	0.262	0.253	0.096		0.098	0.071			0.020	0.179	0.238	
IV-used			SM	E			SM	E			SM	E
F	12.009	10.292	16.934	11.916	4.153	2.957	7.290	4.431	1.273	3.732	0.292	3.707
Observations	63	56	63	56	59	52	59	52	28	26	28	26

Notes: Coefficients are reported with classic standard errors in brackets. * p < 0.10, ** p < 0.05, *** p < 0.01

We repeat this analysis in columns 5-8, where we exclude Neo-Europes (USA, Canada, Australia, New Zealand), and in columns 9-12, where we exclude both Neo-Europes and African nations. In the first-stage regressions reported in columns 5 and 6, the coefficient of interest on both potential instruments are statistically different from zero, while the magnitude of the negative relationship between the historical disease environment and modern institutional quality is not statistically different (z = -0.13). The coefficient on expropriation risk from the second-stage regressions reported in columns 7 and 8 for the sample excluding Neo-Europes are also statistically different from zero. The strength of the causal relationship between institutional quality and GDP in 1995 is not affected by which instrument was used in the analysis; the coefficients reported in columns 7 and 8 are not statistically different from each other (z = 0.03).

In columns 9 and 10, when we exclude both Neo-Europes and African countries, the instrumental variables, settler mortality and endemicity in 1900, continue to produce results that are not statistically different from one another. The coefficients on settler mortality and endemicity 1900 from the first-stage regressions, reported in columns 9 and 10, are similar in magnitude (z = -0.34). However, one difference in the two reported statistics stands out. In column 9, when the settler mortality data are used, the p-value testing the coefficients' difference from zero is low, while the p-value reported with the coefficient on endemicity 1900 in column 10 is much greater. Likewise, the F-statistic in column 9 using settler mortality is also low, while the F-statistics in column 10 using endemicity in 1900 is more in line with those reported in columns 5 and 6. The greater statistical significance of the coefficient on endemicity 1900 captured by the p-value and the fit of the regression model captured by the F-statistic in column 10 likely reflects the greater variation in the historical disease environment, particularly in South America, captured by the

endemicity 1900 variable.

The results from the second-stage regression are reported in columns 11 and 12, using settler mortality and endemicity 1900, respectively. In comparing the coefficients for expropriation risk across the two columns, they are not statistically different from one another (z = -1.18). When settler mortality is used as the instrument, the coefficient of interest is not statistically different from zero, arising from the relatively poorly explanatory power in the first stage. In column 12, the effect of expropriation risk on income in 1995 is statistically significant and similar in magnitude to the coefficient reported in the baseline analysis from column 4 (z = 0.5).

In summary, we find that endemicity 1900 functions in a similar manner to settler mortality within the context of AJR (2001) in terms of variables used and sample of colonies. However, endemicity 1900 allows for greater variation outside of the Neo-Europes and African nations.

6. Conclusion

In this paper, we contribute to the debate about the development of early institutions, and hence modern economic development, in all countries, including European colonies, non-European colonies and countries that never were colonies. We offer a holistic explanation using malaria endemicity in the year 1900 as an alternative instrumental variable for institutional development on samples of colonies and non-colonies.

In our analysis, we first examine the relationship between endemicity 1900 and regulatory quality in 2000, considering population and cropland in 1900. In the first stage, we find evidence of a statistically significant relationship between endemicity 1900 and regulatory quality. In the second stage, according to our LATE estimate, regulatory quality has a substantial impact on income in countries where the mosquito-borne disease environment affected regulatory quality.

Further, we explore the explanatory power of our instrument, endemicity 1900, while considering geographic characteristics such as temperature, precipitation, latitude, coastline and tropical climate that are associated with mosquito-borne disease rates and other channels linking malaria and underdevelopment. We include malaria ecology as a geographic characteristic to be controlled for in our IV approach. The coefficient of interest on endemicity 1900 is not statistically different from our baseline estimate; however, the F-statistic drops in magnitude compared to the baseline. We find that endemicity 1900 contributes to the development of institutions for the set of countries on the margin, the treatment compliers, apart from those countries' geographic attributes.

Next, we parse our set of countries into two types: colonies and non-colonies. The role of colonists and their susceptibility to malaria has been put forth as a mechanism through which the early disease environment determined institutional quality and influenced economic development. We find evidence that endemicity 1900 is a non-weak instrument for both the sub-samples and that the estimate for colonies is nearly double the magnitude of non-colonies. In LATE terminology, the average treatment effect in particular former colonies for which the malaria environment influenced regulatory quality is twice as large as in non-colonies whose regulatory quality was influenced by endemicity in 1900. Economic development in countries that were never colonized is more robust, although it is still linked to the formation of institutional quality determined by early malaria.

Last, we replicate the evidence in AJR (2001) that early institutions were shaped, at least in part, by colonial settlements, and settlements were affected by mortality. We substitute endemicity 1900 for settler mortality (used by AJR (2001) and find evidence that European settlements and mortality were shaped by historical malaria endemicity. We also find that the

causal relationship between institutional quality and income estimated using the second-stage regression for settler mortality and then for endemicity 1900 are each statistically different from zero but not statistically different from one another. The causal relationship between institutional quality and GDP in 1995 is not affected by which instrument was used in the analysis, with the coefficients reported not statistically different from each other.

AJR's principal finding is that differences in colonial experience could be a source of exogenous difference in contemporary institutions and therefore level of development. Clearly, our findings do not invalidate the causal link based on settlement mechanism argued by AJR. In fact, we confirm AJR's assertion in their conclusion that "the colonial experience is one of many factors affecting institutions."

However, not precluding any role European colonizers may have played, we find a common global effect played by the early malaria environment on early institutions that is independent of colonial history. This strongly suggests the presence of non-colonial origins of institutional quality, pointing to a more general relationship between disease environment, early institutions, and modern development. Because historical human capital formation—based on factors such as household decisions on schooling, migration, and savings—was the basis of high-quality early institutions, the results in this paper show that those household decisions would have been present with or without European colonization. This line of reasoning buttresses the contribution of Glaeser et al. (2004), which emphasizes that Europeans brought human capital to the colonies, rather than institutions.

Our findings lead us to conclude that in low-disease environments, quality institutions would have formed regardless of European colonization. The paper provides strong evidence supporting the need for continued investigation into the determinants of economic growth. Not

only do our results illuminate the absence of an endpoint for this literature, but they also question the conventional centrality of a colonial legacy in the economic success of nations.

References

Acemoglu, D., Johnson, S., & Robinson, J. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. The American Economic Review 91 (5): 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. Quarterly Journal of Economics 117 (4): 1231–1294.

Albouy, D. (2008). The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data. National Bureau of Economic Research Working Paper 14130.

Albouy, D. (2012). The Colonial Origins of Comparative Development: An Empirical Investigation: Comment. American Economic Review 102 (6): 3059–76.

Auer, R. A. (2013). Geography, institutions, and the making of comparative development. Journal of Economic Growth, 18(2), 179-215.

Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: a production function approach. World development, 32(1), 1-13.

Carstensen, K., & Gundlach, E. (2006). The Primacy of Institutions Reconsidered: Direct Income of Malaria Prevalence. World Bank Economic Review 20 (3).

Chanda, A., Cook, C. J., & Putterman, L. (2014). Persistence of fortune: Accounting for population movements, there was no post-Columbian reversal. American Economic Journal: Macroeconomics, 6(3): 1–28.

Easterly, W., & Levine, R. (2016). The European origins of economic development. Journal of Economic Growth, 21(3), 225-257.

Gallup, J. L., Sachs, J. D., & Mellinger, A. D. (1999). Geography and economic development. International Regional Science Review, 22(2): 179–232.

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

Glaeser, E. L., La Porta, F., Lopez-de-Silanes, R. & Schleifer, A. (2004). Do Institutions Cause Growth? Journal of Economic Growth 9 (3): 271–303.

Gooch, E. (2017). The Impact of Reduced Incidence of Malaria and other Mosquito-borne Diseases on Global Population. Journal of Development Economics, 124: 214-228.

Guo, Z., et al. (2015). Argoecology, Demographics, Markets, Farming Systems, Agricultural R&D: HarvestChoice 5-Arc Minute Spatial Layers for sub-Saharan Africa Release {Version}. HarvestChoice, IFPRI, University of Minnesota; Available online at http://www.harvestchoice.org/data/.

Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M., & Snow, R. W. (2004). The Global Distribution and Population at Risk of Malaria: Past, Present, and Future. Lancet Infect Dis. 4 (8): 327–38.

Kiszewski, A., et al. (2004). A Global Index Representing the Stability of Malaria Transmission. The American Journal of Tropical Medicine and Hygiene 70 (5).

Klein Goldewijk, K. (2001). Estimating global land use change over the past 300 years; the HYDE database. Glob. Biogeochem. Cycles 15 (2), 417–433.

Klein Goldewijk, K., Beusen, A., Janssen, P. (2010). Long term dynamic modeling of global population and built-up area in a spatially explicit way, HYDE 3.1. Holocene 20, 565–573.

Kuhn, K. G., Campbell-Lendrum, D. H., Armstrong, B., & Davies, C. R. (2003). Malaria in Britain: Past, present, and future. Proceedings of the National Academy of Sciences, 100(17): 9997–10001.

Lysenko, A. J., & Semashko, I.N. (1968). Geography of Malaria. a Medico-Geographic Profile of an Ancient Disease. Medicinskaja Geografija, 25–146.

Packard, R. (2007). The making of a tropical disease. Baltimore: The Johns Hopkins University Press.

Rigobon, R., & Rodrik, D. (2005). Rule of law, democracy, openness, and income. Economics of transition, 13(3), 533-564.

Putterman, L., & Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. The Quarterly Journal of Economics, 125(4): 1627–1682.

Sachs, J. & Malaney, P. (2002). The Economic and Social Burden of Malaria. Nature 415: 680-685.

Sachs, J. D. (2012). Reply to Acemoglu and Robinson's response to my book review. http://jeffsachs.org/(2012/12/reply-to-acemoglu-and-robinsons-response-to-my-book-review. Accessed 11 January 2016.

Spolaore, E., & Wacziarg, R. (2009). The diffusion of development. The Quarterly Journal of Economics, 124(2): 469–529.