Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Institutions Do Not Rule: Reassessing the Driving Forces of Economic Development

Jinfeng Luo
And
Yi Wen

Working Paper 2015-001A
https://dx.doi.org/10.20955/wp.2015.001

January 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Institutions Do Not Rule:
Reassessing the Driving Forces of Economic Development[1]

Jinfeng Luo
Tsinghua University

Yi Wen
Federal Reserve Bank of St. Louis
& Tsinghua University

(This version: January 16, 2015)

Abstract: We use cross-country data and instrumental variables widely used in the literature to show that (i) institutions (such as property rights and the rule of law) do not explain industrialization and (ii) agrarian countries and industrial countries have entirely different determinants for income levels. In particular, geography, rather than institutions, explains the income differences among agrarian countries, while institutions appear to matter only for income variations in industrial economies. Moreover, we find it is the stage of economic development (or the absence/presence of industrialization) that explains a country's quality of institutions rather than vice versa. The finding that institutions do not explain industrialization but are instead explained by industrialization lends support to the well-received view among prominent economic historians—that institutional changes in 17th and 18th century England did not cause the Industrial Revolution.

Keywords: geography, institutions, development, income gaps, industrialization.

JEL Code: O11, P16, P51.

[1] We thank seminar participants at the St. Louis Fed for comments and Judy Ahlers for editorial assistance.

1. Introduction

Why are some countries so rich and others so poor?[2] This may be the single most important question in economics since Adam Smith (1776), but economists today are still sharply divided in their answers. An old school of thought maintains that economic development depends ultimately on the geographic environment or humans' natural living conditions (see Bodin, 1533; Montesquieu, 1748; Myrdal, 1968). This ancient concept is reincarnated in Diamond's (1997) famous book, *Guns, Germs and Steel*: *The Fates of Human Societies*. Diamond uses new evidence from anthropology, biology, and geography to argue that continental terrain and uneven distribution of domesticated plants and animals led to differences in grain production and its spread/transmission across human settlements, which ultimately explains the huge difference in economic, political and military powers across continents. Simply put, long-term economic performance is determined by geographic factors.

However, the institutional school disagrees. This school of thought insists that (i) institutions, especially property rights and the rule of law, are the fundamental cause and the only important determinant of long-term economic performance and (ii) geographical factors work (if at all) only through the channel of institutions (see, e.g., North, 1981; Acemoglu and Robinson, 2001; Easterly and Levine, 2003; Rodrik et al., 2004).

In the meantime, researchers from the geographic school are not convinced by the arguments and evidence presented by the institutional school. They argue that both institutional and geographic factors matter but geography has played a critical role in economic development (Gallup and Sachs, 2001; McCord and Sachs, 2013). Scholars in this camp also insist on multidimensional factors in determining long-term economic performance, instead of a single-factor framework. For example, in addition to institutions and geography, factors such as culture, economic policy, human capital, and specific historical events are also important (Glaeser et al., 2004; Becker and Woessmann, 2009; Tabellini, 2010; Jedwab and Moradi, 2012).

Nonetheless, these schools of thought share one implicit assumption in common: The driving forces of long-term economic performance remain the same throughout history regardless of a country's developmental stage.

We test this key hypothesis here and find that it is false. Our analysis shows that the fundamental determinant of a nation's income level differs at different developmental stages. More specifically, based on widely used cross-country data and instrumental variables, we find that geography is the most powerful and only significant factor in explaining income differences across agrarian economies, while institutional factors appear prominent only in accounting for the income variations among the already industrialized economies.

Our findings make good common sense: In agrarian societies, the productivity of land and labor are inevitably dependent on and vulnerable to water, soil, disease, and climate conditions. Organic production is land intensive and weather (mother-nature) sensitive. Hence, geographic forces must matter a great deal in determining income gaps among agrarian countries. However, after industrialization (due to whatever reasons), tropical countries

[2] The difference in income levels across countries is astonishing. For example, the per capita income gap between Monaco and the Democratic Republic of Congo was almost 700-fold in 2010.

(regions) such as Singapore and Hong Kong became immune to the "geographic curse." These observations indicate that at different developmental stages the forces driving income variations may be fundamentally different. Lumping countries at different developmental stages together in the empirical analysis may lead to inconclusive and even incorrect conclusions.

This paper adopts a different approach by allowing for the possibility of different driving forces at different developmental stages to explain cross-country income variations. Specifically, we divide the full country sample into two subsamples—agrarian countries and industrial countries—based on each country's share of agriculture value added (AVA) in gross domestic product (GDP) (the AVA-to-GDP ratio), or rural population share. This approach decomposes the global income differences into three dimensions: (i) the income variation across agrarian countries, (ii) the income variation across industrial countries, and (iii) the income gap between agrarian countries and industrial countries.

We find that geographic factors are crucial in explaining income differences among agrarian countries but not among industrial countries, whereas institutional factors are significant for income variations across industrial countries but not across agrarian countries. This finding provides a plausible resolution to the controversy between the geographic and institutional schools of thought and explains why the more recent literature tends to find both institutional and geographic factors important when lumping together the two country groups in their analysis (see, e.g., McCord and Sachs, 2013).

More intriguing and important questions naturally emerge from our framework: What causes industrialization? What determines the income gaps between the two country groups? And what forces have prevented agrarian societies from becoming industrialized? To answer such questions, we take a preliminary step by constructing proxies of developmental stages and rerun our regression analyses based on the full sample and full set of instrumental variables. Surprisingly, we find geography, rather than institutions, explains a country's development stage or lack of industrialization. In other words, geography appears to be the ticket to industrialization, while institutions are not. This finding may explain why so many agrarian countries that intended to copy Western institutions to kick-start industrialization often failed to industrialize, lending support to the argument (Pomeranz, 2000) that Europe's 19th century divergence from the old agrarian equilibrium owes much to the fortunate location of coal rather than to institutions.

Using the AVA-to-GDP ratio (or the rural population share) to index/proxy industrialization may seem dubious since such variables are highly correlated with per capita GDP. However, what is important is that we find that institutions are completely insignificant in explaining the AVA-to-GDP ratio (or rural population ratio) once geography is taken into account, suggesting that development stages (or the absence/presence of industrialization) are not the same thing as per capita income levels.

The finding regarding the cause/determinant of industrialization is somewhat puzzling because at least for the industrial countries we find that institutional quality matters greatly for income variations. Why is it that institutions do not help explain industrialization? This motivates us to look further into the possibility that institutions are endogenous to economic development—they are the consequence, rather than the cause, of industrialization. That is, we conjecture that the stage of economic development determines a country's institutional quality, which in turn feeds back into income growth for already industrialized nations. We

conduct a series of analyses to explore this theoretical possibility and our empirical findings support this hypothesis.

The finding that institutions do not explain developmental stage (or the absence/presence of industrialization) is nonetheless not new. Many prominent economic historians have already cast serious doubts on the notion that formal institutions (such as property rights and the rule of law) caused the Industrial Revolution (see, e.g., Robert Allen, 2009; Gregory Clark, 2007; Deirdre McCloskey, 2010; Christine MccLeod, 1988; Joel Mokyr, 2008; and Kenneth Pomeranz, 2000; among others).

One critical puzzle left unsolved by our study is the question regarding the conditions under which geographically disadvantageous nations could become industrialized. For example, why were tropical regions such as Singapore and Hong Kong able to industrialize at some point in history? Why was China able to kick start an industrial revolution since 1980s while unable to do so earlier, say back in 19th century or early 20th century, given that its geographic conditions have changed little since the English Industrial Revolution? The geographic school cannot answer these questions. Nor can the institutional school based on the dichotomy of inclusive/exclusive institutions, as China still remains a socialistic and authoritarian society despite 35 years of hyper growth. This puzzle is related intrinsically to the puzzle of the Industrial Revolution itself: What triggered the Industrial Revolution? Why did it happen in Europe instead of China or India? And why did it start first in England instead of France and Italy? These questions are beyond the scope of this paper and we leave them for future research.[3]

The remainder of the paper is organized as follows. Section 2 briefly reviews related literature as a background for the geography versus institutions debate. Section 3 provides descriptions of the data and our empirical models. Section 4 presents our analytical frameworks, the results, and robustness tests. Section 5 discusses the causal relationship between institutions and industrialization. Section 6 concludes with discussions for further research.

2. Related Literature

The institutional school emphasizes the importance of institutions, such as property rights, the rule of law, and inclusive political systems on long-term growth. The fundamental theory behind this view is that secure property rights are beneficial and conducive to investment in both physical and human capital, technology innovations, and efficient resource allocations through markets, thus enabling faster productivity growth and a higher level of income (North, 1981). A limited set of empirical evidence appears to favor this view. For example, cross-country studies confirm a positive causal relationship between property right protection and economic performance (see Rodrik, 1999; Acemoglu and Robinson, 2001; Acemoglu and Johnson, 2005; Acemoglu, Johnson and Robinson, 2001, 2005). Micro data also show that good institutions boost investment and output at the regional and firm level (Johnson et al, 2002; Iyer and Banerjee, 2005).

Within this strand of the literature, the work most closely related to ours is that of Acemoglu, Johnson and Robinson (AJR, 2001). Based on Curtin's (1989) historical record on the European settlement mortality rate in the colonies, these authors propose using the settlement

[3] Interested readers are referred to Wen (2015) and the references therein.

mortality rate in the colonial period as an instrumental variable for modern institutions (which are correlated with national income). They argue that for countries with lower settlement mortality rates, it is more likely and worthwhile for the settlers to replicate Western institutions. Since institutions tend to persist into the future, the importance and causal effects of institutions on long-run development can then be established by regressing modern economic performances of the postcolonial countries on measures of modern institutions instrumented by the settlement mortality rates in the early colonial period. Based on this approach, AJR (2001) report a significant and positive effect of institutions on long-term economic performance. This particular instrumental variable has since been widely adopted in many related studies to test the "institution hypothesis" and the "geography hypothesis." This follow-up literature shows that when institutions are controlled for, all other factors (such as geography and trade) seem to lose significance in explaining cross-country income differences, thus support AJR's theory that institutions are the only important factor in determining long-term economic performance (see, e.g., Easterly and Levine, 2003; Rodrik et al., 2004). [4]

A competing view argues that geographic factors matter fundamentally to development. The geography hypothesis has a long history that can be traced back at least to Bodin (1533), Montesquieu (1748), and Myrdal (1968). There has been a recent reincarnation of this thesis after Diamond's (1997) best-selling and provocative book. In general, this strand of the literature maintains that geography determines a nation's income level and paths of development. Institutions may also matter, but they are ultimately shaped by geography. Geography can affect development in many ways. In addition to determining transportation costs of trade and the availability of natural resources for nutrition and production, geography also affects human health, work effort, and technologies. For instance, Montesquieu (1748) and Marshall (1890) believe climate is important for work effort and labor productivity. Myrdal (1968), Diamond (1997) and Sachs (2001) emphasize that the geographic environment constrains technologies available to a society, especially in agriculture.

In recent years, the long-run economic effects of diseases induced by geography have also generated much attention (Sachs and Malaney, 2002). Gallup, Sachs and Mellinger (1999) provide a good review on related theoretical works. For relevant empirical evidence, see Sachs and Malaney (2002), Sachs (2003), McCord and Sachs (2013), and Datta and Reimer (2013).

This paper is also related to the literature on the Industrial Revolution. The works of AJR (2005) and Acemoglu and Robinson (2012), following the seminal work of North (1981), suggest that the establishment of property rights and inclusive political systems paved the way for the Industrial Revolution in late 18th century England and 19th century Europe and the United States, and can explain the dismal failures of development for many developing countries in the 20th century. Recent researchers, however, cast doubt on this view. By comparing the market performance of England and China (Yangtze River Delta region) on the eve of the Industrial Revolution, Shiue and Keller (2007) report that the role of institutions in the Industrial Revolution was ambiguous. Pomeranz (2000) argues that

[4] A potential drawback of using the settlement mortality rate as a valid instrumental variable for institutions is that it inevitably contains information on geography (e.g., epidemic). To mitigate this problem, this paper also adopts legal origins as an additional instrument for institutions. See Section 3 for further discussion.

Europe's 19th century escape from the Malthusian trap (in sharp contrast to China's inability to industrialize despite similar market environments and property rights) owes much to the fortunate location of coal and Atlantic trade rather than to institutions. Easy access to coal as a substitute for timber allowed Europe to grow along a resource-intensive, labor-saving path. Boldrin and Levine (2008) note that property rights did not afford great wealth to European inventors at the time of the Industrial Revolution, so the channel for institutions to incentivize technology innovation and industrialization is weak. Allen (2009) argues that the unique geographic condition of easily accessible cheap coal in England, in conjunction with its exceptional high wages, rather than institutions, determined that the Industrial Revolution first took place in England instead of other European or Asian countries.

3. Data

Many geographic factors can affect long-term economic performances. One possible geography channel that affects development is localized disease. For example, malaria risk presents a particular disease burden on humans because it affects labor productivity and population growth. In general, studies show a significant negative correlation between diseases and per capita income (McCarthy et al., 2000; Gallup and Sachs, 2001; Sachs and Malaney, 2002; Datta and Reimer, 2013). Hence, following Sachs (2003), and Carstensen and Gundlach (2006), we use malaria risk, or the risk of malaria transmission, as a proxy for geography. Malaria risk is a good proxy because it is closely related to local exogenous ecological conditions, specifically the type of mosquito vectors and the climate conditions (Kiszewski et al., 2004).[5] Our measure of malaria risk is from Sachs (2003), who calculates the population ratio in areas with malaria transmission risk based on the World Health Organization's (WHO) global malaria map published in 1994. Another measure of geography, fatal malaria risk, is gathered in a similar manner: It multiplies the malarial risk index by an estimate of the proportion of national malaria cases that involve the fatal mosquito species—*Plasmodium falciparum*—as opposed to three largely nonfatal species of the malaria pathogen.[6] Such variables are also widely used in papers of the institutional school to control for the effects of geography when analyzing the role of institutions (see Acemoglu and Robinson, 2001; Rodrik et al, 2004).

Geographic diseases such as malaria are persistent throughout human history, hence tend to correlate with 17-19 century settlement mortality risk. However, using this variable as proxy for geography should not introduce estimation bias for our purpose because (i) our country sample for malarial risk includes non-colonial countries and is twice as large as the European-colonial countries involved with the settlement mortality risk, and (ii) we instrument malarial risk using purely geographical factors such as malaria ecology and this instrument variable is also not limited to early European settlement countries and covers countries more than twice as many as those used by AJR (2001). Also, as argued by AJR (2001), although malaria could affect settlers' mortality rate, it should not have direct influence on current income performances of the ex-colonies other than through institutions. Hence, once institutions are

[5] Malaria is closely associated with climate region, especially in tropical areas, suggesting that it is a geography-related disease by nature. The underlying biologic reason this is that malaria is intrinsically a disease limited to warm environments because a key part of the life cycle of the parasite (sporogony) depends on a high ambient temperature.

[6] Only 4 species of plasmodium can infect human beings: Plasmodium vivax, P. malariae, P. ovale and P. falciparum. Among these, P. falciparum is the most fatal species.

controlled for, including malaria risk (or instrumented malaria risk) in the analysis should not reject or diminish the conclusions reached by AJR (2001) if their theory is correct.

For institutional variables, we choose various measures of property rights, institutional restrictions on government power, and the rule of law to proxy institutions. In particular, three measures are adopted as a proxy for quality of institutions: (i) the strength of protection against expropriation risk published by the Political Risk Services Group, (ii) the Institutional Index proposed by Kaufmann et al. (2004), and (iii) a measure of the rule of law contained in the institutional index measure. All are widely used in the literature as proxies for institutions (see AJR, 2001; Easterly and Levine, 2003; Rodrik et al., 2004). Protection against expropriation risk published by Political Risk Services Group measures the differences in institutions originating from different types of states and state policies (this index ranges from 1 to 10, a higher value indicates less expropriation risk). The Institution Index proposed by Kaufmann et al. (2004) is a composite indicator several elements that capture the strength of property rights protection, government effectiveness, political stability, and so on. The rule of law index is one indicator in the Institution Index: It measures the protection of people and property against violence or theft, the independence and effectiveness of judges and contract enforcement. Both the Institution Index and the Rule of Law variables assume values between -2.5 and 2.5, with higher scores indicating better institutional quality.

Since these proxies for geography and institutions are all subject to endogeneity problems, proper identification strategies to establish causal relations are crucial. For example, there is a strong probability of reverse causality from economic development to disease controls and the population elimination effect of diseases. Although malaria is region specific, a country with more advanced technology and higher productivity could, without doubt, control transmission vectors more efficiently, thus lowering malaria risk (Acemoglu and Johnson, 2007; Bleakley, 2010). Similarly, since establishing and enforcing property rights and the rule of law are very costly, sound institutions may emerge only when their benefits outweigh their costs (North, 1981). As a result, the more developed economies are more inclined to establish better and more sophisticated institutions.[7]

Hence, we rely on instrumental variables to control for the possibility of reversed causality. The instrumental variables that we adopt are widely used in the existing literature, which greatly facilitates comparison between our work and the existing literature. In particular, we rely on two widely accepted instrumental variables to mitigate the endogeneity problem of institutions: the settlement mortality rate constructed by AJR (2001) and the legal origin series constructed by La Porta et al. (1998). As mentioned previously, the underlying logic for settlement mortality rate to serve as a valid instrument for institutions is that the colonization strategy was influenced by the feasibility of settlements. In places where the disease environment was not favorable to European settlers, the cards were stacked against the creation of neo-Europes and the formation of the extractive state was more likely. Assuming that institutions, once established, tend to persist through history and that settlement mortality rate can affect development only through this channel, it is thus a reasonable instrument for institutions and hence its popularity in the existing literature.

[7] Property rights may not mean much for countries without industrial capital and modern financial wealth.

However, an apparent drawback of the settlement mortality rate is that it inevitably contains information on disease itself (such as malaria). To mitigate this shortcoming, we also adopt the legal origin series as an alternative instrument for institutions. Roughly speaking, there are two main legal systems in the world: common law and civil law. There is a consensus that common law, which originated in England, affords stronger protection over property rights (La Porta et al., 2008). The legal origin index classifies countries based on the origin of their legal system. An additional benefit of this alternative instrument variable is that for many countries, the legal system is imposed by the colonizers and thus is relatively exogenous.

To mitigate the endogeneity problem for malaria risks, we need a source of variations that contribute to malaria transmission but are not connected to economic development other than through this channel. We know that malaria is transmitted by mosquitos, but most mosquitos are immune to the plasmodium parasite and only some species of mosquitos of the genus anopheles transmit malaria. In addition, some anopheles species, especially those in sub-Saharan Africa, show a high preference for blood meals from humans (anthropophagy) as opposed to animals such as cattle (Sachs, 2003). Based on these facts, Kiszewski et al. (2004) constructed an ecologically based variable, called malaria ecology (ME), that can predict malaria risk. This instrumental variable includes information on the vector type, species abundance and temperature that are exogenous to public health interventions and economic conditions. However, for robustness analysis, another instrument for malaria risk is also adopted—the share of a country's population in temperate ecozones, as proposed by Sachs (2003).

To summarize, for our benchmark model we have three measures of institutional qualities: (i) protection against expropriation risk (*avexpr*), (ii) the institution index (*kk*), and (iii) the rule of law (*rule*) plus two measures of geography: (i) malaria risk (*mal94p*) and fatal malaria risk (*malfal94*). We have two instrumental variables for institutions: (i) settlement mortality rate (*logmort*) and (ii) the legal origin (*leg_bri*); and two instrumental variables for geography: (i) malaria ecology (*ME*) and (ii) population share of ecozones (*kgptemp*).

Our first goal is to reassess the dominant theories of cross-country income variations. In line with the existing literature, we proxy a country's income level by PPP-adjusted per capita GDP. As shown in Table 1, the discrepancy of income across countries is conspicuous. In 2010, the average income of the world is about $5,000. For agrarian countries and industrial countries, however, the mean is approximately 1,700$ and 20,000$, respectively, so the average gap is more than 10-fold. Moreover, the standard deviations of income levels across countries are large. In terms of log income, the standard deviation is 1.0 for agrarian nations, 0.8 for industrial nations, and 1.5 for the full country sample. Our first goal is to reassess the explanatory powers of institutions and geography on such income gaps and variations across countries.

The second goal of this paper is to investigate the causes of industrialization. The proxy for industrialization is a dummy variable based on a country's AVA-to-GDP ratio. Countries with a high AVA-to-GDP ratio (say above 10%) are classified as agrarian and those with an AVA-to-GDP ratio equal or below the cutoff are classified as industrial. The 10% cutoff is our rule of thumb but nonetheless arbitrary, so we also conduct sensitivity analyses for other cutoff values in Section 4. We show that adjusting the cutoff value up or down does not affect our results. As an additional robustness analysis, we also use the AVA-to-GDP ratio itself as

our proxy for industrialization (or the degree of industrialization), instead of the dummy variable. Furthermore, we also use rural population share as the proxy for industrialization (or the degree of industrialization), and our results remain robust.

Our final goal is to show that institutions, although important for explaining income variations among industrialized economies, are endogenous to development stages. And we study the determinants of institutional qualities.

Table 1 provides summary statistics for the main variables used in our analyses, such as the number of observations, the mean, and the standard deviation (SD). Column (1) pertains to the full sample, column (2) the agrarian country sample, and column (3) the industrial country sample. Notice that missing values for some variables (especially the settlement mortality rate) cause the changes in the number of observations across these data samples.

[Insert Table 1 here]

4. Analytical Framework and Main Results

4.1. Benchmark Model

Our benchmark model for explaining income variations across countries (both the full sample and the subsamples) is specified as follows:

$$\log y_i = u + \alpha \mathrm{INS}_i + \beta \mathrm{GEO}_i + \gamma \mathrm{X}_i + \epsilon_i, \tag{1}$$

where y_i is income (PPP measured GDP per capita) of country i, INS_i and GEO_i are measurements of institutions and geography, respectively, X_i contains control variables and ϵ_i is a random disturbance term. Naturally, α and β are coefficients of interest. We adopt a standard two-stage least squares (2SLS) regression with instrumental variables to estimate the above model.

To study the causes of industrialization or the factors determining the presence/absence (or degree) of industrialization, we simply replace the dependent variable in Equation (1) with the proxy of industrialization (e.g., the AVA-to-GDP ratio or the industrialization dummy, which takes the value of 1 for agrarian countries and 0 otherwise). That is,

$$R_i = u + \alpha \mathrm{INS}_i + \beta \mathrm{GEO}_i + \gamma \mathrm{X}_i + \epsilon_i, \tag{2}$$

where R_i is the proxy for industrialization. The explanatory variables on the right-hand side of equation (2) are identical to those in equation (1).

Before we present the detailed results, it is useful to gain some intuition and anticipate our results by exploring the bivariate relationship between institutions (or geography) and income. Figure 1 plots institution (and geography) against log income. The top panels pertain to the two different institutional measures: the rule of law (*rule*) and protection against expropriation risk (*avexpr*); the bottom panels pertain to the two different geography measures: malaria risk (*mal94*) and fatal malaria risk (*malfal94*). All panels correspond to the full sample. All plots show a clear positive (negative) correlation between institutions

(geography) and income. Thus, both institution and geography have the potential to explain the cross-country income variations.

[Figure 1 about here]

Figures 2 and 3 plot the same relationships for the subsamples. The right panel in Figure 2 confirms a strong positive relationship between institution (e.g., rule of law) and income for the industrial countries, but the left panel indicates a much weaker relationship for the agrarian countries. The same pattern holds true for the other alternative measures of institutions.

[Figure 2 about here]

In contrast, the left panel in Figure 3 shows a strong negative relationship between malaria risk and income for the agrarian countries, while the right panel shows that this correlation is much weaker for the industrial countries. In particular, while most agrarian countries show high malaria risk, only 3 industrialized countries show relatively high malaria risk (greater than 0.5).

[Figure 3 about here]

The above patterns are further confirmed by simple OLS regressions of equation (1) with both the full sample and the subsamples. In Table 2, columns (1) to (3) correspond to the full sample regression, columns (4) to (6) correspond to the industrialized countries sample, and columns (7) to (9) correspond to the agrarian countries sample. The full sample regression shows that both institutions and geography are significant in explaining cross-country income differences, confirming results in the existing literature. However, the subsample regressions show that institutions are significant only for industrial countries, while geography is clearly significant for agrarian countries but the effect of institutions is ambiguous. For example, the coefficient of rule of law (*rule*) is around 0.9 in both the full sample and the industrial countries sample, but it shrinks to 0.36 in the agrarian countries sample (and the SD nearly doubles). More importantly, the coefficient of expropriation risk (*avexpr*) (employed by Acemoglu and Robinson in their classical analysis) completely loses its significance in the agrarian countries sample.[8]

[Table 2 about here]

We now address the possible endogeneity problem by using 2SLS to re-estimate the models. The identification involves using (i) settlement mortality rate and legal origin as instruments for institutional quality and (ii) malaria ecology and climate zone as instruments for malaria risk. In the first-stage regression, the proxies for institutional quality and geography are each regressed on all 4 exogenous (instrument) variables (or a subset of it) to obtain:

$$\mathrm{INS}_i = u + \boldsymbol{Z_i\theta} + \epsilon_i^{\mathrm{INS}}, \tag{3}$$

$$\mathrm{GEO}_i = u + \boldsymbol{Z_i\delta} + \epsilon_i^{\mathrm{GEO}}, \tag{4}$$

[8] We obtain almost identical results for the other proxies of institutions or geography.

where $\boldsymbol{Z_i}$ is a vector of the 4 exogenous instrumental variables (2 for institutions and 2 for geography). The predicted values of INS_i and GEO_i from Equations (3) and (4) are then used as the independent variables in Equations (1) and (2) in the second-stage OLS regressions.

Since we have 3 proxies for institutions and 2 for geography, in principle we can have 6 different combinations of institutions and geography (or 6 possible institution-geography pairs) to conduct our analysis, given the 4 instrumental variables. However, one of the most important instrumental variables for institutions, the settlement mortality risk, has a short sample size (only about half of the sample size of the instruments for geography), we also report results when only the geographic variable is instrumented.

4.2. Determinants of Income Differences

4.2.1. Full Sample Analysis

To facilitate comparison with the existing literature, we proceed as follows. First, we run our model with the full sample without distinguishing between agrarian and industrial economies. We then run our model with subsamples and show the dramatically different conclusions that can emerge from the analysis.

Table 3 shows results for 9 different specifications for the full sample. Columns (1) to (3) pertain to the specification where only geography is instrumented. This specification not only permits a significantly larger sample size but also goes against the geographic hypothesis because instrumenting institutions tends to magnify their significance and create downward bias on geography.[9] For each of the three columns, the independent variable for geography is malarial risk (*mal94p*). However, the independent variable for institutional quality differs across the three columns: changing from protection against expropriation risk (*avexpr*) in column (1) to the rule of law (*rule*) in column (2) and to the institutional index (*kk*) in column (3). In line with the results of Sachs (2003), all specifications from columns (1) to (3) show that geography is highly significant in explaining global income differences even after controlling for institutional quality.

[Table 3 about here]

Columns (4) to (6) in Table 3 pertain to the same specifications in columns (1) to (3) except here both geography and institutions are instrumented, as in equations (3) and (4). The results based on the full set of instruments are the same as in columns (1)-(3), namely, geography is highly significant even after controlling for institutions. We can compare our results with those of AJR (2001). In their baseline model with expropriation risk as the proxy for institutions (instrumented by the settlement mortality rate), the coefficient of institution is 0.94 and the R-square is 0.27 (see column (1) of Table 4 of their paper). In contrast, when geography (malaria risk) is controlled for, the coefficient of expropriation risk is reduced sharply to around 0.43 (see column (1)) and the R-square is increased sharply to 0.74.

Consider two typical countries in the sample, Nigeria and Chile. Based on AJR (2001), institutions account for 206 log points, or more than 80% of the log income difference

[9] The institutional variable can have a stronger correlation with the dependent variable than does the instrument variable.

between Nigeria and Chile when expropriation risk serves as the proxy for institutions (in their data, the log income difference between the two countries is 253 log points for the year of 1995, slightly larger than our income year of 2010). However, based on our results in column (4) of Table 3, institutions now only account for 94 log points, or 43% of the log income differences (in our data, the log income difference between these two countries is 218 log points in 2010). Geography, on the other hand, accounts for 192 log points or 88% of the income difference between Nigeria and Chile, making geography appear more important than institutions in explaining the income gap. Also notice that the coefficient of geography (malaria risk) increases further in columns (5) and (6) when the rule of law (*rule*) or the institutional index (*kk*) are used instead as the proxy for institutions, compared with column (4) (as well as with columns (2) and (3), respectively).

The conclusion from specifications in columns (4)-(6) in Table 3 (or columns (1)-(3)) is thus clear: If the influence of geography on development is ignored, too much explanatory power on the cross-country income differences would be incorrectly attributed to institutions. This omitted variable bias is a general shortcoming of the existing literature that argues for the importance of institutions in long-run economic growth and development, consistent with the findings of Gallup and Sachs (2001) and McCord and Sachs (2013), among others.

Columns (7) to (9) check the robustness of the results reported in columns (4) to (6) by replacing malaria risk (mal94p) with fatal malaria risk (malfal94). Since a fatal disease tends to reduce the size of population in addition to decreasing labor productivity, other things equal it tends to increase per capita income. Hence, a reduction in the coefficient of geography and an increase in that of institutions may be expected. The results in columns (7) to (9) confirm such expectations. However, despite this adverse effect of fatal malaria risk on per capita income, geography is still highly significant in explaining the cross-country income variations, as under the previous specifications, without significant changes in the R-Square values—all R-square values fall within the range of [0.64, 0.75], significantly higher than the value of 0.27 obtained by AJR (2001).

Hence, the results based on the full sample under nine different model specifications all point to the fact that the existing institutional literature has overstated the significance of institutions for development. However, before proceeding further to demonstrate its demise, we discuss why so many prominent researchers found that only institutions matter (or institutions matter the most) in their empirical analyses based on data similar to ours (see, e.g., Rodrik et al., 2004). In our view the conflicts arise from several sources. The first source of conflict relates to the choices of the proxies for geography. For instance, the often-cited paper of Rodrik et al. (2004) uses latitude as the proxy for geography in their benchmark model. Latitude is an aspect of geography but certainly not representative and first-order important for explaining economic performances (Diamond, 1997). Thus, using only latitude as the proxy for geography creates measurement errors and omitted variable bias. This problem is also captured by the R-square values. The R-squares for all of our specifications presented in Table 3 are much larger than those in the preferred-model specifications of Rodrik et al. (2004), indicating that malaria risk is definitely a much better proxy for geography than latitude and enhances the explanatory power of the empirical models significantly.

Second, the econometric method matters. For instance, Easterly and Levine (2003) proxy geographic endowments by the settlement mortality rate, latitude, resources, and a landlocked

dummy and use these variables as instruments for institutions in first stage regressions. However, they fail to use geographic variables in the second-stage regression, thus putting geography at disadvantage or unfair position compared to institutions. Hence, their claim that geography only explains income through institutions is invalid because of the potential omitted variable bias.

The third source of conflict relates to differences in interpretations and subjective priors when facing the same empirical results. For example, in discussing the robustness of their results, Rodrik et al (2004, p.151) admitted that "malaria appears to have a strong, statistically significant, and negative effect on income." Yet they still decided to "attach somewhat less importance to these results" and gave three reasons for doing so: (i) Malaria is debilitating rather than fatal, so its effect should be weak. (ii) The use of ME as instrument for geography is questionable because "the original source of the index…has no discussion of exogeneity at all." And (iii), malaria is especially severe in Africa so it is difficult to separate the effect of malaria from a regional dummy. These arguments are far-fetched. First, fatal diseases, such as the Black Death, do reduce population, so diseases may contribute positively to per capita income (Acemoglu and Johnson, 2007). Malaria is debilitating but not fatal; thus, it lowers labor productivity without killing the population, so its adverse effect on per capita income should be stronger than the effect of fatal diseases. Second, the authors who constructed the ME index directly stated that "the new index will be useful in measuring the extent of causation running from malaria to poverty because the index can be used as an instrumental variable in regressions of economic growth and income levels on malaria endemicity" (Kiszewski et al., 2004). Third, in Section 4.3 we also test and reject the third argument of Rodrik et al. (2004) regarding the validity of malaria as a good proxy for geography because of its similarity to a regional dummy.

4.2.2. Subsample Analysis

We now consider the subsample analyses, one of the main focuses of our paper. Table 4 shows the results for the agrarian country sample under various specifications similar to those in Table 3. Namely, in columns (1) to (3) only geography is instrumented but institutions are not (in order to increase the sample size of the institutional instruments), whereas in columns (4) to (6) both geography and institutions are instrumented. Specifically, in columns (1) to (3) the proxy for geography is the same (malaria risk), but the proxy for institutions differs in each column, with expropriation risk (avexpr) in column (1), the rule of law (rule) in column (2), and institutional index (*kk*) in column (3). Columns (4) to (6) repeat the same exercises in columns (1) to (3) by adding the two institutional instruments as in equation (4): settlement mortality rate (*logmort*) and legal origin (*leg_bri*). The specifications in columns (7) to (9) repeat the exercises in columns (4) to (6) after replacing malaria risk with fatal malaria risk as the proxy for geography.

[Table 4 about here]

Columns (1) to (3) in Table 4 show that geography is the only significant variable in explaining cross-country income differences even though institutions are not instrumented. The coefficient of malaria risk under the three specifications is -2.36 (SD 0.87), -1.73 (SD 0.4) and -2.17 (SD 0.91), respectively; all are significant at the 1% or 5% level. In sharp contrast, the coefficient of institutions, regardless of how measured, is insignificant and sometimes

even has the wrong sign. This important pattern is unaffected when both geography and institution are instrumented. More specifically, in columns (4), (5), and (6) (parallel to columns (1), (2), and (3), respectively), malaria risk is highly significant at the 1% level, although the magnitude of its coefficient is lowered slightly to -1.69, -1.7, and -1.8, respectively, but with smaller standard errors and significantly higher R-squares. Institutions, on the other hand, remain insignificant across all specifications.

In specifications (7), (8), and (9) in Table 4, malaria risk is replaced by fatal malaria risk (*malfal94*) to proxy for geography and obtain similar results. The coefficient of fatal malaria risk is around -1.4 (with high significance at the 1% level), which is slightly lower compared with malaria risk but the R-squares remain roughly unchanged, whereas institutions remain insignificant. The reason for this slightly reduced magnitude of coefficient, as discussed earlier, might be that a more severe strain of malaria could cause more immediate deaths and offset the decline in per capita income. However, despite such an adverse effect on per capita income, the coefficient of institutions remains insignificant.

To summarize, across all specifications with various proxies for geography and institutions, geography is highly significant with stable coefficients across specifications in the agrarian country subsample. However, institutions lose their significance completely in explaining the cross-country income variations of agrarian societies compared with the full sample results in Table 3. Sometimes the coefficient of institutions even has the wrong sign, as in specifications (1), (3), and (6).

Also, the coefficient of geography is remarkably stable not only across various specifications but also across the two samples. As we move from the full sample (see Table 3) to the subsample (see Table 4), the coefficient of malaria risk (mal94p) lies constantly in the range of [-1.7, -2.0] in all specifications, indicating a robust prediction that at the margin a 1% increase in malaria risk in the population would translate into a 1.7% to 2% fall in per capita income—a huge effect. Imagine two countries, one without malaria and one with a 100% malaria risk. According to our estimates, there would be roughly fivefold to sevenfold differences in income levels caused by the presence/absence of the malaria risk.[10] For a fatal malaria risk, the difference is fourfold to fivefold.

Table 5 shows the results for the industrial countries across various specifications. The results are completely reversed compared with Table 4. For example, column (1) in Table 5 shows that the coefficient of institutions (with expropriation risk as the proxy) is 0.35 and is highly significant with a SD 0.01. The coefficient of geography (with malaria risk as its proxy), however, is only marginally significant, its magnitude shrinks by more than half compared with that in column (1) of Table 4. In column (4) of Table 5 where both geography and institutions are instrumented, malaria risk loses its significance completely in the industrial countries group. In sharp contrast, the coefficient of institutions remains highly significant at the 1% level. This pattern holds true across all specifications in columns (4) to (9) in Table 5. Namely, institutions appear to be the only significant factor in explaining income difference across industrial countries regardless of the institutional measures used. Geography, on the contrary, loses significance almost completely regardless of how it is measured and, in some

[10] The index value of malaria risk is between 0 and 1. Thus for countries without a malaria risk and with a 100% malaria risk, the income gap is $\frac{y_1}{y_2} = e^{\beta}$, roughly 5 to 7 times.

cases, even has the wrong sign (e.g., see the specifications where institutions are measured by the rule of law or the institutional index in columns (2) and (3) or (5) and (6)). This pattern is unaffected when fatal malaria risk replaces malarial risk as the proxy for geography. For example, comparing columns (4), (5), and (6) with columns (7), (8), and (9), respectively, shows that the institutional coefficient remains highly significant and barely changed when the proxy of geography switches from malaria risk to fatal malaria risk, but the geography coefficient remains insignificant and unstable (in some cases even with the wrong sign).

[Table 5 about here]

The stability of the institutional coefficient is remarkable as we move from the full sample analysis in Table 3 to the subsample analysis in Table 5. In the full sample (see Table 3) the coefficient of institutions lies in the range of [0.67, 1.12], and in the industrial sample (see Table 5) the range is [0.69, 1.17].[11] Thus, in both the full sample and the industrial countries sample, a 1unit improvement in institutional quality can generally lead to about a 0.7- to 1.2-unit difference in log income, or twofold to threefold differences in income levels. Nonetheless, even though geography is not significant in the industrial countries sample, it still reduces the effects of institutions by as much as 50% in terms of the size of the institutional coefficient, compared with the results in AJR (2001).

To conclude, the results in Tables 4 and 5 clearly show that the determinants of income are fundamentally different for countries at different developmental stages. In agrarian societies, geography rules because labor productivity in autarkic agrarian economies depends in a fundamental way on geographic conditions, such as the quality of land, climate, water resources, and diseases, rather than on cross-country institutional differences. Institutions may affect how labor is organized and incentivized through the division of labor and specialization in societies, but such effects are not important when the basic form of labor organization is the autarkic family instead of highly specialized large-scale factories and financial companies. Once industrialized, however, geography no longer constrains labor productivity because economic activities have shifted to large-scale manufacturing and modern services that depend on capital rather than on land, on electricity rather than on seasonal water flows, and on machinery technology rather than on animals.

The findings in the two subsample analyses in Tables 4 and 5 provide the explanation for the mixed results in the full sample analysis in Table 3. That is, when agrarian countries and industrial countries are combined, then both geography and institutions may appear significant in explaining development, leading to the controversy between the institutional school and the geographic school depending on the proxies and instrumental variables used. But such results are artifacts of the sample biases, driven by the questionable underlying assumption that both agrarian countries and industrial countries share the same fundamental forces of income determination. Once we abandon such an implicit assumption and allow a clear distinction between the driving forces of income determination for agrarian countries and industrial countries, the results become clearcut and the puzzles seem to disappear.

4.3. Robustness Analysis

[11] The coefficients of "protection against expropriation risk" are scaled up to match the unit of measurement for the other variables to make them comparable.

4.3.1. Alternative Cutoffs or Criteria

The first and perhaps the most important concern with our results is the cutoff (partition criterion) of agrarian and industrial countries. The 10% cutoff in the AVA-to-GDP ratio may appear sensible but somehow arbitrary.[12] Here we consider two alternative criteria. The first criterion simply raises the cutoff of the AVA-to-GDP ratio to 15%, and the second criterion uses a country's rural population ratio of 30% as the cutoff to classify agrarian/industrial countries.[13] Under these two alternative partition methods, we re-do the subsample analyses as in Tables 4 and 5. Table 6 shows the new results; columns (1) to (4) correspond to the first criterion based on the 15% AVA-to-GDP ratio cutoff, and columns (5) to (8) correspond to the second criterion based on the rural population ratio.[14] Despite the significantly different criteria (in particular, the 30% rural population share generates quite different subsamples than the AVA-to-GDP ratio), the results in Table 6 are quite similar to our previous results in that both the sign and the magnitude of the coefficients under interest are largely unaffected by the new cutoff values or the new partition methods.

[Table 6 about here]

4.3.2. Africa-Specific Effect?

Another concern of robustness, as noted by Rodrik et al. (2004), is that when malaria risk is used to proxy for geography, it is difficult to distinguish the pure effect of malaria from other nongeographic effects specific to Africa. This is a reasonable concern because malaria is a disease of warm climate and thus especially severe in Africa. We control this possible region-specific effect by adding an Africa continent dummy to the regression analysis for the agrarian country sample. Table 7 shows that adding the African dummy has little effect on the original results: The geography coefficient remains highly significant and stable across all specifications, while institutions remain insignificant in all cases. In addition, the Africa dummy shows no significance in most cases except for marginal significance at the 10% level in 1 of the 9 specifications (i.e., the first specification in column (1)).

[Table 7 about here]

4.3.3. Additional Controls

So far we have relied on a simple framework with widely used proxies for institutions/geography and their instrumental variables to establish causal relations between institutions/geography and development. These proxies and instrumental variables, however, may suffer from the omitted variable bias. Because of the potential omitted variable bias problem, the effects of other factors on economic development might be incorrectly attributed

[12] For example, consider some borderline countries. In 2013, agriculture's share in GDP was about 14% for India and Indonesia, 12% for Thailand, 11% for Sri Lanka, 10% for China, Ecuador Malaysia, and Ukraine, 9% for Georgia and Turkey, 7% for Argentina, and 6% for Brazil and Romania.

[13] Consider the same group of countries as in the previous footnote. In 2013, the rural population share was 68% for India, 48% for Indonesia, 52% for Thailand, 82% for Sri Lanka, 47% for China, 37% for Ecuador, 27% for Malaysia, 31% for Ukraine, 47% for Georgia, 28% for Turkey, 9% for Argentina, 15% for Brazil and 46% for Romania. See http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS.

[14] To conserve space, we run only 4 regressions using each alternative partition method in each sub-sample instead of the original 9 regressions.

to geography and institutions. For example, a potentially important omitted variable is international trade or the openness to trade. Sachs and Warner (1995) and Dollar and Kraay (2004) suggest that international trade can be the major source of growth for poor countries. A second omitted variable might be the identity of the main colony countries, which is emphasized as an important factor for long-term prosperity in La Porta et al. (1999) and Landes (1998). A third possible omitted variable is natural resources such as oil. We know that some countries are rich simply because of their natural endowment in oil resources. Yet another possible omitted variable is human capital.

To address these concerns about omitted variable bias, we add 7 additional control variables (one at a time) to our model to check the robustness of our previous results. Specifically, we choose control variables that are potentially correlated with both economic performances (e.g., income per capita) and institutional quality (or geographic conditions) in the baseline model reported in Tables 3 to 5. Table 8 shows the results based on the 7 additional control variables. To streamline the presentation, in Table 8 we report only the specifications where the rule of law (rule) is the proxy for institutions and malaria risk (mal94p) is the proxy for geography, and where both geography and institutions are instrumented by the full set of instrumental variables as in equations (3) and (4). Since the results based on alternative proxies of institutions/geography are broadly similar, they are not reported but are available upon request.

Table 8 has four panels (or sub-tables). Panel A pertains to the full sample and is analogous to Table 3, Panel B pertains to the agrarian country sample and is analogous to Table 4, and Panel C pertains to the industrial country sample and is analogous to Table 5 (we defer discussions of Panel D to a later section). Overall, across all panels (samples) and specifications in each panel, our previous results remain remarkably robust to the addition of extra controls. That is, in the full sample both institutions and geography appear to be significant in explaining the cross-country income differences, but in the agrarian country sample only geography is significant and in the industrial country sample only institutional quality is significant. Since each additional control also adds a story of its own, we discuss each of these cases with additional control variables (case by case) below.

Columns (1) to (3) of Table 8 (through all panels A, B and C with different sample specifications) pertains to a measure of the country's openness to trade as the additional control. We find no support for the hypothesis that international trade is important for cross-country income differences once geography and institutions are taken into account. Specifically, in column (1) we use trade share (the ratio of total trade volume to nominal GDP, all measured in U.S. dollars) as the proxy for openness. Since trade share is endogenous and can be affected by income, in column (2) we instrument the trade share by the "predicted trade shares" from the bilateral trade equation in Frankel and Romer (1999) to deal with the potential endogeneity problem of the trade share. From column (1), the trade share shows significance in explaining income discrepancy in agrarian countries (see Panel B), but is not significant in explaining income discrepancy for the industrial country sample (see Panel C) and the full sample (see Panel A). However, when the trade share is instrumented using the Frankel-Romer (1999)-type instrument variable, it loses significance completely in all three samples shown in Panels A, B, and C.

Column (3) in Table 8 replaces a country's trade share with the fraction of the country's total population living within 100 km of the coast. Shaped by geography, coastal countries usually have more advantages in international trade than landlocked countries, so coastal population share (pop100km) may be viewed as an exogenous proxy for openness. Column (3) shows that this proxy is also insignificant except for the agrarian country sample (see Panel B).

Most importantly, regardless of the significance of the control variable, controlling for a country's openness to trade in columns (1) to (3) has little effect on our previous estimates of the effects of institutions/geography on economic development. For example, in Panel A the coefficient of institutions falls in the range [0.65, 0.69], whereas that for malaria risk is in the range of [-2.13, -1.96], very similar to those reported in Table 3 and highly significant at the 1% level. Similarly, the stability of both the significance and the magnitude as well as the sign of the coefficient of geography/institutions is evident in Panels B and C. Hence, our previous results remain intact.

Columns (4) and (5) of Table 8 control for the identity of the main colonized countries. Specifically, in column (4) we add a dummy variable for British colonies and in column (5) both the British colony dummy and the French colony dummy are controlled. This addition of colony dummies has little effect on our results. In all panels A to C, the coefficient of colonial origin is insignificantly different from 0.

Column (6) in Table 8 controls for natural resources, where we added a dummy variable for major oil producers to control for the effect of natural resources on development. "Major oil producer" is defined as the top 15 net oil exporters in the world. As expected, oil producer tend to have higher per capita income and lower AVA-to-GDP ratios. However, column (6) shows that our previous results remain intact. Namely, (i) the coefficients of both institutions and geography appear significant in the full sample, (ii) only geography is significant for the agrarian country sample, and (iii) only institutions are significant for the industrial country sample. Also, the coefficient of oil is insignificant in both the agrarian countries sample and the industrial countries sample, but significant at the 5% level in the full country sample, while the magnitudes and signs of institutions/geography remain stable across all specifications.

Column 7 in Table 8 considers human capital. Human capital may be important for explaining economic development. For example, countries with lower settler mortality rates may mean there is a higher human capital stock of European settlers and their descendants, which could influence the colony's long-run economic performance (Glaeser et al., 2004). Omitting human capital thus could lead to a biased conclusion supporting the institutional theory. To address this issue, in column (7) we add the average years of schooling as a control variable to control for the channel of human capital on development. Specifically, in line with Glaeser et al. (2004), we proxy human capital by average years of schooling for the population aged 25 or above, averaging every 5 years from 1960 to 2000 from Barro and Lee (2000). Column (7) shows the results, which convey some interesting information and new findings. First, human capital is significant in explaining income variations in both the full sample and the agrarian country sample, but not in the industrial country sample. Second, institutions lose significance after controlling for human capital. In the full sample, the coefficient of institutions shrinks to about one-third of its original value and becomes insignificant (see Panel A). In the industrial countries sample, the magnitude of the

institutional coefficient remains high but the SD is 7 times larger, which renders institutions insignificant in explaining the income differences across industrial countries (see Panel C). These results suggest that institutions and human capital are deeply entangled (see Gennaioli et al., 2013, and Gallego et al., 2014, for recent discussions on this issue). Therefore, the rule of law may also capture some effects of human capital when it is instrumented by the settlement mortality rate, which may predict the human capital of the settlers. Third and most importantly, our previous finding that countries at different stages have different driving forces of growth remains robust.

Another potentially omitted variable is the influence of religion on economic development, which European settlers and colonizers brought to the New World before and after the Industrial Revolution. For example, Max Weber held that Protestant ethic is a key determinant for economic growth. Column (8) in Table 8 controls this effect by adding the fraction of Protestants in a country's population. Clearly, this has zero effect on our previous results and the influence of religion on economic development is also zero.

[Table 8 is about here]

4.3.4. Validity of Instrumental Variables and Summary

It has long been debated whether settlement mortality is a well-behaved instrument for institutions. For instance, Glaeser et al. (2004) argue that settlement mortality provides information mainly about the European settlers themselves (e.g., their human capital), rather than their home country's formal institutions. Albouy (2012) also notes some drawbacks regarding the construction method of this indicator. To avoid such criticisms, we also adopted a country's legal origin as an alternative instrument for institutions in our analyses with essentially the same results: Namely, the driving forces of economic development differ at different developmental stages. Geography determines a country's income level if it is in the agrarian stage and institutions rule if it is in the industrialized stage.

To summarize, the results in Section 4.3 and its subsections show that our conclusion reached in Section 4.2—that economies at different development stages have different determinants of national income—is robust. At the agrarian stage when land and human muscle are the two most important production factors, geography is the dominant force in determining economic performances. However, at the industrial stage when capital-intensive manufacturing is important, institutions then become the dominant factor in explaining income levels. Other factors, such as international trade, natural resources, colony experience, human capital, religion, and so on, are either unimportant or they work through the channels of geography and institutions to impact economic development. Therefore, the long-standing controversy between the institutional school and the geography school seems rooted in the implicit assumption that economies at different developmental stages share the same driving forces of economic growth. Our analyses above show robustly that such an assumption is false and soundly rejected by the data.

5. What Explains Industrialization?

Our analyses raise even more important questions: What determines a country's development stage? Or what causes industrialization? In other words, if the driving forces of development

are so different before and after industrialization, what then determines the absence/presence of industrialization? Despite the tremendous income gains from industrialization, why have only a handful of countries successfully industrialized since the English Industrial Revolution 200 years ago?

These questions are extremely important not only for understanding the mechanism of economic development in general, but also for shedding light on solving the "ultimate and elusive puzzle" in economic history that "has inspired generations of scholars to lifetimes of fruitless pursuit" (Clark, 2012, p.85): What caused the Industrial Revolution, and why did it happen first in late 18th century England rather than in 17th or 18th century China or India?

Notice that these questions are not identical to the question of what determines a country's per capita income. National or per capita income can differ significantly even among industrialized countries, such as between Spain and the United States or between Greece and Germany. But the Industrial Revolution is what really set the world apart, not merely the income levels but the mode of production. Countries without industrialization remain in the Malthusian equilibrium where Mother Nature dictates harvest, and countries with industrialization can withstand natural forces and with the possibility of permanent growth. Hence, explaining the success and failure of industrialization is far more critical and fundamental to development than merely raising per capita income levels. Countries can raise per capita income through more international trade and selling more natural resources without undergoing industrialization.

In the remainder of this paper we take a preliminary step in addressing the important question of what causes industrialization. We use the framework proposed in Equation (2) with the dependent variable replaced by an industrial dummy. In constructing this dummy variable, we define countries with an AVA-to-GDP ratio less than a cutoff value (such as 10%) as industrial countries and those above this ratio as agrarian countries. Since the independent variable is a dummy, we use the probit-IV method to estimate equation (2). Since some countries in our sample have long been industrialized (say in the 18th and 19th centuries) and we lack data on institutions that far back, we drop these countries from our sample. The remaining sample size is somewhat smaller but still significantly larger than the subsamples for agrarian or industrial countries analyzed in Tables 4 and 5.[15] The results are reported in Table 9, which now has only 7 columns instead of the 9 columns in Table 3.[16]

[Insert Table 9 Here]

Table 9 shows that across all specifications, geography is highly statistically significant in explaining industrialization, while institutions are not significant at all in explaining the absence/presence of industrialization. Consider the specifications in columns (1) through (3) first, where institutions are not instrumented, which tends to put geography at a disadvantage

[15] Hereafter we return to the full sample analysis.

[16] As a standard practice, we estimate this probit-IV model by the maximum likelihood method for its efficiency advantage. This, however, inevitably introduces computational difficulty and may cause convergence problems, especially with multiple endogenous variables (see Woodridge, 2010, for more discussion). Thus, specifications that do not lead to convergence under iterations are not reported in Table 6.

and institutions at an advantage.[17] In column (1) the coefficient of malaria risk is about -2.4 (with SD 0.4), significant at the 1% level, whereas the coefficient of protection against expropriation risk is not only insignificant but also has the wrong sign. In column (2), the rule of law shows the right sign but is again insignificant. Geography, on the other hand, has a coefficient of -1.7 (with SD 0.6) and remains significant at the 1% level. In column (3), we use the combination of institution index and fatal malaria risk. The coefficient of fatal malaria risk is -1.8, consistent with the result in column (2); the SD increases to about 1.1 but is still marginally significant at the 10% level. In column (4), geography remains significant at the 1% significance level and institutions remain insignificant. As in column (4), columns (5) through (7) report results when both geography and institutions are instrumented. Again, the institutional coefficient remains insignificant while that of geography remains significant at the 1% or 5% significance level. In column (7), the coefficient of the institutional index (kk) also has the wrong sign.

To facilitate economic interpretations, we can translate the coefficients into marginal effects on the probability of industrialization by evaluating it at some reference point. If we choose the full sample mean as our reference point (i.e. we focus on the average country), the coefficient implies that a 1-percentage-point increase in malaria risk would lead to a 0.8-percentage-point decrease in the probability of industrialization under specification (4) and a 0.63-percentage-point decrease under specification (5). These are huge effects. After all, the malarial risk is 50 percentage points larger in some areas of Africa than in Western Europe (where such risk is nearly 0). This large difference in malarial risk (50%) can translate into a 30 to 40 percent higher/lower probability of industrialization if the effects are extrapolated linearly. Columns (6) and (7) show results pertaining to fatal malaria risk: The coefficient is about -1.5 to -2 (SD 0.7). Thus, in terms of marginal effects, a 1-percentage-point increase in fatal malaria risk can lead to a 0.52 to 0.66-percentage-point decrease in the probability of industrialization for a typical country (evaluated at the sample mean), slightly smaller than the previous case. Overall, Table 6 shows that geography is the only statistically and economically significant variable in explaining industrialization, whereas institutions are insignificant both statistically and economically.

To check the robustness of our results, we raise the cutoff AVA-to-GDP ratio from 10% to 15% and find no significant changes in our results. We also replace the 10% AVA-to-GDP ratio by the 30% rural population share as the cutoff value of agrarian/industrial countries and find our results remain intact.

Considering that the probit-IV model hinges on the ad hoc cutoff values imposed, in Table 10 we instead use the AVA-to-GDP ratio directly as the dependent variable. This variable is continuous and thus indicates the extent or degree of industrialization instead of the binary stage with/without industrialization. This continuous variable also allows us to use the standard 2SLS regression method as in Table 3.

Table 10 shows that institutions continue to show no statistical significance in explaining the extent/degree of industrialization (except that the rule of law occasionally shows marginal

[17] As explained in Table 3, the uninstrumented institutional variables are more closely correlated with dependent variables due to potential endogeneity or the reverse causality problem. Such cases are nonetheless analyzed because the uninstrumented model has a much larger sample size than the instrumented model.

significance, as in columns (2) and (8)). However, geography (either malaria risk or fatal malaria risk) is always highly significant at the 1% level. This is in sharp contrast to the results in Table 3, where a PPP-based measure of per capita GDP is the dependent variable. Moreover, across various specifications from column (1) to (9), the geography coefficient is remarkably stable and consistent whereas the coefficient of institutions is unstable and even tends to have the wrong sign. Specifically, protection against expropriation risk (*avexpr*, the variable widely used in the existing literature as a proxy for institutions, as in AJR, 2001) and the institutional index (kk) are not only insignificant across all specifications, but also show the wrong sign in columns (1), (4), and (6). The rule of law (*rule*) basically has the right sign across specifications and is the only institutional variable that has high significance in columns (2) when it is not instrumented, but it also loses significance once instrumented by the settlement mortality rate and legal origin in column (5), except in column (8) where it shows marginal significance.

According to these estimates, differences of about 20 to 30 percentage points in the AVA-to-GDP ratio could occur between a country with no malaria risk and a country with 100% malaria risk. Recall that in 2013 the AVA-to-GDP ratio was about 14% for India and Indonesia; 12% for Thailand; 11% for Sri Lanka; 10% for China, Ecuador, Malaysia, and Ukraine; 9% for Georgia and Turkey; 7% for Argentina; and 6% for Brazil and Romania. Such a large magnitude means that poor geographic conditions can exert a powerful force on an agrarian society to keep it permanently in a Malthusian poverty trap and deny it any chance of industrialization. Therefore, it is geography, not institutions, that determines a country's developmental stages (as captured by the industrial dummy) or its extent/degree of industrialization (as captured by the AVA-to-GDP ratio). Such results are in sharp contrast to the claims of the institutional school (e.g., see AJR, 2001).

[Table 10 about here]

Further Robustness Analyses. To rule out potential omitted variable bias, we also add additional control variables to equation (2) as in the previous analysis in Section 4.3. Panel D of Table 8 shows that across all specifications with different control variables, geography remains the only significant variable (at the 1% significance level) in explaining the developmental stages or degrees of industrialization. Institutions, international trade, colony identities, natural resources, religion, and human capital all matter very little once geography is controlled for.

These results are surprising, especially given that the AVA-to-GDP ratio is highly correlated with per capita GDP (the correlation is -0.84) and that institutions are significant in explaining per capita income variations in the full sample. In particular, in the case of per capita GDP, we showed in Section 4 (Table 3) that both institutions and geography matter (in the full sample) in explaining cross-country income gaps, except when human capital is included as an additional control variable, in which case institutions are no longer significant but geography and human capital are significant (see Panel A of Table 8). This sharp difference in the results suggests that the AVA-to-GDP ratio contains information about a country's stage of economic development or its industrial structure that is very different from the PPP-based measure of per capita GDP. The results tell us that institutions do not matter at all for explaining a country's developmental stage or degree/extent of industrialization, even though they may help determine a country's per capita income once it is industrialized.

6. Institutions and Development: A "Chicken Versus Egg" Problem?

The result that institutions do not explain industrialization or the extent/degree of industrialization is somewhat puzzling given that institutions do appear to explain cross-country income variations for the industrial economies (i.e., for nations with an AVA-to-GDP ratio below a threshold level), and especially given that a country's AVA-to-GDP ratio is strongly correlated with its per capita income. To confirm the strong correlation between the AVA-to-GDP ratio and per capita income and their close relationships with institutions, Table 11 shows a simple OLS regression in which the extent of industrialization (AVA-to-GDP ratio or the rural population ratio) is regressed on geography and institutions (but without using instrumental variables). As the table shows, a strong negative correlation exists between institutions and the AVA-to-GDP ratio.

[Table 11 about here]

However, such a correlation does not imply causation. Our empirical analysis in Section 5 showed that institutions (once instrumented by the settlement mortality rate and/or legal origin) cannot explain a country's development stage or its lack/presence of industrialization. This motivates us to explore the possibility of reversed causality: It may be that economic development or the absence/presence of industrialization determines institutional qualities, rather than vice versa.

To test this possibility, we run regressions with institutions as the dependent variable but with the AVA-to-GDP ratio as the independent variable. To address potential endogeneity problems, we use the predicted values of the AVA-to-GDP ratio as the independent variable. Specifically, we use geography as the instrument for AVA-to-GDP ratio in a three-stage least squares (3SLS) regression. In other words, we first run 2SLS regressions (as in Table 10) and collect the predicted values of the AVA-to-GDP ratio explained by the instrumented geographic factors, and in the third stage we regress institutions on this geography-instrumented AVA-to-GDP ratio. This 3SLS regression model takes advantage of the fact that geography has already been established (after being instrumented by the instrumental variables for geography in equation (4)) as an exogenous causal factor in determining industrialization (see Section 5). In addition, geography is likely to be exogenous to institutions but may affect institutions only through its influence on development.

The third-stage regression results are reported in columns (1) to (4) in Panel A of Table 12. According to our estimates, a 1-percentage-point exogenous decrease in the AVA-to-GDP ratio can cause the rule of law indicator to increase by 0.06 points (significant at the 1% level) or increase the protection against expropriation risk indicator by more than 0.1 points (significant at the 1% level). These effects are consistent with each other since the magnitude of the latter institutional index is [0, 10], while the former is [-2.5, 2.5]).

Because it is difficult to understand how geography could shape institutions directly, it should be a valid instrument variable for industrialization (the AVA-to-GDP ratio). However, just in case this exogeneity assumption is invalid, we also add other geographic variables as controls in the 3SLS regression and report the results in columns (5) to (8) of Panel A of Table 12. Across all specifications in columns (5) to (8), the level of industrialization (*agval_mal94p_hat* or *agval_malfal94_hat*) is highly significant in explaining the quality of

institutions, even after geography itself is controlled for in the third-stage regression. Although the magnitude of the impact is reduced to about 0.04, the estimated coefficient still remains highly significant at the 1% level. This implies that if the AVA-to-GDP ratio decreases from 70% to 20% in a country, its institutional quality as measured by the rule of law would increase by 2 points. Such an increase is an extraordinary improvement in institutions because more than 100 countries are located in the interval [-1, 1] for the rule of law indicator. In other words, a 50-percentage-point decrease in the AVA-to-GDP ratio can move a country's institutional quality from -1 (the very bottom of the 100 countries) to 1 (the very top of the 100 countries). Moreover, the additional control variables (the geographic factors) are largely insignificant in explaining institutional qualities once the industrialization index is included in the independent variable, suggesting that geographic conditions affect institutional qualities only through industrialization.

Panel B of Table 12 repeats this exercise using our alternative measure of industrialization (the rural population ratio), and the results are basically the same. Combining the results in Tables 9, 10, and 12, we view them as strong evidence against the institutional school. Institutions do not rule, although they do facilitate further income growth once a country is successfully industrialized or lowers its AVA-to-GDP ratio (or rural population share) below a critical level. Our results instead favor the hypothesis that economic development or the extent of industrialization explains improvement in institutions, and institutional quality, in turn, feeds back to income levels for industrialized nations.

[Table 12 about here]

To support the position that the stage of development can further influence income levels only through better institutions (once countries are industrialized), we run the following regression in which the per capita income levels of industrial countries (in the subsample) are the dependent variable, and the instrumented AVA-to-GDP ratio and instrumented institutions are the independent variable (analogous to the subsample analysis in Table 5 but with the instrumented AVA-to-GDP ratio added as an additional control). The results show that the instrumented AVA-to-GDP ratio (or rural population ratio) is insignificant in explaining the income differences among the industrialized countries once institutions are taken into account, suggesting that once a country is industrialized, better institutional quality, although by itself unable to explain industrialization, can nevertheless feed back on development and further stimulate income growth and enhance the fruits of industrialization.

7. Conclusion and Future Research

The pursuit to uncover the driving forces behind cross-country income gaps and long-term development has sharply divided economists into two major camps: One emphasizes geographic factors, while another stresses institutions. Each school of thought has its own theoretical foundations and empirical support, but they share an implicit hypothesis in common—the driving forces remain the same regardless of a country's stage of economic development. Such an implicit hypothesis implies a theory that the process of development in human history is a continuous improvement in income levels and that structural changes do not dictate the influences of geography and institutions on national income.

This paper first tested this implicit hypothesis and found it rejected by data. The truth is instead that economies are characterized more fundamentally by their developmental stages and countries at different developmental stages have different determinants for income levels. In particular, geographic factors effectively explain the cross-country income gaps among agrarian countries, while institutional factors largely account for the income differences across industrial economies.

Such results make good economic sense. When economies remain agrarian, their technology and mode of production are not equipped to withstand natural disasters and environment conditions. Hence, natural forces such as germs are important for determining human productivity and income levels. However, once industrialized, such natural forces should matter less in determining a country's economic performance because of improvements in the mode of production, medical knowledge, and technologies. In this case, non-geographic factors such as institutions should gain importance for income levels.

However, the above analysis pushed the question only one step further and raised even more fundamental and intriguing questions: What causes industrialization in the first place? Or what determines the absence/presence of industrialization? We found that it is geography, rather than institutions, that explains the absence/presence of industrialization—the opposite of claims made by institutional theories. Geography is important not only in determining cross-country income differences in agricultural societies, but also as the ticket for entering the industrial club. Once a country is industrialized, however, institutions replace geography and become the most important driving force of further income changes.

Even more strikingly, we found that institutions are endogenous rather than exogenous to economic development. In particular, it is the absence/presence and extent of industrialization that determine the quality of institutions. In other words, good institutions are the consequence of industrialization, rather than vice versa. This finding also makes good economic sense. Institutions all have costs for their creation and enforcement: They emerge only when their benefits exceed these costs. For example, in an agrarian world where trade is confined to villages and trade volumes are limited by high transport costs, trade will be anarchic and unstructured. But when trade volumes rise (say after the great voyage and discovery of America), there is more incentive to create "free-trade" institutions (including navy forces) that facilitate (protect) long-distance trade.

Our studies thus lead to the following theory: Geography causes income divergences across agrarian societies and thus determines the likelihood of industrialization. Industrialization and subsequent development determine the quality of institutions, which in turn leads to further development and income growth in industrialized countries.

The empirical evidence presented in this paper has thus turned the institutional school upside-down. Rather than being the fundamental driving force of development in developing countries, we instead find institutions the consequence of development. This finding is consistent with the well-received view among prominent historians that changes in the 17th- and 18th-century English institutions were not the cause of the Industrial Revolution (see, e.g., Allen, 2009; Clark, 2007; McCloskey, 2010; MccLeod, 1988; Mokyr, 2008; and Pomeranz, 2000, among others). In particular, McCloskey (2010, p.320) notes that "property was very fully developed especially in land and in personal possessions" in England for centuries

before the Industrial Revolution, and that "incentives of a strictly economic sort did not change between 1000 and 1800." Therefore, the institutional changes in property rights during and after the Glorious Revolution were minimal at the best compared with the private property rights already established in England long before 1688. Pomeranz (2000) documents that private property and production factor mobility in 17th- and 18th-century China, especially its Yangtze River Delta region, were no worse protected than those in Europe and England in that period. He thus attributes the fundamental cause of the Industrial Revolution to the fortunate location and distribution of coal in England and global trade after the discovery of the New World.

However, important questions remain unanswered in this paper because regression analyses are reduced-form approaches without providing the underlining mechanisms for the established causal linkages. In particular, under what conditions geographically disadvantaged nations can become industrialized? How exactly does the stage of development improve institutional quality?

Therefore, we remain cautious and open about the implications of our findings, and our results merit further investigation and scrutiny by future research. Several additional possibilities could explain our results. First, the conventional instrumental variables for institutions (such as the settlement mortality rate used by AJR, 2001) may be invalid when geographic factors are present. Second, the conventional definitions and measures of institutions are too simplistic and thus unable to capture real business incentives and innovative activities stimulated by informal institutions (see Rodrik et al., 2004, for further discussions). Third, there may be an omitted third class of variables responsible for industrialization and economic take-off that is missing from the analysis. Possible missing variables include (i) coordinated beliefs and bourgeois spirits (Mokyr, 2009; McClosky, 2010); (ii) entrepreneurial skills and cultural genes (Clark, 2007); and (iii) good governance and correct industrial policies (which are arguably key to Japan's successful industrialization during the Meiji Restoration, South Korea's rapid growth under Park's dictatorship regime between 1960 and the late 1970s, and China's recent 35-year growth miracle despite extractive institutions).[18] These possibilities are on our agenda of future research.

[18] For detailed account and explanations of China's growth miracle and why China's development path sheds light on solving the puzzle of the Industrial Revolution, see Wen (2015).

References

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

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

Acemoglu, D., & Johnson, S. (2005). Unbundling Institutions. *Journal of Political Economy*, *113*(5), 949-995.

Acemoglu, D., Johnson, S., & Robinson, J.A. (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *American Economic Review*, *95*(3) 546-579.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a Fundamental Cause of Long-run Growth. *Handbook of economic growth*, *1*, 385-472.

Acemoglu, D., & Johnson, S. (2007). Disease and Development: The Effect of Life Expectancy on Economic Growth. *Journal of Political Economy*, *115*(6), 925-985.

Acemoglu, D., & Robinson, J. A. (2012). *Why Nations Fail: The Origin of Power, Prosperity, and Poverty*. New York: Crown Business.

Acemoglu, D., Gallego, F. A. & Robinson, J. A. (2014). Institutions, Human Capital, and Development. *Annual Review of Economics*, *6*(1), 875-912.

Albouy, D. Y. (2012). The Colonial Origins of Comparative Development: An Empirical Investigation: Comment. *American Economic Review*, *102*(6), 3059-3076.

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press.

Barro, R. J., & Lee, J. W. (2001). International Data on Educational Attainment: Updates and Implications. *Oxford Economic Papers*, *53*(3), 541-563.

Becker, S. O., & Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, *124*(2), 531-596.

Bleakley, H. (2010). Health, Human Capital, and Development. *Annual Review of Economics*, *2*(1), 283-310.

Boldrin, M., & Levine, D. K. (2008). *Against Intellectual Monopoly*. Cambridge: Cambridge University Press.

Bodin, J. (1533). *Les six livres de la République*. de l'imprimerie de Jean de Tournes.

Carstensen, K., & Gundlach, E. (2006). The Primacy of Institutions Reconsidered: Direct Income Effects of Malaria Prevalence. *World Bank Economic Review*, *20*(3), 309-339.

Clark, G. (2012). A Review Essay on The Enlightened Economy: An Economic History of Britain 1700–1850 by Joel Mokyr. *Journal of Economic Literature*, *50*(1), 85-95.

Curtin, P. D. (1989). *Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century*. New York: Cambridge University Press.

Datta, S. C., & Reimer, J. J. (2013). Malaria and Economic Development. *Review of Development Economics*, *17*(1), 1-15.

Diamond, J. (1997). *Guns, Germs and Steel: The Fates of Human Societies*. New York: W. W. Norton.

Easterly, W., & Levine, R. (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*, *50*(1), 3-39.

Frankel, J. A., & Romer, D. (1999). Does Trade Cause Growth? *American Economic Review*, *89*(3), 379-399.

Gallup, J. L., Sachs, J. D., & Mellinger, A. D. (1999). Geography and Economic Development. *International Regional Science Review*, *22*(2), 179-232.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human Capital and Regional Development. *Quarterly Journal of Economics*, *128*(1), 105-164.

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

Johnson, S., McMillan, J., & Woodruff, C. (2002). Property Rights and Finance. *American Economic Review*, *92*(5), 1335-1356.

Jedwab, R., & Moradi, A. (2012). Colonial Investments and Long-term Development in Africa: Evidence from Ghanaian Railroads. Mimeo, George Washington University.

Iyer, L., & Banerjee, A. (2005). History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *American Economic Review*, *95*(3), 1190-1213.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2004). Governance matters III: Governance Indicators for 1996, 1998, 2000, and 2002. *World Bank Economic Review*, *18*(2), 253-287.

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. (2004). A Global Index Representing the Stability of Malaria Transmission. *American Journal of Tropical Medicine and Hygiene*, *70*(5), 486-498.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1998). Law and Finance. *Journal of Political Economy*, *106*(6), 1113-1155.

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2008). The Economic Consequences of Legal Origins. *Journal of Economic Literature*, *46*(2), 285-332.

Landes, D. S. (1999). *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: W.W. Norton.

McCarthy, F. D., Wolf, H. C., & Wu, Y. (2000). Malaria and Growth. World Bank Policy Research Working Paper, No. 2303.

McCloskey, D. N. (2010). *Bourgeois Dignity: Why Economists Can't Explain the Modern World*. Chicago: University of Chicago Press.

McCord, G. C., & Sachs, J. D. (2013). Development, Structure, and Transformation: Some Evidence on Comparative Economic Growth (No. w19512). National Bureau of Economic Research.

Mokyr, J. (2006). The Great Synergy: The European Enlightenment as a Factor in Modern Economic Growth. *Understanding the Dynamics of a Knowledge Economy* (pp. 7-41). Cheltenham: Edward Elgar Publishing.

Mokyr, J. (2008). The Institutional Origins of the Industrial Revolution. *Institutions and Economic Performance* (pp. 64-119), Cambridge: Harvard University Press.

Mokyr, J. (2009). The Enlightened Economy: British and the Industrial Revolution 1700-1850. New Haven, CT: Yale University Press.

Montesquiedeu, C. S. (1748). *The Spirit of the Laws*. Digireads. com Publishing.

Myrdal, G. (1968). *Asian Drama, an Inquiry into the Poverty of Nations*. New York: Pantheon.

North, D. C. (1981). *Structure and Change in Economic History*. New York: W. W. Norton.

Rodrik, D. (1999). Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses. *Journal of Economic Growth*, *4*(4), 385-412.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*(2), 131-165.

Pomeranz, K. (2000). *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton, NJ: Princeton University Press.

Sachs, J. D. (2001). Tropical Underdevelopment (No. w8119). National Bureau of Economic Research.

Sachs, J., & Malaney, P. (2002). The Economic and Social Burden of Malaria. *Nature*, *415*(6872), 680-685.

Sachs, J. D. (2003). Institutions Don't Rule: Direct Effects of Geography on Per Capita Income (No. w9490). National Bureau of Economic Research.

Shiue, C. H., & Keller, W. (2007). Markets in China and Europe on the Eve of the Industrial Revolution. *American Economic Review*, *97*(4), 1189-1216.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edwin Cannan's annotated edition.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, *70*(1), 65-94.

Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association*, *8*(4), 677-716.

Wen, Y. (2015). The Making of an Economic Superpower: Unlocking China's Secrets of Rapid Industrialization. Mimeo, Federal Reserve Bank of St. Louis.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT press.

Figures and Tables

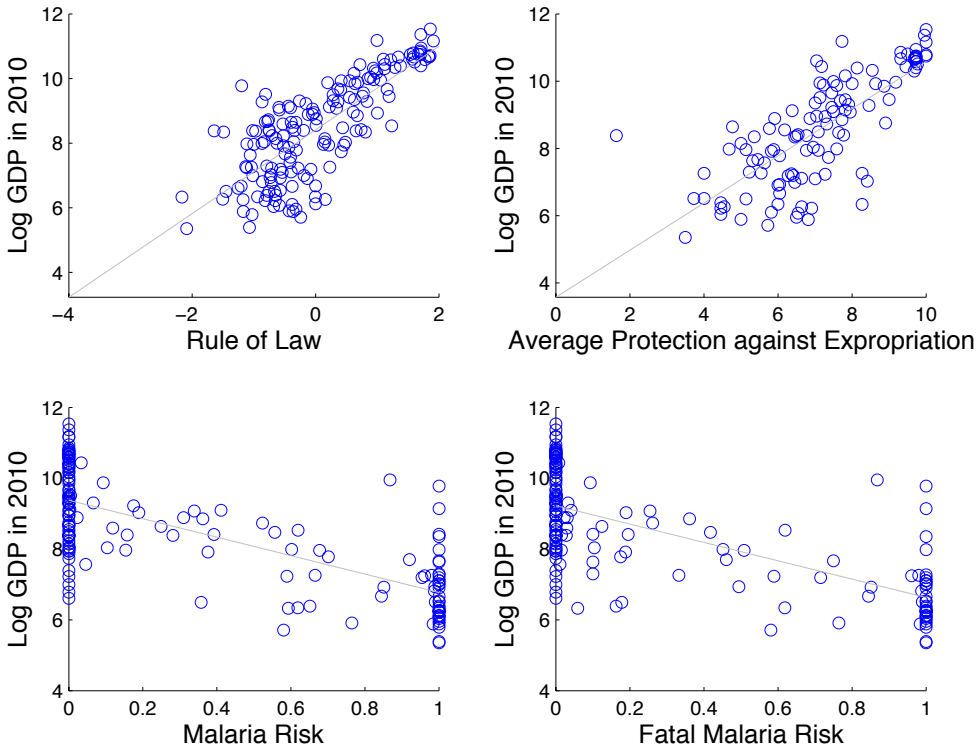

Figure 1. Full-sample scatterplot of institutions, geography, and log GDP per capita in 2010.

The top panels plot the bivariate relationship between institutional variables (rule of law (left) and protection against expropriation (right)) and log GDP per capita in 2010. The bottom panels plot the bivariate relationship between geographic variables (malaria risk (left) and fatal malaria risk (right)) and log GDP per capita in 2010. One circle represents a country observation.

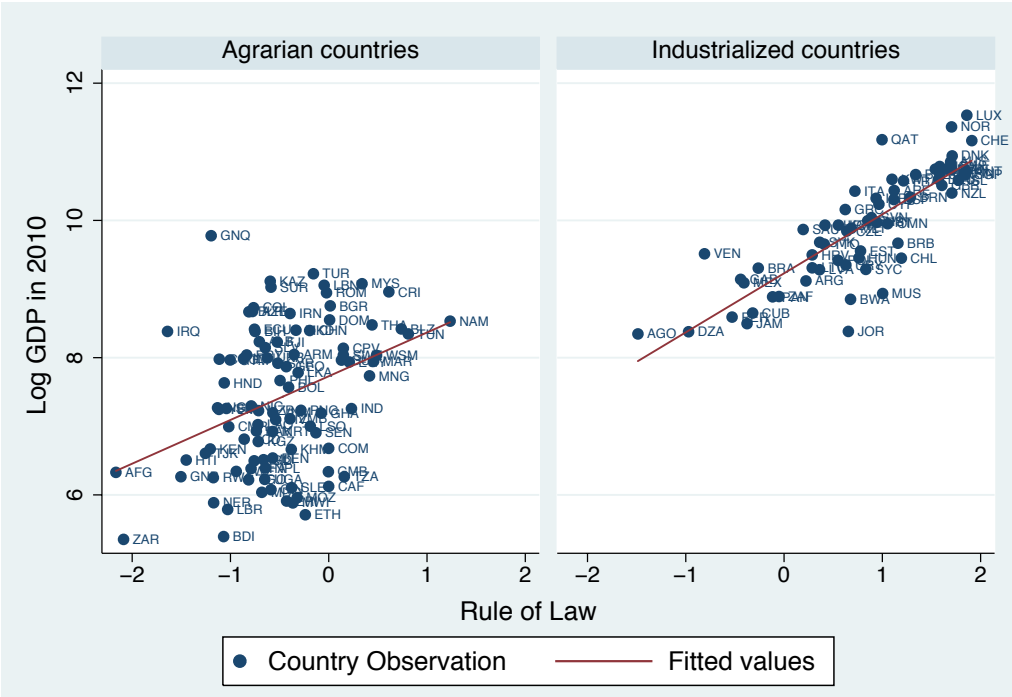

Figure 2. Subsample scatter plot of rule of law and log income.

The left panel shows the bivariate relationship of the rule of law and log income in the agrarian countries sample. The right panel shows the bivariate relationship of the rule of law and log income in the industrial countries sample. The subsamples are divided based on the AVA-to-GDP ratio with a cutoff of 10%. Countries are labeled by the International Standards Organization (ISO) 3-digit alphabetic codes.

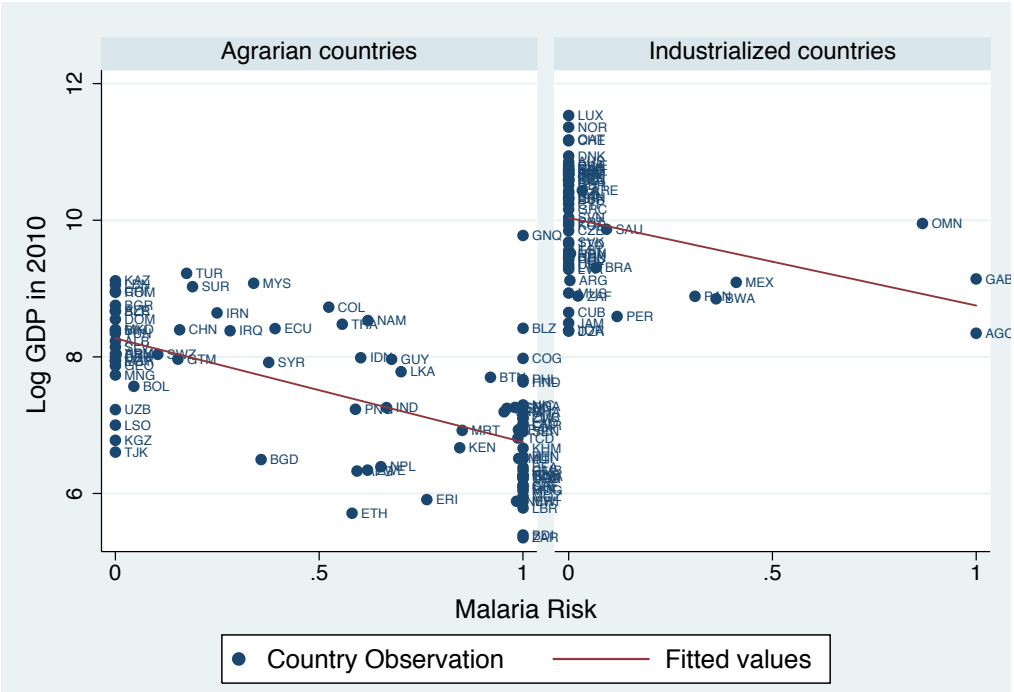

Figure 3. Subsample scatterplot of malaria risk and log income.

The left panel shows the bivariate relationship of malaria risk and log income in the agrarian countries sample. The right panel shows the bivariate relationship of malaria risk and log income in the industrial countries sample. The subsamples are divided based on the AVA-to-GDP ratio with a cutoff of 10%. Countries are labeled by the International Standards Organization (ISO) 3-digit alphabetic codes.

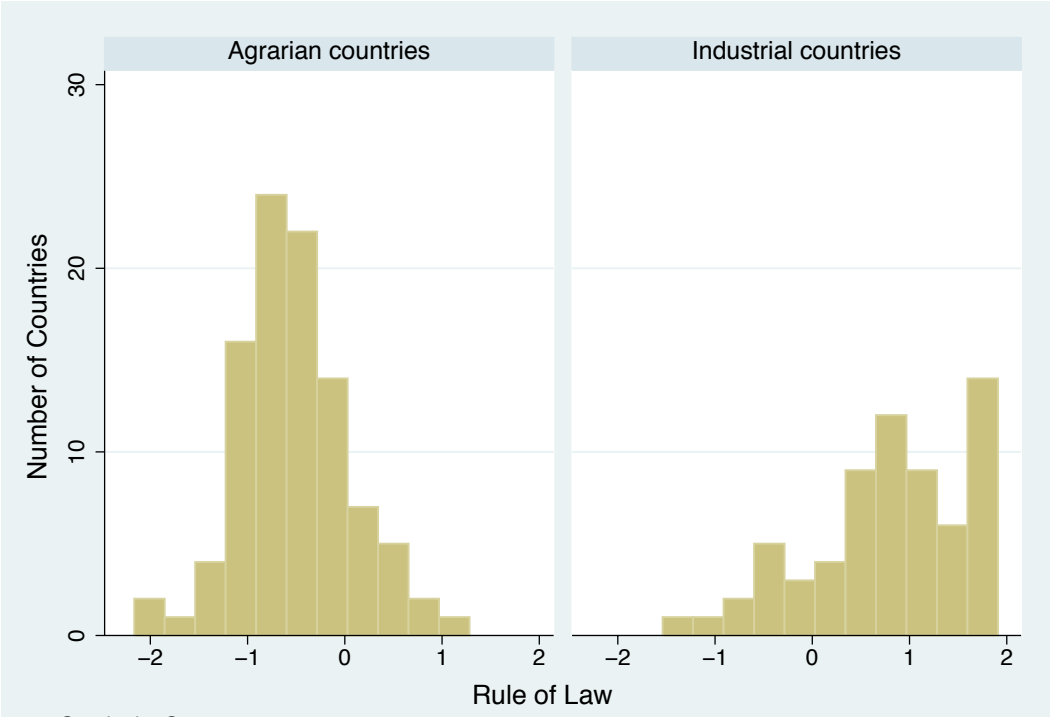

Figure 4. Histogram of institutional quality in subsamples

The left panel plots the histogram of the rule of law index in the agrarian countries sample. The right panel plots the histogram of rule of the law index in the industrial countries sample. The subsamples are divided based on the AVA-to-GDP ratio with a cutoff of 10%.

Table 1. Summary Statistics

	Full sample		Agrarian countries		Industrialized countries	
	Obs.	Mean/SD	Obs.	Mean/SD	Obs.	Mean/SD
Log GDP per capita (2010, PPP measured)	184	8.507	100	7.433	66	9.922
(*loggdp10*)		(1.538)		(1.023)		(0.804)
AVA-to-GDP ratio (average of 1990 and	173	18.40	104	27.50	69	4.681
2000) (*agval*)		(15.27)		(13.22)		(2.639)
Protection against expropriation risk (average	117	7.104	62	6.047	53	8.295
of 1990 and 2000) (*avexpr*)		(1.781)		(1.401)		(1.392)
Institutions index (*kk*)	72	-0.0905	48	-0.346	23	0.395
		(0.684)		(0.445)		(0.802)
Rule of law (*rule*)	171	0.00804	98	-0.523	66	0.788
		(0.939)		(0.590)		(0.812)
Malaria risk (*mal94p*)	165	0.361	94	0.576	61	0.0704
		(0.436)		(0.431)		(0.219)
Fatal malaria risk (*malfal94*)	165	0.306	94	0.488	61	0.0572
		(0.420)		(0.441)		(0.212)
Settlement mortality rate (*logmort*)	85	4.591	57	5.036	26	3.737
		(1.302)		(1.091)		(1.266)
Legal origin (*leg_bri*)	201	0.303	102	0.304	69	0.290
		(0.461)		(0.462)		(0.457)
Population ratio in temperate climate zone	145	0.337	88	0.153	54	0.632
(*kgptemp*)		(0.434)		(0.326)		(0.428)
Malaria ecology (*ME*)	173	3.656	101	5.572	62	0.812
		(6.416)		(7.559)		(2.448)

Notes: Data sources are as follows. GDP per capita and agriculture value added are from the World Bank database. The rule of law index and institutions index are from Kaufman et al. (2004). Protection against expropriation risk index and settlement mortality rate are from AJR (2001). Malaria risk, fatal malaria risk, and malaria ecology are from Sachs (2003) and Kiszewski et al. (2004). Legal origin is from La Porta et al. (1998).

Table 2. Results of Simple OLS regressions

	Full sample			Industrial countries			Agrarian countries		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)								
Rule	0.916***		0.969***	0.879***		0.877***	0.361**		0.369**
	(0.0817)		(0.0782)	(0.0782)		(0.0757)	(0.149)		(0.143)
Avexpr		0.420***			0.509***			0.0980	
		(0.0547)			(0.0590)			(0.0703)	
mal94p	-1.517***	-1.911***		0.0958	-0.215		-1.396***	-1.685***	
	(0.179)	(0.223)		(0.291)	(0.333)		(0.204)	(0.234)	
malfal94			-1.513***			0.0914			-1.436***
			(0.178)			(0.291)			(0.190)
Constant	8.973***	6.307***	8.887***	9.219***	5.751***	9.222***	8.382***	7.830***	8.285***
	(0.0920)	(0.454)	(0.0847)	(0.0946)	(0.513)	(0.0904)	(0.147)	(0.500)	(0.131)
Observations	155	110	155	60	49	60	90	59	90
R-squared	0.739	0.760	0.740	0.725	0.677	0.725	0.436	0.539	0.475

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 3. Results of 2SLS regressions: Full Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria ecology (ME), population ratio in temperate zone (kgptemp), settlement mortality rate (logmort) and legal origin (leg_bri)					
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)								
avexpr	0.354**			0.427***			0.591***		
	(0.144)			(0.149)			(0.125)		
mal94p	-2.287***	-1.850***	-1.868***	-1.921***	-2.064***	-1.970***			
	(0.530)	(0.399)	(0.598)	(0.463)	(0.433)	(0.507)			
rule		0.903***			0.667***			0.924***	
		(0.240)			(0.224)			(0.184)	
kk			0.904*			0.790**			1.117***
			(0.463)			(0.363)			(0.308)
malfal94							-1.391***	-1.582***	-1.474***
							(0.364)	(0.335)	(0.398)
Constant	6.878***	9.055***	8.991***	6.127***	9.152***	9.024***	4.665***	8.841***	8.703***
	(1.216)	(0.150)	(0.271)	(1.184)	(0.21)	(0.245)	(0.944)	(0.157)	(0.181)
Observations	104	141	62	66	74	62	66	74	62
R-squared	0.744	0.750	0.634	0.726	0.687	0.637	0.716	0.694	0.637

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 4. Results of 2SLS regressions: agrarian country sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria ecology (ME), population ratio in temperate zone (kgptemp), settlement mortality rate (logmort) and legal origin (leg_bri)					
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)								
avexpr	-0.132			0.13			0.0853		
	(0.483)			(0.244)			(0.234)		
mal94p	-2.362***	-1.727***	-2.173**	-1.692***	-1.702***	-1.809***			
	-0.873	(0.567)	(0.915)	(0.445)	(0.428)	(0.579)			
rule		0.398			0.19			0.339	
		(1.276)			(0.375)			(0.33)	
kk			-0.94			-0.057			0.242
			(1.631)			(0.795)			(0.658)
malfal94							-1.455***	-1.334***	-1.373***
							(0.35)	(0.306)	(0.394)
Constant	9.643***	8.524***	8.292***	7.597***	8.441***	8.384***	7.553***	8.133***	8.065***
	(3.451)	(0.438)	(0.372)	(1.67)	(0.269)	(0.28)	(1.544)	(0.21)	(0.214)
Observations	58	85	44	45	53	44	45	53	44
R-squared	0.422	0.515	0.295	0.532	0.51	0.49	0.585	0.569	0.57

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 5. Results of 2SLS regressions: industrialized country sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria ecology (ME), population ratio in temperate zone (kgptemp), settlement mortality ratio (logmort) and legal origin (leg_bri)					
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)								
avexpr	0.345***			0.516***			0.530***		
	(0.0999)			(0.0659)			(0.0616)		
mal94p	-0.996*	0.066	0.893	-0.368	0.394	0.821			
	(0.523)	(0.558)	(0.664)	(0.307)	(0.432)	(0.594)			
rule		0.878***			0.840***			0.805***	
		(0.198)			(0.13)			(0.112)	
kk			1.219***			1.174***			1.083***
			(0.319)			(0.244)			(0.19)
malfal94							-0.344	0.282	0.625
							(0.287)	(0.371)	(0.452)
Constant	7.141***	9.215***	8.930***	5.557***	9.186***	8.953***	5.431***	9.224***	9.035***
	(0.868)	(0.194)	(0.212)	(0.538)	(0.127)	(0.185)	(0.494)	(0.106)	(0.137)
Observations	45	53	18	21	21	18	21	21	18
R-squared	0.638	0.762	0.687	0.865	0.809	0.708	0.864	0.814	0.756

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 6. Robustness check: different sample partition methods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Agrarian Countries							
	Cutoff of AVA-to-GDP ratio = 15%				Cutoff of rural population ratio = 30%			
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)							
mal94p	-1.879***		-2.004***		-1.844***		-2.104***	
	(0.662)		(0.645)		(0.494)		(0.714)	
rule			-0.293	0.38			-0.376	0.517
			(0.749)	(0.522)			(1.012)	(0.67)
kk	-0.154	0.406			-0.0761	-0.0772		
	(0.887)	(0.652)			(0.884)	(0.851)		
malfal94		-1.253***		-1.258***		-1.595***		-1.358***
		(0.371)		(0.333)		(0.411)		(0.404)
Constant	8.346***	7.981***	8.292***	8.025***	8.442***	8.139***	8.422***	8.314***
	(0.373)	(0.269)	(0.44)	(0.341)	(0.268)	(0.251)	(0.312)	(0.279)
Observations	39	39	46	46	48	48	57	57
R-squared	0.39	0.541	0.276	0.504	0.515	0.55	0.408	0.543
	Panel B: Industrial Countries							
	The Dependent Variable: Log GDP per capita in 2010 (PPP measured)							
mal94p	0.957		0.282		3.665		4.101	
	(0.636)		(0.499)		(3.965)		(4.908)	
rule			0.999***	0.981***			1.147*	0.938***
			(0.193)	(0.176)			(0.55)	(0.304)
kk	1.355***	1.296***			1.459**	1.304**		
	(0.308)	(0.271)			(0.594)	(0.431)		
malfal94		0.849		0.242		6.864		5.927
		(0.544)		(0.453)		(7.351)		(7.08)
Constant	8.740***	8.799***	8.949***	8.969***	8.552***	8.760***	8.800***	9.080***
	(0.201)	(0.167)	(0.161)	(0.14)	(0.671)	(0.452)	(0.691)	(0.362)
Observations	23	23	28	28	13	13	16	16
R-squared	0.579	0.6	0.613	0.618	0.293	0.462	0.211	0.533

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 7. Robustness check: adding Africa regional dummy to agrarian sample analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria ecology (ME), population ratio in temperate zone (kgptemp) and settlement mortality (logmort)					
	Dependent Variable: Log GDP per capita in 2010 (PPP measured)								
avexpr	-0.101			-0.295			0.0807		
	(0.450)			(0.916)			(0.508)		
mal94p	-2.200***	-1.589**	-2.045**	-2.246*	-1.999***	-1.863***			
	(0.805)	(0.649)	(0.832)	(1.292)	(0.656)	(0.685)			
afri	-0.413*	-0.458	-0.300	-0.376	-0.153	-0.276	-0.105	-0.161	-0.199
	(0.212)	(0.407)	(0.244)	(0.578)	(0.234)	(0.218)	(0.329)	(0.202)	(0.2)
rule		0.520			-0.2			0.388	
		(1.323)			(0.701)			(0.48)	
kk			-0.686			-0.168			0.391
			(1.459)			(0.951)			(0.717)
malfal94							-1.448**	-1.291***	-1.276***
							(0.585)	(0.365)	-0.422
Constant	9.492***	8.663***	8.411***	10.65	8.473***	8.477***	7.610**	8.187***	8.138***
	(3.224)	(0.527)	(0.345)	(6.488)	(0.3)	(0.292)	(3.433)	(0.243)	(0.219)
Observations	58	85	44	45	53	44	45	53	44
R-squared	0.499	0.559	0.400	0.196	0.434	0.499	0.586	0.576	0.577

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 8. Robustness check: additional controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Control for log openness	Control for log openness, along with its instrument of Frankel and Romer (1999)	Control for share of the national population living within 100km of the coast	Control for British colony	Control for British and French colony	Control for major oil producer	Control for years of schooling in population above aged 25	Control for Protestants as % of population
	Panel A: Full Sample. Dependent Variable: Log GDP Per Capita 2010 (PPP Measured)							
rule	0.688***	0.652***	0.687***	0.753***	0.776***	0.589***	0.160	0.585**
	(0.238)	(0.242)	(0.221)	(0.221)	(0.227)	(0.208)	(0.286)	(0.237)
mal94p	-2.132***	-2.043***	-1.957***	-1.935***	-1.794***	-2.144***	-1.608***	-2.170***
	(0.462)	(0.481)	(0.459)	(0.417)	(0.478)	(0.411)	(0.474)	(0.445)
lcopen	0.315	-0.144						
	(0.202)	(0.320)						
pop100km			0.251					
			(0.317)					
coluk				-0.132	-0.233			
				(0.217)	(0.248)			
colfr					-0.251			
					(0.275)			
oil						0.919**		
						(0.375)		
tyr60							0.264**	
							(0.119)	
prot								0.00676
								(0.00721)
Constant	8.004***	9.682***	8.987***	9.148***	9.179***	9.108***	7.830***	9.118***
	(0.754)	(1.157)	(0.312)	(0.210)	(0.205)	(0.209)	(0.689)	(0.214)
Observations	73	73	74	74	74	74	52	74
R-squared	0.686	0.672	0.695	0.687	0.697	0.714	0.882	0.690
	Panel B: Agrarian Country Sample. Dependent Variable: Log GDP Per Capita 2010 (PPP Measured)							
rule	0.264	0.143	0.243	0.400	0.490	0.0800	0.245	0.201
	(0.390)	(0.417)	(0.341)	(0.321)	(0.313)	(0.354)	(0.230)	(0.376)
mal94p	-1.804***	-1.751***	-1.416***	-1.523***	-1.360***	-1.892***	-1.315**	-1.655***
	(0.460)	(0.489)	(0.408)	(0.398)	(0.434)	(0.409)	(0.541)	(0.449)
lcopen	0.453**	-0.0429						
	(0.220)	(0.376)						
pop100km			0.683**					
			(0.311)					
coluk				-0.0372	-0.152			
				(0.210)	(0.259)			
colfr					-0.206			
					(0.265)			
oil						0.756		
						(0.713)		
tyr60							0.297**	
							(0.120)	
prot								-0.00329
								(0.00829)
Constant	6.845***	8.626***	7.999***	8.447***	8.489***	8.497***	7.559***	8.437***
	(0.855)	(1.368)	(0.359)	(0.276)	(0.284)	(0.275)	(0.642)	(0.271)
Observations	52	52	53	53	53	53	33	53
R-squared	0.551	0.502	0.576	0.512	0.513	0.504	0.741	0.514
	Panel C: Industrial Country Sample. Dependent Variable: Log GDP Per Capita 2010 (PPP Measured)							
rule	0.838***	0.837***	0.833***	0.849***	0.838***	0.859***	1.405	0.878***

	(0.132)	(0.133)	(0.134)	(0.145)	(0.135)	(0.128)	(0.977)	(0.144)
mal94p	0.484	0.510	0.334	0.376	0.366	0.412	2.732	0.453
	(0.456)	(0.448)	(0.461)	(0.444)	(0.440)	(0.409)	(3.110)	(0.456)
lcopen	-0.127	-0.168						
	(0.165)	(0.182)						
pop100km			-0.135					
			(0.313)					
coluk				-0.0333	-0.0304			
				(0.213)	(0.218)			
colfr					-0.0366			
					(0.343)			
oil						0.371		
						(0.220)		
tyr60							-0.158	
							(0.305)	
prot								-0.00358
								(0.00575)
Constant	9.648***	9.798***	9.278***	9.197***	9.205***	9.107***	9.822***	9.222***
	(0.614)	(0.686)	(0.249)	(0.142)	(0.142)	(0.134)	(1.409)	(0.143)
Observations	21	21	21	21	21	21	19	21
R-squared	0.814	0.813	0.813	0.809	0.810	0.839	0.585	0.809
Panel D: Full Sample. Dependent Variable: AVA-to-GDP ratio in 1990-2000.								
rule	-1.489	-1.068	-2.635	-4.767*	-4.988*	-1.684	3.984	-1.343
	(2.845)	(2.979)	(2.755)	(2.723)	(2.856)	(2.702)	(4.718)	(3.039)
mal94p	26.84***	25.31***	21.11***	19.94***	20.19***	24.21***	23.68***	24.90***
	(5.557)	(5.973)	(5.748)	(5.222)	(6.118)	(5.379)	(7.873)	(5.721)
lcopen	-5.457**	1.516						
	(2.390)	(3.914)						
pop100km			-6.102					
			(3.977)					
coluk				3.632	3.024			
				(2.629)	(3.072)			
colfr					-1.725			
					(3.483)			
oil						-8.203*		
						(4.916)		
tyr60							-2.771	
							(1.970)	
prot								-0.0733
								(0.0927)
Constant	27.81***	2.449	13.08***	9.162***	9.616***	9.490***	21.76*	9.373***
	(8.895)	(14.06)	(3.903)	(2.628)	(2.603)	(2.727)	(11.45)	(2.750)
Observations	74	74	75	75	75	75	53	75
R-squared	0.571	0.527	0.566	0.563	0.564	0.555	0.694	0.535

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 9. Probit-IV of industrialized country dummy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria Ecology (ME), population ratio in temperate zone (kgptemp), settlement mortality rate (logmort) and legal origin (leg_bri)			
	Dependent Variable: Industrialization Dummy						
avexpr	-0.304			0.0978			
	(0.235)			(0.341)			
mal94p	-2.398***			-2.164***	-2.003**		
	(0.389)			(0.638)	(0.996)		
rule		0.465			0.118	0.525	
		(0.569)			(0.678)	(0.492)	
kk			0.483				-0.409
			(1.101)				(0.865)
malfal94		-1.715***	-1.819*			-1.413**	-1.99***
		(0.589)	(1.074)			(0.711)	(0.724)
Constant	2.603	0.024	0.382	-0.319	0.263	0.033	-0.033
	(1.629)	(0.196)	(0.445)	(2.444)	(0.384)	(0.309)	(0.295)
Observations	84	119	57	60	68	68	57
Wald	39.14	19.84	28.25	15.83	14.68	10.1	11.55

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 10. Regressions on degree of industrialization: full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Instrument variables are malaria ecology (ME) and population ratio in temperate zone (kgptemp)			Instrument variables are malaria ecology (ME), population ratio in temperate zone (kgptemp), settlement mortality (logmort) and legal origin (leg_bri)					
	Dependent Variable: AVA-to-GDP Ratio								
avexpr	0.714			0.0593			-2.114		
	(1.877)			(2.009)			(1.488)		
mal94p	29.87***			27.24***	23.77***	27.01***			
	(6.941)			(6.260)	(5.510)	(6.146)			
rule		-6.893***			-2.227			-5.183**	
		(2.595)			(2.832)			(2.328)	
kk			-6.211			1.247			-3.236
			(4.134)			(4.399)			(3.551)
malfal94		17.29***	17.94***				20.29***	18.18***	20.21***
		(4.324)	(4.972)				(4.359)	(4.253)	(4.592)
Constant	0.208	13.05***	12.02***	6.506	9.006***	7.186**	25.98**	12.57***	11.58***
	(15.74)	(1.550)	(2.181)	(15.97)	(2.670)	(2.968)	(11.25)	(1.992)	(2.093)
Observations	106	142	62	67	75	62	67	75	62
R-squared	0.437	0.497	0.478	0.489	0.588	0.459	0.508	0.540	0.555

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 11. OLS regression results on industrialization

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Dependent Variable: AVA-to-GDP Ratio					
avexpr	-2.740***	-3.081***				
	(0.613)	(0.573)				
mal94p	16.93***		18.14***		14.76***	
	(2.542)		(3.288)		(2.293)	
malfal94		16.82***		18.10***		14.87***
		(2.436)		(3.179)		(2.254)
kk			-3.483	-3.640*		
			(2.191)	(2.135)		
rule					-6.980***	-7.436***
					(1.051)	(1.003)
Constant	29.62***	33.08***	10.93***	12.56***	13.14***	13.92***
	(5.072)	(4.615)	(1.946)	(1.684)	(1.187)	(1.088)
Observations	111	111	67	67	154	154
R-squared	0.594	0.602	0.517	0.527	0.561	0.566
	Panel B: Dependent Variable: Rural Population Ratio					
avexpr	-2.100**	-3.099***				
	(0.982)	(0.950)				
mal94p	35.66***		26.67***		27.38***	
	(4.100)		(5.239)		(3.352)	
malfal94		33.27***		25.01***		26.34***
		(4.084)		(5.226)		(3.382)
kk			-10.97***	-11.86***		
			(3.469)	(3.477)		
rule					-8.005***	-9.093***
					(1.524)	(1.489)
Constant	46.62***	56.59***	37.06***	40.09***	37.79***	39.60***
	(8.126)	(7.661)	(3.081)	(2.743)	(1.716)	(1.612)
Observations	111	111	66	66	155	155
R-squared	0.607	0.586	0.579	0.564	0.567	0.554

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.

Table 12. Regressions of institutions on geography-explained industrialization level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Geography-explained AVA-to-GDP ratio as independent variables							
	rule	*rule*	*avexpr*	*avexpr*	*rule*	*rule*	*avexpr*	*avexpr*
agval_mal94p_hat	-0.0634***		-0.0981***		-0.0369**		-0.0748***	
	(0.00963)		(0.0157)		(0.0149)		(0.0253)	
agval_malfal94_hat		-0.0656***		-0.125***		-0.0392*		-0.0947**
		(0.0128)		(0.0236)		(0.0213)		(0.0421)
mal94p					-0.630**		-0.643	
					(0.274)		(0.546)	
malfal94						-0.481		-0.557
						(0.311)		(0.638)
Constant	0.499***	0.230	8.000***	7.606***	0.499***	0.230*	7.987***	7.591***
	(0.149)	(0.139)	(0.274)	(0.252)	(0.145)	(0.137)	(0.274)	(0.253)
Observations	75	75	67	67	75	75	67	67
R-squared	0.372	0.265	0.374	0.302	0.415	0.289	0.387	0.310
	Panel B: Geography-explained rural population ratio as independent variable							
	rule	*rule*	*avexpr*	*avexpr*	*rule*	*rule*	*avexpr*	*avexpr*
agpop_mal94p_hat	-0.0450***		-0.0689***		-0.0262**		-0.0525***	
	(0.00684)		(0.0110)		(0.0105)		(0.0177)	
agpop_malfal94_hat		-0.0512***		-0.0977***		-0.0306*		-0.0739**
		(0.00998)		(0.0184)		(0.0166)		(0.0329)
mal94p					-0.630**		-0.643	
					(0.274)		(0.546)	
malfal94						-0.481		-0.557
						(0.311)		(0.638)
Constant	0.499***	0.230	8.000***	7.606***	0.499***	0.230*	7.987***	7.591***
	(0.149)	(0.139)	(0.274)	(0.252)	(0.145)	(0.137)	(0.274)	(0.253)
Observations	75	75	67	67	75	75	67	67
R-squared	0.372	0.265	0.374	0.302	0.415	0.289	0.387	0.310

Notes: Standard errors in parenthesis. * indicates p-value < 0.1, ** indicates p-value < 0.5 and *** indicates p-value < 0.01.