The Primacy of Institutions Reconsidered: Direct Income Effects of Malaria Prevalence

Kai Carstensen and Erich Gundlach

Some recent empirical studies deny any direct effect of geography on development and conclude that institutions dominate all other potential determinants of development. An alternative view emphasizes that geographic factors such as disease ecology, as proxied by the prevalence of malaria, may have a large negative effect on income, independent of the quality of a country's institutions. For instance, pandemic malaria may create a large economic burden beyond medical costs and forgone earnings by affecting household behavior and such macroeconomic variables as international investment and trade. After controlling for institutional quality, malaria prevalence is found to cause quantitatively important negative effects on income. The robustness of this finding is checked by employing alternative instrumental variables, tests of overidentification restrictions, and tests of the validity of the point estimates and standard errors in the presence of weak instruments. The baseline findings appear to be robust to using alternative specifications, instrumentations, and samples. The reported estimates suggest that good institutions may be necessary but not sufficient for generating a persistent process of successful economic development.

Economists, historians, and other social scientists have explained the large differences in the standard of living between the world's richest and poorest nations in many different ways. One strand of the literature has emphasized the preeminent role of physical geography in explaining cross-country differences in the level of development. Some recent empirical studies deny any direct effect of geography on development and conclude that institutions dominate all other potential determinants of development (Hall and Jones 1999; Acemoglu, Johnson, and Robinson 2001; Easterly and Levine 2003; Rodrik, Subramanian, and Trebbi

Kai Carstensen and Erich Gundlach are research fellows at the Kiel Institute for the World Economy; their email addresses are kai.carstensen@ifw-kiel.de and erich.gundlach@ifw-kiel.de. The authors thank two anonymous referees and the editor, seminar participants at Aarhus University, Goethe University in Frankfurt, and Hamburg University, participants at the Ninth Convention of the East Asian Economic Association in Hong Kong, and Michael Funke, Charles I. Jones, Rolf J. Langhammer, Dani Rodrik, and Jeffrey Sachs for helpful comments on an earlier draft, as well as William A. Masters for providing the data on frost frequency, Gordon McCord for explaining the malaria data, Stephen Donald for sharing a sample program on instrument selection, and Jean-Louis Arcand for suggesting econometric strategies.

THE WORLD BANK ECONOMIC REVIEW, VOL. 20, NO. 3, pp. 309–339 doi:10.1093/wber/lhl001
Advance Access publication June 8, 2006

2004). Engerman and Sokoloff (1997) and Acemoglu, Johnson, and Robinson (2002) examine how geographic endowments in the Americas may have shaped factor abundance (people per unit of land) and how unequal factor abundance may have shaped persistent institutions imposed by the colonizing powers that enabled the entrenchment of a small group of elites. Because the Industrial Revolution required the broad participation of the population in entrepreneurship and innovation, economies were favored that started with a more equal factor abundance (due to geographic endowments) and hence with institutions that resulted in a less unequal distribution of income and wealth. Overall, this new literature emphasizes that geographic endowments affect the level of development only through their impact on factor abundance, political economy, and institutions and not more directly.

While development economists can easily agree on the relevance of good institutions for successful development and on the indirect role of physical endowments in shaping different institutional outcomes and different paths of development, there is no agreement on the direct role of geography for development in the recent empirical literature. Partly by highlighting the arguments of the older literature and partly by presenting new empirical evidence, Jeffrey Sachs and his coauthors in particular have argued in a series of papers that measures of geography such as disease ecology may directly affect the level of economic development in addition to the undisputed effects of the institutional framework of a country (Bloom and Sachs 1998; Gallup, Sachs, and Mellinser 1999; Gallup and Sachs 2001; Sachs 2001; McArthur and Sachs 2001; Sachs and Malaney 2002; Sachs 2003). The main disagreement in the current debate concerns the robustness of the empirical evidence presented by Sachs and his coauthors, which has been directly rejected by Acemoglu, Johnson, and Robinson (2001) and Rodrik, Subramanian, and Trebbi (2004) and is in conflict with the studies that favor the primacy of institutions.

A parsimonious baseline specification is used here to reconsider the general econometric limitations of alternative empirical strategies that have been applied to derive clear-cut conclusions with regard to the deep determinants of development. Recent empirical studies have not treated geographic variables such as disease ecology in the same way as measures of institutions, thus probably reducing the chances of the geography hypothesis to prevail. The contribution of this article is to see whether a measure of disease ecology such as malaria prevalence, which is likely to be an endogenous geography variable, directly explains the level of development independent of a measure of institutions, which is also likely to be an endogenous variable. This appears to be an unsettled question in the empirical literature.

Mainly to keep the empirical analysis tractable in the presence of a limited number of candidates for instrumental variables, explanatory variables other than institutions and disease ecology are ignored, not least

because measures such as the quality of economic policies (Easterly and Levine 2003) or the level of trade integration (Rodrik, Subramanian, and Trebbi 2004) have not been found to exert a direct effect on the level of development independent from the effect of institutions. Instead, the emphasis is on the fundamental problems of statistical inference implied by instrumental variable estimation and on tracing the basic reason for the different empirical results on the direct role of disease ecology. After controlling for institutional quality, malaria prevalence is found to have quantitatively important direct negative effects on income. This finding appears to be robust to alternative specifications, instrumentations, and samples.

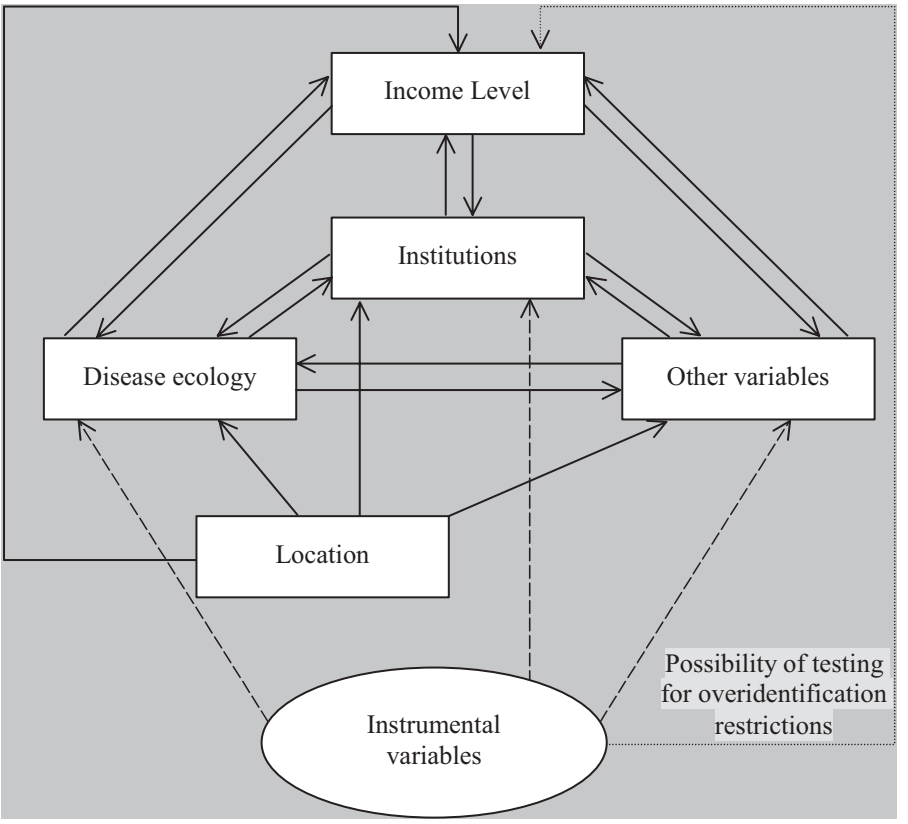
Finding a robust direct effect of malaria prevalence, which is held to be a proxy for the adverse disease ecology of a country, should matter for devising appropriate development policies. For instance, if there is no empirical evidence that malaria prevalence directly affects the level of development in impoverished countries, foreign aid may be targeted mainly to improve policies and institutions. But the finding that there are such direct effects on income means that foreign aid should also be spent on solving biophysical and technological problems that are specific to public health in tropical countries. Especially in Sub-Saharan Africa, but probably also in parts of Asia and Latin America, poor countries may need something in addition to good institutions to generate a persistent process of successful economic development.

I. The Many Possible Links between Geography, Institutions, and Development

The various approaches that have been used to identify the many possible links between geographic endowments, institutions, and level of development can be represented with the help of figure 1. The solid arrows indicate the potential directions of causality between the variables considered, the dashed arrows indicate the relation between the instrumental variables and the endogenous explanatory variables, and the dotted arrow indicates the possibility of testing for overidentification restrictions, if there are more instrumental variables than endogenous variables.

Instrumental variables are needed to identify the direct development effects of institutions and other endogenous explanatory variables. An instrumental variable can be used to identify that part of the variation in the endogenous explanatory variables that is exogenous to the variation in the dependent variable, which here is the level of income per capita. The solid arrows indicate that the instrumental variable method can identify the true causal effect of the endogenous variables on the level of income. If valid instrumental variables are available, unbiased estimates of the causal effects on income of institutions, disease ecology, and other variables may be obtained without having to identify

FIGURE 1. Alternative Links between Geography, Institutions, and Development



Source: Authors' summary.

the potential reverse causality from the level of income to the explanatory variables. In addition, the instrumental variable method can be used in principle to identify the web of relations between the endogenous explanatory variables and the indirect effects of geography on the level of income.

Indirect effects of geography are especially highlighted by Engerman and Sokoloff (1997). In terms of figure 1, they argue that geographic endowments may determine factor abundance (the arrow from "location" to "other variables"), factor abundance may entrench persistent institutions, and institutions may influence the level of development. In line with their hypothesis, Rodrik, Subramanian, and Trebbi (2004) consider whether a measure of location may affect the level of development directly or indirectly through measures of institutions and trade integration. They find only weak statistical evidence for a direct effect of a measure of geography, whereas a measure of the quality of institutions appears as the only important explanatory variable. Their conclusion is

confirmed by Easterly and Levine (2003), who do not use a measure of geography in their specification but instead employ an overidentification test to see whether their set of geography-based instrumental variables can be considered valid. They find no evidence that any of their instrumental variables should be included in their specification, which may also be interpreted as supporting the hypothesis of the primacy of institutions.[1]

With all possibilities taken into account, certain variables may affect the level of income directly, through their effects on other variables, through one channel, or not at all. Empirical estimates of the various potential causal effects depend crucially on the quality of the available instrumental variables, which have to be correlated with the endogenous explanatory variables but uncorrelated with the error term in the structural equation. This condition will be satisfied if the instrumental variables affect the dependent variable only indirectly through the endogenous explanatory variables that are included in the structural equation and if the instrumental variables are not affected by any feedback from the endogenous variables. Thus, almost by definition, valid instrumental variables are difficult to come by in the cross-country empirics of development, because most economic variables are affected by the level of income. Measures of geography play a special role in this context because most of them, such as distance from the equator or temperature, can undoubtedly be considered as exogenous to the level of development and to the endogenous explanatory variables.

The problem is that once a measure of geography has a direct effect on the level of income, it no longer qualifies as a valid instrumental variable. For instance, local temperature may affect the level of income either directly through the location–income link or through its effects on other variables (figure 1). Where there is a direct location–income link, the measure of geography cannot be used as an instrumental variable. And if the remaining available instrumental variables are also mainly measures of geography such as distance from the equator, the independent variation across the exogenous variables could probably turn out to be too small to allow for an empirical identification of all causal effects of interest.

Acemoglu, Johnson, and Robinson (2001) estimate a specification where institutions and malaria prevalence determine the level of income. Since the measure of malaria prevalence is not instrumented but is nevertheless statistically insignificant, they conclude that there is no causal disease–income link. In support of their empirical results, Acemoglu, Johnson, and Robinson dismiss a priori the possibility that tropical diseases such as malaria could have a large effect on the level of development because people living in areas where such diseases are endemic may have developed immunities against the diseases. According to this view, malaria is unlikely to have strong income effects because it is a debilitating rather than a fatal disease, with the risk of severe illness and

1. See Hall and Jones (1999) for the same line of reasoning based on overidentification tests.

death limited mainly to people without any immunity such as children below the age of 5 years and adults who grew up elsewhere, like European settlers.[2]

However, this argument ignores other mechanisms through which malaria may affect the level of income. For instance, one form of immunity against malaria comes at a cost for the adult population. The sickle cell trait provides protection against malaria without serious health complications when inherited from only one parent, but the same allele inherited from both parents leads to sickle cell anemia.[3] Sickle cell anemia generates severe episodes of pain and increasing infections, outcomes that are at least comparable to the direct negative health effects of malaria experienced by people without any immunity. These considerations suggest that due to natural selection, areas with a high prevalence of malaria are likely to be areas with a high prevalence of sickle cell anemia. Some estimates claim that up to 40 percent of the population in tropical Africa may carry the sickle cell trait.[4] Thus, malaria may additionally cause poor health and absenteeism of the workforce through natural selection in favor of a high prevalence of the sickle cell trait.

This is not to deny that traditional studies of the economic costs of malaria find relatively small gross domestic product (GDP) losses on the basis of the total number of cases and the fixed costs of prevention and treatment. But there are at least two other mechanisms through which the pandemic nature of malaria can impose large economic costs, namely by affecting household behavior and macroeconomic variables such as foreign direct investment, trade, and tourism (Sachs and Malaney 2002). In response to the disease, households may increase fertility and thus the dependency ratio, which will reduce GDP per capita. A high fertility rate is also likely to reduce investments in human capital per child, which may reduce possibilities for long-run development. In addition, malaria infections can reduce the cognitive development and learning ability of children, which may further depress the long-run average skill level and thus the level of development. Malaria may also decrease household investment in physical

2. The Anopheles mosquito is the vector that transmits malaria from human to human. The mosquito must first bite an infected person who is sick with malaria. Then, the mosquito must survive several days while the malaria parasite develops in its body. Finally, the infected mosquito must bite another human to complete the circle of infection. About 40 species of the Anopheles mosquito are significantly involved in the transmission of malaria. These species differ substantially with respect to their feeding behavior on humans and their longevity. Thus, all other things constant, the potential for malaria transmission will be high in densely populated regions where the locally dominant Anopheles has developed, through biological evolution, a specific human-biting behavior and a relatively high daily survival rate and has found excellent breeding conditions. Most tropical regions combine these favorable conditions for malaria prevalence. In moderate climatic zones far away from the equator, the breeding conditions are less favorable, and the locally dominant Anopheles mosquitoes are less specialized on human biting and usually less robust. However, differences in the potential for malaria transmission do exist not only across but also within climatic zones, as is shown by a new measure called the stability of malaria transmission (Kiszewski and others 2004).

3. For information on sickle cell anemia, see http://www.scinfo.org/ (March 2006).

4. See http://www.pbs.org/wgbh/evolution/library/01/2/l_012_02.html (March 2006).

capital compared with a situation with less-frequent episodes of illness, lower fertility, and lower dependency ratios.

At the macroeconomic level, malaria appears to suppress the economic linkages between malarious and nonmalarious regions of the world. Foreign investors may avoid malarious regions if the disease burden raises the costs of attracting the needed labor force, and with less foreign investment inflows, there will be fewer possibilities of exploiting comparative advantages by specialization and international trade. For instance, trade in services such as tourism cannot prosper under conditions of high rates of malaria transmission, which will reduce the possibility to import growth-enhancing investment goods. In addition, malaria might be closely related to other diseases, either as a direct causal factor or by rendering individuals more susceptible to other diseases, which would further increase its cumulative economic costs. Malaria prevalence may thus have much larger economic costs than will be visible from calculating only the direct medical costs and forgone earnings, especially if international trade in goods and services and international investment are critical factors for development in an era of closer global economic integration.

To put the geography hypothesis on a more equal footing with the primacy of institutions hypothesis, both the prevalence of malaria and the quality of institutions are used as endogenous explanatory variables, as suggested by Sachs (2003). The main interest is the relative size of the causal effects of the institutions–income link and the disease–income link (figure 1). This approach differs from the approaches by Hall and Jones (1999), Acemoglu, Johnson, and Robinson (2001), Easterly and Levine (2003), and Rodrik, Subramanian, and Trebbi (2004), which do not always instrument their measure of endogenous disease ecology, use an exogenous measure of geography, or rely on an overidentification test only. Because of previous results reported in the literature and a general shortage of plausible instrumental variables, all possible effects that might result from explanatory variables other than institutions and disease ecology are ignored, as are possible links between institutions and disease ecology. Different from Sachs (2003), additional instrumental variables are employed, recently developed econometric tests are used to check the validity of the point estimates and standard errors in the presence of weak instrumental variables, and a test for overidentification restrictions is applied to avoid a potentially unjustified exclusion of exogenous measures of geography from the baseline specification.

II. Specification and Data

In line with previous empirical studies, the following cross-country regression equation is used to estimate the relative effects of institutional quality (*INSTITUTIONS*) and malaria prevalence (*MALARIA*) on economic development, which here is measured by the logarithm of GDP per capita (ln*GDPC*):

$$(1) \qquad \ln GDPC_i = \beta_1 + \beta_2 \cdot INSTITUTIONS_i + \beta_3 \cdot MALARIA_i + \varepsilon_i$$

where ε_i is an error term with zero mean and common variance, and β_2 and β_3 are the coefficients of interest. The research question is whether an estimate of β_3, as represented by the disease–income link in figure 1, is statistically different from zero, negative, and quantitatively important. To better understand where the different results in the literature may come from, the baseline specification [(equation (1)] is re-estimated by paying particular attention to the choice of the variables, the instruments, and the country sample.

The Choice of the Variables

Indicator variables are needed to measure the effects of institutions and geography on the level of economic development. Such indicator variables are necessarily incomplete and erroneous because the three concepts are multidimensional and difficult to measure. Therefore, different indicator variables are used. The dependent variable is either the ln*gdpc* in 1995, which is used by Acemoglu, Johnson, and Robinson (2001), Easterly and Levine (2003), and Rodrik, Subramanian, and Trebbi (2004), or the log of GDP per working age person in 1990 (ln*gdpw*), which appears to be more closely related to the applied growth literature and is used by Hall and Jones (1999).[5]

Institutional quality is measured by one of the following three variables: an average index of the quality of governance in 1996 (*rule*) from Kaufmann, Kraay, and Mastruzzi (2004), the index of government antidiversion policies in 1986–95 (*gadp*) used by Hall and Jones (1999), and the index of protection against expropriation in 1985–95 (*exprop*) used by Acemoglu, Johnson, and Robinson (2001). To measure disease ecology, two measures of malaria prevalence are employed: the proportion of a country's population at risk of malaria falciparum transmission in 1994 (*malfal*) used by Acemoglu, Johnson, and Robinson (2001) or a new index of malaria risk (*malrisk*) suggested by Sachs (2003). The new index is based on the prevalence of nonfatal species of the malaria pathogen (*Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale*), where a relatively higher proportion of malaria vivax is reported for the Americas, Europe, and much of Asia than for sub-Saharan Africa. For international comparisons, the new index may provide a more accurate measure of the share of the population that is at risk of malaria infection than the measure used by Acemoglu, Johnson, and Robinson (2001).

The Choice of the Instrumental Variables

Two premises, both suggested by Acemoglu, Johnson, and Robinson (2001), are used to find an instrumental variable for institutional quality. First, studying the impact of institutions on the level of development has to focus on a sample of

5. See the appendix for detailed descriptions of the data and sources.

former colonies, because only this sample provides the necessary exogenous variation in measures of institutions that can be exploited to estimate a causal effect. Second, the potential endogeneity of any measure of institutional quality should be controlled for by a measure that is correlated with the current variation in the institutional frameworks without being influenced by current economic conditions,[6] and it should only affect the current level of development through its effect on institutions but not directly. In this context, mortality among European settlers in the early nineteenth century appears to be the most plausible instrumental variable that has been suggested to date.

Differences in mortality among early settlers across colonies, which were well known in Europe at the time, may explain the differences in institutional frameworks that were created by the colonizing powers. For instance, regions with low mortality were favored for settlement, and colonies of settlers may have implemented for themselves a set of institutions that resembled the institutions of their home countries by establishing property rights, the rule of law, and checks against government power. In regions where large-scale settlement was not feasible for Europeans because of an unfavorable disease ecology and high rates of mortality, the colonial powers may have imposed a different set of institutions that did not protect private property and did not provide protection against expropriation but instead focused mainly on the extraction of natural resources.[7] Since early settler mortality is certainly independent of current economic conditions and since early institutional frameworks have proved to be fairly persistent over time (Acemoglu, Johnson, and Robinson 2001), settler mortality across former colonies can be used as an instrumental variable that helps to identify the exogenous cross-country variation in current institutional frameworks.

To control for the endogeneity of malaria prevalence, a new measure of malaria ecology (*maleco*) is considered that was developed by Kiszewski and others (2004) and first used for cross-country regressions by Sachs (2003). Since this measure of malaria ecology is built only on the climatic factors and biological properties of each regionally dominant malaria vector, Kiszewski and others (2004) claim that *maleco* is exogenous to public health interventions and economic conditions and thus can be considered as a valid instrumental variable in regressions of economic development on malaria risk.

The index of malaria ecology measures the contribution of regionally dominant vector mosquitos to the *potential* transmission intensity of malaria. Thus, it includes regions where malaria is not currently transmitted but where it had

6. Technically, this means that the control measure should be uncorrelated with the error term of the income equation [equation (1)].

7. The hypothesis advanced by Acemoglu, Johnson, and Robinson (2001) that geographic and climatic conditions were decisive for the adoption of institutions that favored either settlement or resource extraction is in conflict with some historical facts for the colonization of the Americas. An alternative hypothesis favored by Engerman and Sokoloff (1997) emphasizes initial factor abundance as determinants of institutions. See Hoff (2003) for a survey of the issues.

been transmitted in the past or might be in the future.[8] Since the region-specific dominant malaria vector reflects only the forces of biological evolution, it can be considered independent of current economic conditions. That is, terms likely to be affected by economic conditions or public health interventions (*mosquito abundance*, for example) do not enter the calculation of the index. The index reveals that, because of different vector properties, a given malaria intervention is likely to have a smaller impact in the tropics than in more temperate climatic zones, where the vector is less robust and does not specialize in human biting and where the parasite has less fatal infectious consequences.

However, Rodrik, Subramanian, and Trebbi (2004) doubt that *maleco* is actually exogenous to current economic conditions. They object that Sachs (2003) does not detail the construction of the index and point out that Kiszewski and others (2004) do not discuss exogeneity at all.[9] While this critique is technically correct, doubts regarding the exogeneity of *maleco* may not be justified, as discussed in the previous paragraph. Nevertheless, three sets of further instrumental variables that relate to the climatic environment, the influence of Western European languages, and the openness of a country are considered in addition to the two baseline instruments ln*mort* and *maleco*.

Temperature, rainfall, and latitude are additional measures of the climatic environment that can be related to preconditions for the prevalence of malaria. Since a key part of the life cycle of the parasite depends on a high ambient temperature, malaria is intrinsically a disease of warm environments. Malaria also depends on adequate conditions of mosquito breeding, mainly pools of clean water from rainfall. Hence, the prevalence of frost (*frost*), measured as the proportion of a country's land receiving 5 or more frost days in winter, or the degree of humidity (*humid*), measured as the highest temperature during the month when average afternoon humidity is at its highest, may be considered as appropriate instrumental variables that are exogenous to economic conditions. In addition, distance from the equator as measured by the absolute latitude of a

8. Abstracting from all detail, the construction of the index proceeds in two basic steps. First, the regionally dominant Anopheles mosquito is identified across countries in which malaria is or has been endemic. The criteria for the identification of the dominant vector are its longevity and its human-biting habit. Second, the index of malaria ecology is calculated as $\left(a_i^2 p_i^E\right)/(-\ln p_i)$, where i is the identity of the dominant malaria vector, a is the proportion of vector i biting people [0,1], p is the daily survival rate of vector i [0,1], and E is the length of the extrinsic incubation period in days, which depends mainly on average temperature and differs between *Plasmodium falciparum* and *Plasmodium vivax*. Hence, the index value for a specific country is measured as a function of climatic factors that determine the required habitat of the dominant vector and of biological properties of the region-specific dominant vectors.

9. Information on the construction of the malaria transmission index (malaria ecology) is available online at http://www.earth.columbia.edu/about/director/malaria/index.html (March 2006). A previous version of the text describing the construction of the index may have contributed to the impression that *maleco* is not purged of endogeneity, because it stated that a measure of mosquito abundance is included in the calculation. However, observed mosquito abundance enters the index of malaria ecology only as a screen for precipitation data, where the independently identified dominant malaria vector is assumed to be absent from the specific site under consideration if precipitation falls below a certain level per month.

country (*latitude*) may also be used as a proxy for the climatic environment. What has to be taken into account, however, is that these three measures of climatic conditions may be instrumental variables not only for measures of disease ecology but also for measures of institutional quality. This is because settler mortality and thus the design of early institutions were influenced by the prevailing climatic conditions of the colonies. In this context, Acemoglu, Johnson, and Robinson (2001) point out that their work shows why absolute distance from the equator might matter as an instrument for a measure of institutions, as used in Hall and Jones (1999).

Other plausible instrumental variables than measures of geography are more difficult to come by. One possibility is to consider the share of the population that speaks English (*engfrac*) or another Western European language (*eurfrac*) as the first language. As suggested by Hall and Jones (1999), these variables may reflect the different degree of Western European influence on the sample countries and thus may help to identify the exogenous variation in measures of institutions. Since Acemoglu, Johnson, and Robinson (2001) generally question the exogeneity of these variables, some formal tests of exogeneity are provided when using them in the checks of the robustness of the baseline results.

Furthermore, instrumental variables that are used for checking robustness relate to a country's trade openness. More open countries may have better institutions because openness may encourage less arbitrary government behavior, especially toward property rights. Thus, exogenous measures of openness could be used as instrumental variables for measures of institutions. Two measures of openness are employed: the proportion of land area that is within 100 kilometers of the coast (*coast*), which is taken from McArthur and Sachs (2001), and the (log) predicted trade share of a country (*trade*), which is constructed by Frankel and Romer (1999) from a gravity model that uses mainly geographical variables to explain actual bilateral trade flows.

Choice of the Sample

The sample of countries is limited to former colonies for which data on early settler mortality are available. Acemoglu, Johnson, and Robinson (2001, table 7, p. 1392) estimate equation 1 for a sample of 62 countries. This sample, however, includes 14 countries that are known to provide unreliable statistics (rated as D countries in Summers and Heston 1991), two countries that are very small (less than 1 million inhabitants in 1990), and one country that depends mainly on oil production. These countries are removed from the sample. Thus, baseline results are reported for a smaller but probably more reliable sample of 45 former colonies that are not statistical terra incognita, small, or dependent on oil production. By contrast, previous studies that took issue with the Acemoglu, Johnson, and Robinson result on the primacy of institutions (McArthur and Sachs 2001; Easterly and Levine 2003; Sachs 2003; Rodrik, Subramanian, and Trebbi 2004) increased the Acemoglu, Johnson, and Robinson sample size but

disregarded data quality. As a robustness check of the baseline findings on sample size, a larger sample of countries with additional observations on settler mortality (Acemoglu, Johnson, and Robinson 2000) is also included.

III. Baseline Estimation Results

To begin, equation 1 is estimated by two-step least squares (2SLS) using ln*gdpc* as the dependent variable, *rule* and *malfal* as explanatory variables, and ln*mort* and *maleco* as instrumental variables. This baseline specification is close to specifications in the literature. In particular, (a) ln*gdpc* is used as the dependent variable by Acemoglu, Johnson, and Robinson (2001), Easterly and Levine (2003), and Rodrik, Subramanian, and Trebbi (2004); (b) *rule* is used as a measure of institutional quality by Easterly and Levine (2003) and Rodrik, Subramanian, and Trebbi (2004); and (c) *malfal* is used as a measure of malaria prevalence by Acemoglu, Johnson, and Robinson (2001). The results are presented in column 1 of table 1.

The point estimates have the expected signs and are quantitatively important. The point estimate of β_2 reflects the change in log output per capita associated with an one-unit increase in the index of governance quality. Thus, $\beta_2 = 0.89$ implies that a difference of 0.1 in the governance index is associated with a 8.9 percent cross-country difference in output per capita. To show the potential magnitude of the estimated effect of the measure of institutions on economic performance, two countries in the sample that represent about the 70th and the 30th percentile of the governance index are compared, South Africa with an index value of 0.21 and Ecuador with a value of −0.40. This difference is predicted to result in a 0.54 log-point difference [(0.21 + 0.40) times 0.89] between the log per capita GDPs of the two countries. That is, the per capita GDPs of South Africa and Ecuador are predicted to differ by a factor of about 1.7 due to institutional differences, whereas their sample per capita GDPs differ by a factor of about 2.7.

The point estimate of β_3 reflects the change in log output per capita associated with an one-unit increase in malaria prevalence. Thus, $\beta_3 = -1.04$ implies that the per capita GDPs of Paraguay and Pakistan, which represent roughly the 40th and the 70th percentile of the highly stratified distribution of the malaria index (with percent values of 0.001 for Paraguay and 0.49 for Pakistan), should differ by a factor of about 1.7 due to the differences in the proportion of the population that lives with the risk of malaria infection, whereas the sample per capita GDPs of these two countries differ by a factor of about 2.6.

The point estimates are statistically significant, with estimated standard errors of 0.18 for β_2 and 0.30 for β_2, which imply *t*-statistics of 5.04 and −3.46. These values indicate statistical significance at the 5 percent level when using *t*-tests based on conventional asymptotic theory. The reported conventional confidence intervals contain the unknown true parameters with a

TABLE 1. Baseline Estimation Results

	1		2		3		4	
Explanatory variables	*rule*	*malfal*	*rule*	*malrisk*	*gadp*	*malfal*	*exprop*	*malfal*
Estimated coefficients								
Two-step least square	0.89	−1.04	0.78	−1.31	3.31	−1.48	0.55	−1.03
Standard error	0.18	0.30	0.22	0.42	0.60	0.22	0.12	0.34
Fuller[a]	0.89	−1.05	0.78	−1.29	3.31	−1.48	0.53	−1.07
Standard error	0.17	0.30	0.22	0.40	0.59	0.21	0.12	0.33
Bounds of 95 percent confidence intervals								
Conventional								
Upper	0.53	−1.65	0.33	−2.16	2.09	−1.91	0.30	−1.72
Lower	1.25	−0.43	1.23	−0.47	4.52	−1.04	0.79	−0.34
Conditional likelihood ratio								
Upper	0.44	−1.83	0.05	−2.85	1.53	−2.07	0.29	−1.75
Lower	1.43	−0.18	1.39	−0.23	4.99	−0.87	1.29	0.75
Number of observations	45		45		45		45	
Number of instruments	2		2		2		2	
First-stage statistics								
F-statistic	28.12	42.31	28.12	20.38	23.37	41.78	17.49	42.31
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Partial *R*	0.57	0.67	0.57	0.49	0.53	0.67	0.45	0.67
Shea partial *R*	0.37	0.43	0.33	0.28	0.37	0.46	0.23	0.34
Weak-instrument test[b]								
Cragg–Donald	11.45		7.91		10.33		5.74	
Critical value	7.03		7.03		7.03		7.03	

Note: All specifications are estimated with the instrumental variables ln*mort* and *maleco* for the small sample of 45 countries.

[a]Denotes the Fuller estimator with correction parameter $c = 1$ proposed by Hausman, Stock, and Yogo (2005).

[b]The Cragg–Donald statistic (Cragg and Donald 1993) is used by Stock and Yogo (2002) for weak-instrument tests. If the Cragg–Donald statistic exceeds the critical value, then a standard significance test with nominal size of 5 percent has a maximal size of 10 percent.

Source: Authors' analysis based on data described in the text.

confidence level of 95 percent. Thus, the true coefficient of the governance index is estimated to be between 0.53 and 1.25, and the true impact of malaria prevalence is estimated to be between −1.65 and −0.43. However, these results depend on instrument relevance, as emphasized by the recent literature on weak instruments (Staiger and Stock 1997; Moreira 2003). If the instrumental variables are only weakly correlated with the endogenous explanatory variables, conventional asymptotic theory no longer holds, and statements about statistical significance and inference may lead to the wrong conclusions. The relevance of the instrumental variables has to be checked to allow for statements about statistical significance.

The first-stage regressions supply valuable information about the relevance of the instrumental variables. Highly significant *F*-statistics of 28.1 and 42.3 for the

first-stage regressions of *rule* and of *malrisk* are reported (see table 1). In addition, both the usual partial R^2 and the Shea (1997) partial R^2 are far above zero. These test results point to strong instruments, but even large first-stage *F*-statistics can be misleading. For example, the two instruments ln*mort* and *maleco* may not carry sufficient independent information, which could make it difficult to identify distinct effects of *rule* and *malrisk*. To this end, a statistic proposed by Cragg and Donald (1993) is computed, which represents the relevance of the weakest instrument. Using weak-instrument asymptotic theory, Stock and Yogo (2002) show that a conventional significance test on β with a nominal size of 5 percent has an actual size of 10 percent or more, and is thus severely distorted, if the Cragg–Donald statistic is below 7. Since the Cragg–Donald statistic equals 11.45, the results of the baseline specification are not affected by weak-instrument problems.

As a further robustness check for specifications with potentially weak instruments, a modified limited information maximum likelihood estimator (Fuller 1977) is applied. This Fuller estimator with modification parameter $\alpha = 1$ is more robust to the presence of weak instruments than 2SLS (Hahn, Hausman, and Kuersteiner 2004; Hausman, Stock, and Yogo 2005). The Fuller point estimates are almost identical to the 2SLS estimates, with estimates of 0.89 for β_2 and -1.05 for β_3. In addition, 95 percent confidence intervals are computed based on inverted conditional likelihood ratio (CLR) tests that take any weak-instrument problem into account (Moreira 2003).[10] The CLR intervals turn out to be only slightly larger than the confidence intervals based on conventional asymptotic theory reported above, ranging from 0.44 to 1.43 for β_2 and from -1.83 to -0.18 for β_3. In particular, when the CLR confidence intervals are turned into significance tests, both β_2 and β_3 are individually statistically significant, because the confidence intervals do not include zero.

Before proceeding with further robustness checks, it is worth summarizing the results obtained with the baseline specification. The point estimates for the effects of institutional quality and malaria prevalence have the expected signs and are economically important. They do not appear to suffer from a weak-instrument problem and so appear statistically significant as well. Identifying a direct effect of a measure of disease ecology on the level of development conflicts with the evidence presented by Hall and Jones (1999); Acemoglu, Johnson, and Robinson (2001); Easterly and Levine (2003); and Rodrik, Subramanian, and Trebbi (2004) and confirms the evidence presented by Sachs (2003).

10. Since there are two endogenous explanatory variables, the approach of Moreira (2003) delivers a bivariate confidence region from which two univariate confidence intervals are calculated by the projection method put forward by Dufour (1997). A Matlab program that accomplishes this task is available on request.

IV. Robustness

The results of the baseline specification are subjected to a number of robustness checks, beginning with the effects on the baseline results of alternative measures of the dependent variable and of the endogenous explanatory variables. Other robustness checks assess the inclusion of alternative and additional instrumental variables, the validity of the baseline instruments ln*mort* and *maleco*, and the impact on the results when a larger sample of countries is used.

Effects of Alternative Variables

Estimation results for specifications with alternative explanatory variables are presented in columns 2–4 of table 1. In column 2, institutional quality is still measured by the governance index (*rule*), but malaria prevalence is now measured by the risk of infection with the nonfatal malaria pathogen (*malrisk*). The main difference from the baseline specification is the smaller weight for institutional quality and the larger weight for malaria prevalence. This difference may be due simply to estimation uncertainty, which has increased compared with the baseline specification, as indicated by the larger confidence intervals. Moreover, the weak-instrument problem is of slightly greater relevance than before, as indicated by a smaller Cragg–Donald statistic, which nevertheless still exceeds the critical value of 7. Despite somewhat weaker test statistics, all general conclusions drawn from the baseline specification are confirmed by the specification with *malrisk* as well. Therefore, all further-reported specifications use *malfal* as the measure of malaria prevalence.[11]

In column 3, the governance index is replaced by the index of government antidiversion policies (*gadp*) as a measure of institutional quality, while malaria prevalence is measured by the risk of infection with malaria falciparum (*malfal*). The point estimate for β_2 is considerably larger than in the baseline specification, but this is due mainly to the smaller variance of *gadp* compared with *rule*. The point estimate for β_3 is also absolutely larger than in the baseline specification, but the difference is not substantial if estimation uncertainty is taken into account. The economic significance of these estimates can be shown again for the country pairs discussed above. With a point estimate of 3.31 for β_2, the empirical model predicts that the per capita GDPs of South Africa and Ecuador differ by a factor of 1.7 due to differences in institutional quality, whereas their sample per capita GDPs differ by a factor of about 2.7. With a point estimate of -1.48 for β_3, the model predicts that the per capita GDPs of Paraguay and Pakistan differ by a factor of 2.1 due to differences in malaria prevalence, whereas their sample per capita GDPs differ by a factor of about 2.6. The statistical significance of these estimates can be inferred from both the conventional and the CLR confidence intervals, which do not include zero. In addition,

11. Detailed results based on specifications with *malrisk* are available on request.

there is no weak-instrument problem, as indicated by a large Cragg–Donald statistic.

In column 4, the governance index is replaced by the risk of expropriation (*exprop*), while the risk of infection with malaria falciparum (*malfal*) remains the measure of malaria prevalence. This is the specification analyzed by Acemoglu, Johnson, and Robinson (2001, table 7), who obtain an insignificant effect of malaria prevalence with their sample of countries (without instrumenting *malfal*). For the re-estimated equation, the point estimate of β_2 is smaller than in the baseline specification, but this may be explained by the estimation uncertainty and the larger variance of *exprop* compared with *rule*. The point estimate of β_3 is virtually unchanged. At first sight, both estimates appear statistically significant, as indicated by low standard errors (2SLS and Fuller). However, the Cragg–Donald statistic of 5.74 indicates a weak-instrument problem. The 2SLS estimator may be biased and the conventional confidence intervals may be inadequate. While the point estimates and standard errors remain virtually unchanged when the robust Fuller estimator is used, the CLR confidence intervals of Moreira (2003) indicate that the estimate of the coefficient on *exprop* is statistically significant but the estimate of the coefficient on *malfal* is not.

However, this result does not necessarily imply that malaria prevalence does not have an effect on per capita income. Rather, it indicates that it may not be possible to identify an independent effect with sufficient precision, given the relatively small sample size. This view can be supported by two observations. First, a correlation coefficient of 0.6 shows that the instrumental variables ln*mort* and *maleco* are strongly correlated, which does not leave much information in one instrument that is independent of the other. While this information appears to be sufficient for the previous specifications, it turns out to be insufficient for the current specification, as indicated by lower Shea partial R^2s than before.

Second, the power of the significance test for β_3 (which is derived from its estimated confidence interval) might be low, probably due to the weak-instrument problem. While not much is known about the power of significance tests in the presence of weak instruments, power is certainly lower than in the conventional strong-instrument case, due to reduced estimation precision. More specifically, the power of a significance test for β_3 using conventional asymptotic theory should give an upper bound for the power of a significance test under weak-instrument asymptotic theory. Fortunately, power for the former test can be easily calculated following the approach by Andrews (1989). An interval with power below 0.5 is calculated to see in which region of true parameter values $\beta_3 \neq 0$ the test can be expected not to reject the wrong null hypothesis $\beta_3 = 0$. The interval turns out to be [−0.67, 0.67]. This implies that true parameter values of β_3 between −0.67 and 0.67 have a better chance to be undetected than to be detected. An interval with power below 0.95 turns out to be [−1.23, 1.23]. This implies that only true parameter values of $|\beta_3| > 1.23$ are likely (with

probability above 95 percent) to be found statistically significant. The 2SLS point estimate of β_3 is only -1.03.

Since the correct power of a significance test based on weak-instrument asymptotic theory is likely to be overstated in this exercise, the lower interval limits of -0.67 for power regions below 50 percent and of -1.23 for power regions below 95 percent are only upper bounds for the correct but unknown interval limits. That is, the power of the significance test for β_3 appears to be quite low, given that parameter values for β_3, such as the reported point estimate, are economically important but statistically difficult to distinguish from zero. Thus, finding the coefficient estimate of β_3 to be statistically insignificant is probably due to the low power of the significance test rather than to the unimportance of malaria prevalence for economic development.[12] This conclusion is also corroborated by the statistically significant point estimates of β_3 in columns 1 and 3 of table 1.

Effects of Additional Instrumental Variables

To exclude the possibility that the reported results are driven by the choice of instrumental variables, the analysis is replicated using additional sets of instrumental variables related to the climatic environment (*frost*, *humid*, and *latitude*), the Western European influence (*eurfrac* and *engfrac*), and openness (*coast* and *trade*). The analysis is restricted to the baseline specification, where institutional quality is measured by the governance index (*rule*) and malaria prevalence is measured by the risk of infection with malaria falciparum (*malfal*). The results are presented in table 2.

Once the additional climate instruments (*frost*, *humid*, and *latitude*) are included (table 2, column 1), the effect of institutional quality is found to be stronger than in the baseline specification (table 1, column 1), whereas the effect of malaria prevalence is weaker. Both differences are small and can be explained by estimation uncertainty. The Hansen test does not reject the three overidentifying restrictions that arise from the fact that the two endogenous regressors are now estimated with five instruments. Taken at face value, this test result would imply that the exogeneity restrictions on the instruments appear valid and that there are no direct effects on the level of economic development from the additional instruments.

The result of the Hansen test should be viewed with care if there are signs of a weak-instrument problem, because then the usual inference based on the χ^2 distribution of the test statistic would no longer hold. There is conflicting evidence on the presence of weak instruments. The Cragg–Donald statistic, which is much smaller than the critical value, does indicate a weak-instrument problem. But the weak-instrument test of Hahn and Hausman (2002), which can be applied only for overidentified equations, does not reject the null hypothesis

12. The reason for the low power is probably that the small sample is not informative enough to identify an independent effect of malaria prevalence in this specification.

TABLE 2. The Impact of Additional Instrumental Variables

	Instruments							
	1		2		3		4	
	Baseline + climate		Baseline + Europe		Baseline + openness		All instruments	
Explanatory variables	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Estimated coefficients								
Two-step least square	0.97	−0.90	0.86	−1.14	0.81	−1.09	0.84	−1.08
Standard error	0.16	0.29	0.16	0.27	0.16	0.29	0.13	0.25
Fuller	0.98	−0.89	0.86	−1.14	0.81	−1.09	0.84	−1.08
Standard error	0.16	0.29	0.16	0.27	0.16	0.29	0.14	0.26
Bounds of 95 percent confidence intervals								
Conventional								
Upper	0.65	−1.48	0.55	−1.69	0.49	−1.67	0.57	−1.58
Lower	1.30	−0.32	1.18	−0.59	1.13	−0.51	1.10	−0.58
Conditional likelihood ratio								
Upper	0.60	−1.60	0.46	−1.88	0.34	−1.95	0.48	−1.79
Lower	1.54	0.02	1.31	−0.39	1.27	−0.25	1.22	−0.35
Number of observations	44		44		44		44	
Number of instruments	5		4		4		9	
Hansen test (OIR)	2.27		1.05		2.38		6.08	
p-value	0.52		0.59		0.30		0.53	
First-stage statistics								
F-statistic	13.74	16.74	20.53	30.33	19.96	22.98	19.58	16.05
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Partial R^2	0.64	0.69	0.68	0.76	0.67	0.70	0.84	0.81
Shea partial R^2	0.48	0.51	0.49	0.54	0.46	0.48	0.67	0.64
Weak-instrument test[a]								
Cragg–Donald	6.29		8.49		7.21		5.97	
Critical valaue	19.45		16.87		16.87		27.51	
Hahn–Hausman test	0.21	−0.22	0.30	−0.31	−0.12	0.12	0.21	−0.21
p-value	0.83	0.83	0.76	0.76	0.91	0.91	0.83	0.83

(*Continued*)

TABLE 2. Continued

	Instruments							
	1		2		3		4	
	Baseline + climate		Baseline + Europe		Baseline + openness		All instruments	
Explanatory variables	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Instrument selection criteria[b]								
Donald–Newey MSE	0.42		0.37		0.47		0.10	
BIC	−8.99		−6.50		−3.68		−18.84	
HQIC	−5.67		−4.28		−1.46		−11.08	

Note: The sets of instruments are baseline (ln*mort* and *maleco*), climate (*frost*, *humid*, and *latitude*), Europe (*eurfrac* and *engfrac*), and openness (*coast* and *trade*).

[a]The Cragg–Donald statistic (Cragg and Donald 1993) is used by Stock and Yogo (2002) for weak-instrument tests. If the Cragg–Donald statistic exceeds the critical value, then a standard significance test with nominal size of 5 percent has a maximal size of 10 percent. The weak-instrument test by Hahn and Hausman (2002) is based on the normalized difference between bias-adjusted two-step least squares estimators (B2SLS) for an equation and its reverse equation, where the left-side variable and the endogenous right-side variable are interchanged.

[b]The Donald–Newey instrument selection criterion (Donald and Newey 2001) is the expected average mean-squared error of the 2SLS estimator. BIC and HQIC are the Bayesian and Hannan–Quinn information criteria (Andrews 1999) for the choice of instruments.

Source: Authors' analysis based on data described in the text.

of strong instruments. On balance, it can be concluded that the results of the baseline specification are not rejected by adding additional climate instruments, because significant effects of institutional quality and malaria prevalence are obtained if the robust CLR confidence intervals are used for inference.[13]

Adding the Western European instruments (table 2, column 2) and the openness instruments (table 2, column 3) to the baseline instruments also does not change the results by much. Both institutional quality and malaria prevalence exert highly significant effects, even if a potential weak-instrument problem is taken into account, as indicated by the Cragg–Donald test. Again, the overidentifying restrictions are not rejected, and the conclusions of the baseline model are confirmed, as indicated by the CLR confidence intervals. This result is not altered when all instruments are included together (table 2, column 4). The parameters remain highly significant, and the Hansen test still does not reject the overidentifying restrictions. Moreover, a Hansen difference test cannot reject the additional overidentifying restrictions of column 4 in table 2 over those of the columns 1, 2, and 3, leading to test statistics of 3.81 (p-value 0.43), 5.03 (p-value 0.41), and 3.7 (p-value 0.59). Thus, including additional instrumental variables based on climate, Western European influence, and openness does not change the conclusions relative to those obtained from the baseline specification.

As a further robustness check to see which set of instrumental variables is favored, three formal instrument selection criteria are applied—the expected average mean squared error criterion of Donald and Newey (2001), the Bayesian information criterion, and the Hannan–Quinn information criterion of Andrews (1999). All test results suggest choosing the full instrument set.[14] However, it must be taken into account that these criteria are not designed for specifications with weak instruments, where they may lead to the wrong conclusions. As a general tendency in the estimation results, the weak-instrument problem appears to increase with the number of instruments, at least if the Cragg–Donald statistic is taken as the benchmark. This may reflect an overfitting problem when using many instruments, notably in table 2, column 4, where nine instrumental variables are employed for a sample of 45 countries. Nevertheless, the point estimates are quite robust to the number of instruments included. This indicates that the results of the baseline specification are not driven by the choice of instruments and that the baseline specification should be preferred to minimize overfitting and the weak-instrument problems.

13. The 95 percent CLR confidence interval includes zero but is a borderline case. For example, a 93 percent CLR confidence interval does not include zero.

14. The values of Bayesian and Hannan–Quinn information criteria equal zero for all just-identified specifications and are therefore reported only for overidentified specifications.

Validity of Ln*mort and Maleco*

So far, the article has not questioned whether ln*mort* and *maleco* are valid instrumental variables for institutional quality and malaria prevalence. Both variables have been criticized as flawed instrumental variables in the literature. For instance, Albouy (2004) argues that Acemoglu, Johnson, and Robinson (2001) measure settler mortality (ln*mort*) imprecisely. He constructs a "high revision" variable (ln*mort2*) that he claims exhibits improved geographic relevance, statistical precision, and cross-country comparability. When Albouy (2004) re-estimates the effect of institutional quality on economic development with the revised instrument (ln*mort2*), he obtains a severe weak-instrument problem, which results in a failure to measure any statistically significant effect of institutional quality on income.

When ln*mort* is replaced by ln*mort2*, the point estimates remain almost unchanged (table 3, column 1). However, the Cragg–Donald statistic becomes quite small, indicating the presence of weak instruments. Therefore, the CLR confidence intervals are used for inference. They turn out to be very large and in the case of *malfal* to include zero. Somewhat surprisingly, the point estimate of the coefficient on institutional quality is still statistically significant, even though the first-stage statistics indicate a severe drop in explanatory power for the first-stage regression of *rule* on the instruments relative to the baseline estimate.[15]

To improve the first-stage regression results, the set of instrumental variables is augmented with the measures of Western European influence (*eurfrac* and *engfrac*). Their inclusion increases the explanatory power in the first-stage regressions considerably (table 3, column 2). However, there is still a weak-instrument problem according to the Cragg–Donald statistic. But the CLR confidence intervals indicate statistically significant effects of both institutional quality and malaria prevalence in this specification, and the point estimates remain almost unchanged compared with the baseline model.

Thus, the revised settler mortality variable constructed by Albouy (2004) exhibits a smaller degree of instrument relevance than the original variable used by Acemoglu, Johnson, and Robinson (2001). While replacing the original with the revised variable leads to almost unchanged point estimates, the statistical significance of *malfal* becomes questionable. Nevertheless, it is still possible to come up with statistically significant and economically meaningful estimates of the effects on income of institutions and malaria prevalence without using the original mortality variable if the measures of Western European influence are added to the list of instrumental variables. Thus, the baseline results do not depend on using the original mortality variable as an instrument.

Rodrik, Subramanian, and Trebbi (2004) question the exogeneity of the second baseline instrumental variable, *maleco*. Therefore, a specification is

15. In the first-stage regression of *rule*, the revised variable ln*mort2* is highly significant with a *t*-value of −3.66 but much less so than in the baseline model.

TABLE 3. The Validity of Ln*mort* and *Maleco*

	Instruments							
	1		2		3		4	
	ln*mort2* + *maleco*		ln*mort2* + *maleco* + Europe[a]		ln*mort* + Europe		ln*mort2* + Europe[a]	
Explanatory variables	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Estimated coefficients								
Two-step least square	1.12	−0.85	0.89	−1.15	0.75	−1.35	0.77	−1.39
Standard error	0.32	0.41	0.19	0.28	0.20	0.36	0.21	0.35
Fuller	1.07	−0.90	0.89	−1.15	0.75	−1.33	0.77	−1.37
Standard error	0.29	0.38	0.19	0.28	0.20	0.35	0.21	0.33
Bounds of 95 percent confidence intervals								
Conventional								
Upper	0.48	−1.69	0.51	−1.72	0.33	−2.08	0.34	−2.09
Lower	1.77	−0.02	1.27	−0.58	1.16	−0.62	1.20	−0.69
Conditional likelihood ratio								
Upper	0.34	−1.86	0.36	−1.96	−0.04	−2.72	0.01	−2.61
Lower	9.93	8.80	1.57	−0.29	1.34	−0.36	1.42	−0.46
Number of observations	45		44		44		44	
Number of instruments	2		4		3		3	
Hansen test (OIR)			1.683		0.149		0.000	
p-value			0.431		0.700		1.000	
First-stage statistics								
F-statistic	10.181	41.89	9.12	26.16	23.92	31.21	12.41	23.45
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Partial R^2	0.33	0.67	0.48	0.73	0.64	0.70	0.48	0.64
Shea partial R^2	0.13	0.27	0.35	0.52	0.29	0.32	0.27	0.36

(*Continued*)

TABLE 3. Continued

	Instruments							
	1		2		3		4	
	ln*mort2* + *maleco*		ln*mort2* + *maleco* + Europe[a]		ln*mort* + Europe		ln*mort2* + Europe[a]	
Explanatory variables	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Weak-instrument test[b]								
Cragg–Donald	3.27		5.18		4.65		4.76	
Critical value	7.03		16.87		13.43		13.43	
Hahn–Hausman test			0.10	−0.10	0.28	−0.28	0.33	0.10
p-value			0.92	0.92	0.78	0.78	0.74	0.92
Instrument selection criteria[c]								
Donald–Newey MSE	7.95		0.65		1.91		1.67	
BIC			−5.72		−3.57		−3.78	
HQIC			−3.50		−2.46		−2.67	

[a]Refers to the instruments *engfrac* and *eurfrac*.

[b]The Cragg–Donald statistic (Cragg and Donald 1993) is used by Stock and Yogo (2002) for weak-instrument tests. If the Cragg–Donald statistic exceeds the critical value, then a standard significance test with nominal size of 5 percent has a maximal size of 10 percent. The weak-instrument test by Hahn and Hausman (2002) is based on the normalized difference between bias-adjusted two-step least squares estimators (B2SLS) for an equation and its reverse equation, where the left-side variable and the endogenous right-side variable are interchanged.

[c]The Donald–Newey instrument selection criterion (Donald and Newey 2001) is the expected average mean-squared error of the 2SLS estimator. BIC and HQIC are the Bayesian and Hannan–Quinn information criteria (Andrews 1999) for the choice of instruments.

Source: Authors' analysis based on data described in the text.

estimated that excludes *maleco* and uses only (original) settler mortality and the measures of Western European influence as instrumental variables (table 3, column 3). The results indicate that this specification creates a weak-instrument problem according to the Cragg–Donald statistic. But the point estimates remain statistically significant[16] and quantitatively similar, compared with the baseline specification. This suggests that finding a significant influence of malaria prevalence on economic development does not hinge on the use of *maleco* as an instrument. Moreover, when *maleco* is included as an additional instrumental variable, its validity is not rejected by a Hansen difference test,[17] which provides statistical evidence against the endogeneity of *maleco* as presumed by Rodrik, Subramanian, and Trebbi (2004). Basically, the same findings emerge if the revised settler mortality variable (ln*mort2*) is used (table 3, column 4).[18]

Sample Size

As a final robustness check, results for a larger sample of countries are reported in appendix (table A.1). First, the baseline specification is re-estimated (column 1). Compared with the baseline sample of 45 countries (table 1, column 1), the point estimates and interval estimates are by and large the same. Thus, independent of the sample of countries, the estimated coefficients on both institutional quality and malaria prevalence appear economically important and statistically significant. Also as before, the Cragg–Donald statistic does not signal a weak-instrument problem for this specification. In contrast to the previous estimates, there is a reduced fit of the first-stage regressions that may reflect the presumed weak data quality for some of the countries in the larger sample, especially for the governance index (*rule*).

Similar conclusions can be drawn from the Acemoglu, Johnson, and Robinson (2001) specification in column 2 of table A.1, where institutional quality is measured by expropriation risk (*exprop*). Compared with the estimates for the smaller sample of countries (table 1, column 4), the point estimates change slightly. The most important result, the statistical insignificance of the coefficient on malaria prevalence, also shows up in the large sample. Again, this can be traced to a severe weak-instrument problem, which affects mainly the first-stage equation for *exprop*.

Using additional instrumental variables also does not change any previous insights for the larger sample of countries. In column 3 of table A.1, results are presented for the baseline specification augmented by all available instruments. The point and interval estimates change only slightly compared with the previous estimates, both

16. The 95 percent CLR confidence interval for institutional quality includes zero but is a borderline case. For example, a 93 percent CLR confidence interval does not include zero.

17. Detailed results are available on request.

18. A potential problem with the specification reported in table 3, column 4, is signaled by the Hansen statistic that is essentially zero. This could indicate that the instruments are highly correlated with each other or that there are too many instruments. However, the correlation of the instruments is far below one, and each instrument is significant in at least one first-stage regression. In fact, ln*mort2* and *engfrac* are highly significant in the regression for *rule*, and ln*mort2* and *eurfrac* are highly significant in the regression for *malfal*.

institutional quality and malaria prevalence remain statistically significant, and the overidentifying restrictions are not rejected. Finally, using the revised settler mortality variable (ln*mort2*) and excluding malaria ecology (*maleco*) from the set of instrumental variables again leads to almost identical results in the large sample (table A.1, column 4) compared with the smaller sample (table 3, column 4).

The conclusion from the robustness checks is that the results of the baseline specification are not sensitive to changes in the explanatory variables, the set of instruments, or the sample of countries.[19] Therefore, the hypothesis derived from the baseline specification is maintained. Both institutions and malaria prevalence appear economically important determinants of the level of development.

V. Conclusion

The reported empirical results suggest that the Sachs hypothesis of direct income effects of malaria prevalence cannot be dismissed as easily as claimed in recent studies by Acemoglu, Johnson, and Robinson (2001) and Rodrik, Subramanian, and Trebbi (2004). Different from Acemoglu, Johnson, and Robinson (2001), statistically significant effects on income of malaria prevalence are estimated once its potential endogeneity is controlled for, and there appears to be no empirical evidence that the instrumental variable used by Sachs (2003) is invalid, as presumed by Rodrik, Subramanian, and Trebbi (2004). For given effects of institutional quality, the estimated direct negative income effects of malaria prevalence are quantitatively important. This result appears to be robust to using alternative measures of institutions and malaria prevalence, alternative and additional instrumental variables, and an alternative sample of countries.

Taken at face value, the results imply that institutions do not dominate all other potential determinants of development. An emphasis on good governance, even if it can be successfully implemented in poor countries, will probably not suffice to achieve improved economic performance. As argued by Sachs and his coauthors in various papers, subsidized research on tropical diseases and direct assistance from foreign donors for interventions against diseases may be needed to advance the development of poor countries, which otherwise may not escape the restrictions imposed on them by adverse geographic endowments. All this is certainly not to deny that good institutions would make such interventions possible in the first place or at least would make them more productive, but the findings of this article point out that good institutions alone are not necessarily a sufficient recipe for successful economic development.

19. Regarding variations in the sample size, the inclusion of sub-Saharan African countries is key to the presented identification strategy and to the validity of the instrumental variables, as it is in Acemoglu, Johnson, and Robinson (2001). The reason is that the variance across the sample stems mainly from the difference between the sub-Saharan African countries and the other countries. The reported direct effects of malaria on income may thus explain why some previous cross-country studies reported a negative coefficient of the dummy variable for sub-Saharan Africa (Collier and Gunning 1999).

APPENDIX. Definitions and Sources of Variables

Variable	Definition	Source
coast	Proportion of land area within 100 kilometers of the sea coast	Gallup, Sachs, Mellinger (1999), here taken from McArthur and Sachs (2001)
engfrac	Proportion of the population speaking English	Hall and Jones (1999)
eurfrac	Proportion of the population speaking one of the major languages of Western Europe: English, French, German, Portugese, or Spanish	Hall and Jones (1999)
exprop	Index of protection against expropriation in 1985–95; limited to 64 countries but includes the Bahamas and Vietnam, which are not included in *socinf*; measured on a scale of 1 to 10	Acemoglu, Johnson, and Robinson (2001), p. 1398
frost	Proportion of a country's land receiving 5 or more frost days in that country's winter, defined as December through February in the Northern hemisphere and June through August in the Southern hemisphere; measured on a scale of 0 to 1	Masters and McMillan (2001)
gadp	Index of government antidiversion policies; calculated as an unweighted average of five variables: law and order, bureaucratic quality, corruption, risk of expropriation, and government repudiation of contracts; measured on a scale of 0 to 1	Hall and Jones (1999)
humid	Highest temperature during the month when average afternoon humidity is at its highest; measured in degrees Celsius	Parker (1997)
latitude	Distance from the equator as measured by the absolute value of country-specific latitude in degrees	Hall and Jones (1999)
ln*gdpc*	Real GDP per capita, adjusted for purchasing power parity (PPP), 1995; measured in international dollars	World Bank, Development Indicators CD-ROM (2002)
ln*mort*	Settler mortality rates in colonies in the early nineteenth century, fourth mortality estimate (72 countries, excluding France and the United Kingdom); measured as death rate among 1,000 settlers, where each dead settler is replaced with a new settler	Acemoglu, Johnson, Robinson (2001, p. 1398), and Acemoglu, Johnson, Robinson (2000)
ln*mort2*	Revised estimate of ln*mort*	Albouy (2004)

(*Continued*)

APPENDIX. Continued

Variable	Definition	Source
maleco	Combines climatic factors and biological properties of the regionally dominant malaria vector into an index of the stability of malaria transmission, which is called malaria ecology; the index of malaria ecology is measured on a highly disaggregated subnational level and then averaged for the entire country and weighted by population; the index ranges from 0 to 31.5 (Burkina Faso); for details see text; dataset as of October 27, 2003	Kiszewski and others (2004), here taken from http://www.earth.columbia.edu/about/director/malaria/index.html#datasets
malfal	Proportion of a country's population at risk of falciparum malaria transmission in 1994; measured on a scale of 0 to 1. Revised dataset as of October 27, 2003	Sachs (2003), here taken from http://www.earth.columbia.edu/about/director/malaria/index.html#datasets
malrisk	Proportion of each country's population that lives with the risk of malaria transmission, involving three largely nonfatal species of the malaria pathogen (*Plasmodium vivax*, *Plasmodium malariae*, and *Plasmodium ovale*); measured on a scale of 0 to 1; dataset as of October 27, 2003	Sachs (2003), here taken from http://www.earth.columbia.edu/about/director/malaria/index.html#datasets
rule	Average governance indicator based on six aggregated survey measures: voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption	Kaufmann Kraay, and Mastruzzi (2004)
trade	Natural log of the Frankel–Romer predicted trade share, based on measures of population size and geography	Hall and Jones (1999)

TABLE A-1. Results for a Larger Sample of Countries

	Instruments							
	1		2		3		4	
	ln*mort* + *maleco*		ln*mort* + *maleco*		All instruments		ln*mort2* + Europe[a]	
Explanatory variables	*rule*	*malfal*	*exprop*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Estimated coefficients								
Two-step least square	0.92	−0.95	0.61	−0.81	0.78	−1.09	0.77	−1.26
Standard error	0.24	0.31	0.19	0.40	0.15	0.24	0.23	0.32
Fuller	0.90	−0.96	0.57	−0.88	0.78	−1.09	0.77	−1.25
Standard error	0.24	0.30	0.17	0.36	0.15	0.24	0.23	0.31
Bounds of 95 percent confidence intervals								
Conventional								
Upper	0.43	−1.57	0.23	−1.61	0.48	−1.56	0.30	−1.90
Lower	1.40	−0.33	0.99	−0.01	1.08	−0.62	1.24	−0.62
Conditional likelihood ratio								
Upper	0.40	−1.62	0.31	−1.49	0.38	−1.75	0.07	−2.22
Lower	1.64	−0.11	2.25	2.25	1.22	−0.45	1.52	−0.35
Number of observations	64		62		61		61	
Number of instruments	2		2		9		3	
Hansen test (OIR)					5.37		0.17	
p-value					0.61		0.68	
First-stage statistics								
F-statistic	19.51	58.65	11.17	56.76	15.59	23.39	12.44	25.93
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Partial R^2	0.39	0.66	0.28	0.66	0.73	0.80	0.40	0.58

(*Continued*)

TABLE A-1. Continued

	Instruments							
	1		2		3		4	
	ln*mort* + *maleco*		ln*mort* + *maleco*		All instruments		ln*mort2* + Europe[a]	
Explanatory variables	*rule*	*malfal*	*exprop*	*malfal*	*rule*	*malfal*	*rule*	*malfal*
Shea partial R^2	0.21	0.35	0.10	0.25	0.56	0.61	0.24	0.35
Weak-instrument test[b]								
Cragg–Donald	7.97		3.36		6.68		5.99	
Critical value	7.03		7.03		27.51		13.43	
Hahn–Hausman test					0.37	−0.38	0.29	−0.30
p-value					0.71	0.71	0.77	0.77

[a]Refers to the instruments *engfrac* and *eurfrac*.

[b]The Cragg–Donald statistic (Cragg and Donald 1993) is used by Stock and Yogo (2002) for weak-instrument tests. If the Cragg–Donald statistic exceeds the critical value, then a standard significance test with nominal size of 5 percent has a maximal size of 10 percent. The weak-instrument test by Hahn and Hausman (2002) is based on the normalized difference between bias-adjusted two-step least squares estimators (B2SLS) for an equation and its reverse equation, where the left-side variable and the endogenous right-side variable are interchanged.

Source: Authors' analysis based on data described in the text.

REFERENCES

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2000. "The Colonial Origins of Comparative Development: An Empirical Investigation." Working Paper 00-22. Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.

———. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–401.

———. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4):1231–94.

Albouy, David. 2004. "The Colonial Origins of Comparative Development: A Reexamination Based on Improved Settler Mortality Data." Department of Economics, University of California, Berkeley.

Andrews, Donald W. K. 1989. "Power in Econometric Applications." *Econometrica* 57(5):1059–90.

———. 1999. "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation." *Econometrica* 67(3):543–64.

Bloom, David E., and Jeffrey D. Sachs. 1998. "Geography, Demography, and Economic Growth in Africa." *Brookings Papers on Economic Activity* 2:207–95.

Collier, Paul, and Jan Willem Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37(1):64–111.

Cragg, John G., and Stephen G. Donald. 1993. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory* 9(2):222–40.

Donald, Stephen G., and Whitney K. Newey. 2001. "Choosing the Number of Instruments." *Econometrica* 69(5):1161–91.

Dufour, Jean-Marie. 1997. "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models." *Econometrica* 65(6):1365–87.

Easterly, William, and Ross Levine. 2003. "Tropics, Germs, and Crops: How Endowments Influence Economic Development." *Journal of Monetary Economics* 50(1):3–39.

Engerman, Stanley L., and Kenneth I. Sokoloff. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States." In Stephen Haber, ed., *How Latin America Fell Behind*. Stanford, Calif.: Stanford University Press.

Frankel, Jeffrey A., and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89(3):379–99.

Fuller, Wayne A. 1977. "Some Properties of a Modification of the Limited Information Estimator." *Econometrica* 45(4):939–54.

Gallup, John Luke, Jeffrey D. Sachs, and Andrew Mellinger. 1999. "Geography and Economic Development." *International Regional Science Review* 22(2):179–232.

Gallup, John Luke, and Jeffrey D. Sachs. 2001. "The Economic Burden of Malaria." *The American Journal of Tropical Medicine & Hygiene* 64(1–2, Suppl.):85–96.

Hahn, Jinyong, and Jerry Hausman. 2002. "A New Specification Test for the Validity of Instrumental Variables." *Econometrica* 70(1):163–89.

Hahn, Jinyong, Jerry Hausman, and Guido Kuersteiner. 2004. "Estimation with Weak Instruments: Accuracy of Higher-Order Bias and MSE Approximations." *Econometrics Journal* 7(1):272–306.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1):83–116.

Hausman, Jerry, James H. Stock, and Motohiro Yogo. 2005. "Asymptotic Properties of the Hahn-Hausman Test for Weak Instruments." *Economics Letters* 89(3):333–342.

Hoff, Karla. 2003. "Paths of Institutional Development: A View from Economic History." *The World Bank Research Observer* 18(2):205–26.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2004. "Governance Matters III: Governance Indicators for 1996–2002." *World Bank Economic Review* 18(2):253–87.

Kiszewski, Anthony, Andrew Mellinger, Pia Malaney, Andrew Spielman, Sonia Ehrlich Sachs, and Jeffrey D. Sachs. 2004. "A Global Index of the Stability of Malaria Transmission Based on the Intrinsic Properties of Anopheline Mosquito Vectors." *American Journal of Tropical Medicine and Hygiene* 70(5):486–98.

Masters, William A., and Margaret S. McMillan. 2001. "Climate and Scale in Economic Growth." *Journal of Economic Growth* 6(3):167–86.

McArthur, John W., and Jeffrey D. Sachs. 2001. *Institutions and Geography: A Comment on Acemoglu, Johnson, and Robinson (2000)*. NBER Working Paper 8114. National Bureau of Economic Research, Cambridge, Mass.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71(4):1027–48.

Parker, Philip M. 1997. *National Cultures of the World. A Statistical Reference*. Westport, Conn.: Greenwood Press.

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. 2004. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth* 9(2):131–65.

Sachs, Jeffrey D. 2001. *Tropical Underdevelopment*. NBER Working Paper 8119. National Bureau of Economic Research, Cambridge, Mass.

———. 2003. *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income*. NBER Working Paper 9490. National Bureau of Economic Research, Cambridge, Mass.

Sachs, Jeffrey D., and Pia Malaney. 2002. "The Economic and Social Burden of Malaria." *Nature* 415(6872):680–85.

Shea, John. 1997. "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *Review of Economics and Statistics* 79(2):348–52.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables with Weak Instruments." *Econometrica* 65(3):557–86.

Stock, James H., and Motohiro Yogo. 2002. *Testing for Weak Instruments in Linear IV Regressions*. NBER Technical Working Paper 284. National Bureau of Economic Research, Cambridge, Mass.

Summers, Robert, and Alan Heston. 1991. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988." *Quarterly Journal of Economics* 106(2):327–68.

World Bank. 2002. *World Development Indicators 2002*. CD-ROM. Washington, D.C.