The Primacy of Institutions Reconsidered:

The Effects of Malaria Prevalence in the Empirics of Development*

by

Kai Carstensen and Erich Gundlach

Kiel Institute for World Economics, Germany

revised, August 2005

Corresponding author:
Erich Gundlach
Kiel Institute for World Economics
P.O. Box 4309
24100 Kiel
Germany

email: egundlach@ifw.uni-kiel.de
phone: +49 431 8814284
fax: +49 431 8814500

*We would like to thank three anonymous referees, seminar participants at Aarhus University and at Goethe University in Frankfurt, 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 version; as well as William A. Masters for providing the data on frost frequency, Gordon McCord for advice on the malaria data, and Stephen Donald for a sample program on instrument selection.

The Primacy of Institutions Reconsidered:
The Effects of Malaria Prevalence in the Empirics of Development

ABSTRACT

Some recent empirical studies deny any direct performance effects of measures of geography and conclude that institutions trump all other potential determinants of development. For given effects of institutional quality, our empirical results indicate quantitatively important direct negative performance effects of a measure of disease ecology, namely malaria prevalence. This finding appears to be robust to using alternative specifications, instrumentations, and samples. We conclude from our estimates that implementing good institutions appears to be necessary but not sufficient to generate a persistent process of successful economic development.

Key Words: Economic development, institutions, malaria prevalence, instrument selection

JEL Classification: O1, O4

1. Motivation

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 respected strand of the literature has emphasized the preeminent role of physical geography in explaining cross-country differences in the level of development. Earlier studies along this line of reasoning include Lee (1957) and Kamarck (1976), more recent contributions include Landes (1998), Olsson and Hibbs (2005), and Diamond (1997). By contrast, the vast empirical growth literature of the last two decades has largely neglected physical geography as a relevant dimension of analysis, with the notable exception of Masters and McMillan (2001). What is more, some recent empirical studies that do consider measures of geography deny any direct impact of geography on the level of economic development (Hall and Jones 1999, Acemoglu et al. 2001, Easterly and Levine 2003, Rodrik et al. 2004).

We think that these four studies (henceforth HJ 1999, AJR 2001, EL 2003, and RST 2004) are key contributions to the empirical literature in emphasizing the primacy of institutions over geography and other potential determinants of development. All four studies suggest that measures of geographic endowments will not affect the level of development directly but only through their effect on institutions. This line of reasoning is further supported by Engermann and Sokoloff (1997) and Acemoglu et al. (2002) who discuss, for historical examples from the Americas, how geographic endowments may have shaped factor endowments (persons per unit of land), and how unequal factor endowments in turn may have shaped persistent institutions imposed by the colonizing powers that have enabled the entrenchment of a small group of elites. Given that the Industrial Revolution required a broad participation of the population in entrepreneurship and innovation, economies which started with more equal factor endowments (due to geographic endowments) and hence with institutions that resulted in a less unequal distribution of income and wealth probably therefore began to realize faster persistent per capita growth than more unequal economies. Hence this view also emphasizes that geographic endowments only affect the level of development through their impact on factor endowments and institutions, but not directly.

While development economists can easily agree on the relevance of good institutions for successful development, and on the indirect role that physical endowments may have played 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, especially Jeffrey Sachs and coauthors have argued in a series of papers that measures of geography that focus on the 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 et al. 1999, Gallup and Sachs 2001, Sachs 2001, McArthur and Sachs 2001, Sachs and Malaney 2002, Sachs 2003). The main disagreement in the present debate is about the robustness of the empirical evidence presented by Sachs and his coauthors, which is in conflict with the key studies that favor the primacy of institutions and has been directly rejected by AJR (2001) and RST (2004).

With our contribution, we do not primarily attempt to refute a specific empirical result of the four key papers in the ongoing debate. Alternative samples, specifications, and instrumentations can easily lead to a multiplicity of sometimes contradictory empirical results. In the absence of an empirical model that could serve as an unambiguous measure of reference, such differences in empirical results are difficult to evaluate. Rather than focusing on specific reported findings, we use a parsimonious baseline specification reconsider the general econometric limitations of alternative empirical strategies that have been used to derive clear-cut conclusions with regard to the deep determinants of development. We argue that the recent key papers have not treated geographic variables such as measures of disease ecology in the same way as measures of institutions, thus giving the geography hypothesis probably a smaller chance to prevail.

Our major contribution 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. In our view, the jury is still out on this specific question. Mainly to keep the empirical analysis tractable in the presence of a rather limited amount of plausible instrument variables, we ignore other explanatory variables than institutions and disease ecology, not least because such measures likes the quality of economic policies (EL 2003) or the level of trade integration (RST 2004) have not been found to exert a direct effect on the level of development independent from the effect of institutions. Instead we emphasize the fundamental problems of statistical inference implied by instrumental variable estimation and try to trace out the basic reason for the different empirical results regarding the direct role of measures of disease ecology. Finding a robust direct effect of malaria prevalence, which is

held to be a proxy for the adverse disease ecology of a country, would substantially matter for devising appropriate development policies.

For instance, foreign aid may mainly be targeted on initiating policy reform and on improving institutions in impoverished countries if there is no empirical evidence that malaria prevalence directly affect the level of development. But given that there are such direct performance effects of the disease ecology; foreign aid may also be spend on solving biophysical or 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.

2. The Many Possible Links Between Geography, Institutions, and Development

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

The common leitmotiv of the four key papers is that IVs are needed to identify the direct development effects of institutions and other endogenous explanatory variables. An IV 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 is here the level of income per capita. The solid arrows with a number attached to them are meant to indicate that the IV method in principle identifies the true causal effect of the endogenous variables on the level of income. For instance, one may obtain unbiased estimates of the effects of institutions (1), disease ecology (2), and probably other variables (3) on the level of income without having to identify the potential reverse causality from the level of income to the explanatory variables, given that valid IVs are available. In addition, the IV method can be used in principle to identify the web of relations between the endogenous explanatory variables, which are represented by the solid arrows (4)-(9).

Taking all possibilities into account, specific variables may affect the level of income directly and through their effects on other variables, or only through one channel, or not at all. In any case, empirical estimates of the various potential causal effects will crucially depend on the quality of the available IVs. IVs have to be correlated with the endogenous variables, are not allowed to affect the dependent variable, and have to be independent from the dependent variable. Hence almost by definition, valid IVs are difficult to come by in the cross-country empirics of development, just because most economic variables are affected, one way or another, by the level of income. Measures of geography play a special role in this context because most of them, like distance from the equator or temperature, can undoubtedly be considered as exogenous to the level of development and to the endogenous explanatory variables.

Apart from the required correlation between IVs and the endogenous explanatory variables, which has to be taken into account to avoid the weak instruments problem, the more fundamental problem is that once a measure of geographic endowments has a direct effect on the level of income, it would no longer qualify as a valid IV. For instance in Figure 1, geography as measured by location may affect the level of income either directly (solid arrow (10)) or through its effects on other variables such as indicated by the solid arrows (11)-(13). In case (10) identifies a causal effect, the specific measure of location cannot be used as an IV. And if the remaining available IVs are also mainly measures of geographic endowments, 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.

AJR (2001) estimate a specification where institutions and disease ecology (malaria prevalence) determine the income level. Since they do not instrument their measure of disease ecology but nevertheless do not find a statistically significant effect, they conclude that institutions (arrow (1)) but not disease ecology (arrow (2)) affect the level of income. Apart from their empirical results, AJR (2001) a priori dismiss the possibility that a measure of disease ecology like malaria prevalence could actually have a large effect on the level of development. They argue that tropical diseases like malaria are unlikely to be the reason why many countries in Africa and Asia are very poor today just because people in areas where malaria is endemic may have developed various types of immunities against such diseases. According to this view, strong performance effects of malaria are implausible because malaria is a debilitating rather than fatal disease, with the risk of malaria severity and death mainly

limited to non-immunes such as children below the age of five and adults who grew up elsewhere, like European settlers.[1]

But even if the health effects of malaria on the adult population cannot justify large performance effects, which is at least open to debate[2], one form of immunity against malaria also comes at a cost for the adult population. The sickle cell trait provides protection against malaria without serious health complications when inherited from one parent, but the same allele inherited from both parents is fatal, leading to sickle cell anemia.[3] Sickle cell anemia generates severe pain episodes and increasing infections, i.e. outcomes that are at least comparable to the direct negative health effects of malaria experienced by non-immunes. 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 appear to suggest that up to 40 percent of the population in tropical Africa may carry the sickle cell trait.[4] Hence given that a high degree of malaria prevalence reflects natural selection in favor of a high degree of prevalence of the sickle cell trait, there is at least an additional possibility that geographic endowments may affect economic performance directly through poor health and absenteeism of the workforce, and not only through their possible indirect effect on the adoption of specific institutions.

Such indirect effects of geographic endowments are especially highlighted by Engerman and Sokoloff (1997). In terms of Figure 1, they argue that geographic endowments may determine factor endowments (arrow (13)), which may entrench persistent institutions (arrow (5)) that

---

1 The vector that transmits malaria from human to human is a mosquito, more specifically the Anopheles mosquito. The Anopheles mosquito must first bite an infected person who is sick with malaria. Then the Anopheles mosquito must survive several days while the malaria parasite develops in its body. Finally, the infected Anopheles 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. Hence 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 finds excellent breeding conditions. Most tropical regions combine all these favorable preconditions 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 also less robust. However, differences in the potential for malaria transmission do not only exist across but also within climatic zones, as is shown by a new measure called the stability of malaria transmission (Kiszewski and Sachs et al. 2004).

2 See, e.g., Arrow (2004) or WHO (2005).

3 For information on sickle cell anemia, see http://www.scinfo.org/ (August 2005).

4 See http://www.pbs.org/wgbh/evolution/library/01/2/l_012_02.html (April 2004).

may influence the level of development (arrow (1)). In line with this hypothesis, RST (2004) consider whether a measure of location may affect the level of development directly (arrow (10)) or indirectly through measures of institutions and trade integration (arrows (12) and (13)). They find only weak statistical evidence for a direct effect of a measure of geography, whereas a measure of the quality of institutions appears to trump all other explanatory variables. Their conclusion is confirmed by EL (2003), who do not use a measure of geographic endowments in their specification but instead employ an overidentification test to see whether their set of geography-based IVs can be considered as valid (arrow (OIR)). EL (2003) find no evidence that any of their IVs should be included in their specification, which also appears to support the hypothesis of the primacy of institutions.[5]

We follow Sachs (2003) in using the both prevalence of malaria and the quality of institutions as endogenous explanatory variables to put the geography hypothesis on a more equal footing with the primacy of institutions hypothesis. As can be seen from Figure 1, this approach differs from the approaches by HJ (1999), AJR (2001), EL (2003), and RST (2004), which either do not instrument their measure of endogenous disease ecology, use an exogenous measure of geography, or only rely on an overidentification test. We are mainly interested in the relative size of the causal effects that are indicated by arrows (1) and (2) in Figure 1. Due to previous results reported in the literature and due to a general shortage of plausible IVs, we ignore all possible effects that might result from other explanatory variables than institutions and disease ecology, and we also ignore the possible links between institutions and disease ecology. That is, we do not consider the effects indicated by arrows (3)-(13). Different from Sachs (2003), we use additional IVs, apply recently developed econometric tests to check the validity of our point estimates and standard errors in the presence of weak IVs, and we test for overidentification restrictions (OIR) to avoid a potentially unjustified exclusion of exogenous measures of geography from our basic specifications.

3. Specification and Data

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

[5] See HJ (1999) for the same line of reasoning based on overidentification tests.

$$(1) \qquad LNGDPC_i = \beta_1 + \beta_2 \cdot INSTITUTIONS_i + \beta_3 \cdot MALARIA_i + \varepsilon_i ,$$

where ε_i is an error term with zero mean and common variance, and β_2 and β_3 are the coefficients of interest. Our specific research question is whether an estimate of β_3 , as represented by the solid arrow (2) in Figure 1, is statistically different from zero, negative, and quantitatively important. In order to better understand where the different results in the literature may come from, we re-estimate (1), paying particular attention to the choice of the variables, the instruments and the country sample.

The Choice of the Variables

To measure the effects of institutions and geography on the level of economic development, it is necessary to choose indicator variables. Such indicator variables are necessarily incomplete and erroneous because the three concepts are multidimensional and difficult to measure. Therefore, we do not concentrate on a single specification but rather use different alternative indicator variables. As the dependent variable, we take either the log of GDP per capita in 1995 (*lngdpc*), which is used by AJR (2001), EL (2003) and RST(2004), or the log of GDP per working-age person in 1990 (*lngdpw*), which appears to be more closely related to the applied growth literature and is used by HJ (1999).[6]

To measure institutional quality, we take one out of the following three variables: the rule-of-law index for 1996 (*rule*) from Kaufmann et al. (2003) used by EL (2003) and RST (2004), the index of government antidiversion policies in 1986-1995 (*gadp*) used by HJ (1999), and the index of the protection against expropriation in 1985-95 (*exprop*) used by AJR (2001). To measure disease ecology, we employ two measures of malaria prevalence, either the proportion of a country's population at risk of malaria falciparum transmission in 1994 (*malfal*) used by AJR (2001) or a new index of malaria risk (*malrisk*) suggested by Sachs (2003). The new index is based on the prevalence of non-fatal species of the malaria pathogen (Plasmodium vivax, P. malariae, P. ovale). 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 AJR (2001). The reason is that the effects of measures of malaria falciparum may be difficult to separate from the effects of continental dummies because Sub-Saharan Africa has a high proportion of malaria falciparum cases, whereas a

[6] See the Appendix for detailed descriptions of the data and sources.

relatively higher proportion of malaria vivax is reported for the Americas, Europe, and much of Asia.

The Choice of the Instruments

We start with the following two premises, both suggested by AJR (2001), to find an instrument for institutional quality. First, studying the impact of institutions on the level of development has to focus on a sample of 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 present variation in the institutional frameworks without being influenced by present economic conditions, and it should not affect by itself the present level of development. In this context, European settler mortality in the early 19th century appears to be the most plausible instrumental variable that has been suggested to date.

Differences in early settler mortality across colonies, which were well known in Europe at the time of settlement, 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 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 were large-scale settlement was not feasible for Europeans due to an unfavorable disease ecology and high rates of mortality, the colonial powers imposed a different set of institutions that did not protect private property and did not provide protection against expropriation but instead mainly focused on the extraction of natural resources. Since early settler mortality is certainly independent from present economic conditions, and since early institutional frameworks have proved to be fairly persistent over time (AJR 2001), settler mortality across former colonies can be used as an instrumental variable that helps to identify the exogenous cross-country variation in present institutional frameworks.

In order to control for the endogeneity of malaria prevalence, we consider a new measure of malaria ecology (*maleco*) that was developed by Kiszewski and Sachs et al. (2004) and first used for cross-country regressions by Sachs (2003). Since this measure of malaria ecology is only built upon climatic factors and specific biological properties of each regionally dominant

malaria vector, Kiszewski and Sachs et al. (2004) argue 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.

More specifically, the index of malaria ecology developed by Kiszewski and Sachs et al. (2004) is meant to measure the contribution of regionally dominant vector mosquitos to the *potential* transmission intensity of malaria. Hence it includes regions where malaria is not currently transmitted, but where it had been transmitted in the past or where it might be transmitted in the future.[7] Since the region-specific dominant malaria vector only reflects the forces of biological evolution, it can be considered as independent from present economic conditions. That is, terms likely to be affected by economic conditions or public health interventions (like mosquito abundance, for example) do not enter the calculation of the index. The calculated index reveals that due to vector specific 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 and human biting, and here the parasite has less fatal infectious consequences.

However, RST (2004) doubt that *maleco* is actually exogenous to present economic conditions. They criticize that Sachs (2003) does not go into the details of the construction of the index and point out that Kiszewski and Sachs et al. (2004) do not discuss exogeneity at all.[8] While this critique is technically correct, we do not think that doubts regarding the exogeneity of *maleco* are justified. Nevertheless, it would be reassuring if our results could be

---

[7] Abstracting from all detail, the construction of the index proceeds in two basic steps. First, the regionally dominant vector Anopheles 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 according to $\left(a_i^2 p_i^E\right)/\left(-\ln p_i\right)$, 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 mainly depends on average temperature and differs between P. falciparum and P. 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 specific biological properties of the region-specific dominant vectors.

[8] Online information on the construction of the malaria transmission index (malaria ecology) is available at http://www.earth.columbia.edu/about/director/malaria/index.html (August 2005). 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 only enters the index of malaria ecology 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.

based on additional instruments as well. Therefore, we supplement the two baseline instruments *lnmort* and *maleco* with three additional sets of instruments that relate to the climatic environment, the influence of Western European languages, and the openness of a country.

Temperature, rainfall, and latitude are additional measures of the climatic environment that can be related to the 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, usually due to rainfall ending up in puddles and the like. Hence the prevalence of frost (*frost*), measured as the proportion of a country's land receiving five or more frost days in that country's 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 country (*latitude*) may also be used as a proxy for a specific climatic environment. What has to be taken into account is that these three measures of climatic conditions may not only be good instruments for measures of disease ecology, but also for measures of institutional quality. This is because settler mortality and thus the design of early institutions was accordingly influenced by the prevailing climatic conditions of the colonies. In this context, AJR (2001) point out that their result validates using absolute distance from the equator as an instrument for a measure of institutions, as in HJ (1999).

Other plausible IVs than measures of geography are more difficult to come by. One possibility is to consider the fraction of the population that speaks a Western European language (*eurfrac*) or English (*engfrac*) as the first language as IVs for measures of institutions. As suggested by HJ (1999), these variables may reflect the different degree of Western European influence on the sample of former colonies and thus can probably help to identify the exogenous variation in measures of institutions. Since AJR (2001) generally question the exogeneity of these variables, we provide some formal tests of their exogeneity when using them for our robustness checks.

Further IVs that we use for our robustness checks relate to the openness to trade of a country. More open countries may have better institutions because openness may encourage less

arbitrary government behavior, especially with regard to property rights. Hence exogenous measures of openness could probably be used as valid IVs for measures of institutions. We use two proxy measures for openness, namely the proportion of land area of a country that is within 100 km of the open sea 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 mainly uses geographical variables to explain actual bilateral trade flows.

The Choice of the Sample

Our sample of countries is limited to former colonies for which data on early settler mortality are available. AJR (2001, Tab. 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 (QJE 1991)), 2 countries that are very small (less than one million inhabitants in 1990), and 1 country that depends mainly on oil production. We delete these countries from the AJR dataset and thus report baseline results for a smaller but probably more reliable sample of 45 former colonies that are neither statistical terra incognita, nor small, nor dependent on oil production. In contrast, previous papers that take issue with the AJR result of the primacy of institutions (McArthur and Sachs 2002, EL 2003, Sachs 2003, RST 2004) increase the AJR sample size but do not pay attention to the data quality. As a robustness check of our baseline findings, we include results based on a larger sample of countries that uses additional observations on settler mortality reported in Acemoglu et al. (2000).

4. Baseline Estimation Results

In a first step, we estimate (1) by two-step least squares (2SLS) using *lngdpc* as dependent variable, *rule* and *malfal* as explanatory variables, and *lnmort* and *maleco* as IVs. We use this as our baseline specification because it is close to the previous literature. In particular, *lngdpc* is used as dependent variable by AJR (2001), EL (2003) and RST(2004), *rule* is used as a measure of institutional quality by EL (2003) and RST (2004), and *malfal* is used as a measure of malaria prevalence by AJR (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 a one-unit increase

in the rule-of-law index. Hence, $\beta_2 = 0.89$ implies that a difference of 0.1 in the rule-of-law index is associated with a 8.9 percent cross-country difference in output per capita. To see the potential magnitude of the estimated effect of the measure of institutions on economic performance, we compare two countries which represent about the 70th and the 30th percentile of the rule-of-law index in the sample, say South Africa with an index value of 0.21 and Ecuador with an index value of –0.40. This difference in the rule of law 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 actual per capita GDPs differ by a factor of about 2.7 in our sample.

The point estimate of β_3 reflects the change in log output per capita associated with a one-unit increase in malaria prevalence. Hence, our point estimate of $\beta_3 = -1.04$ predicts that the per capita GDPs of Paraguay and Pakistan, which roughly represent 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 actual per capita GDPs of these two countries differ by a factor of about 2.6 in our sample. Hence, for these two country pairs, the differences in institutional quality and malaria prevalence each explain roughly one-half of the respective differences in per capita GDPs.

Our point estimates are statistically significant. In Table 1, we report estimated standard errors of 0.18 for β_2 and 0.30 for β_3, which imply *t*-statistics of 5.04 and –3.46. These values indicate statistical significance at the 5 percent level when we use *t*-tests for significance based on conventional asymptotic theory. We also report conventional confidence intervals that cover the unknown true parameters with a probability of 95 percent. Hence, we should be quite confident that the true impact of our rule-of-law index is between 0.53 and 1.25 and the true impact of our measure of malaria prevalence is between –1.65 and –0.43. However, the validity of these results hinges crucially on instrument relevance as emphasized by the recent literature on weak instruments, see, e.g., Staiger and Stock (1997) and Moreira (2003). If the IVs are only weakly correlated with the endogenous explanatory variables, the conventional asymptotic theory breaks down so that statements about statistical significance and inference

may lead to wrong conclusions. Hence we have to check whether our IVs are indeed strong enough to allow for statements about statistical significance.

The first-stage regressions supply valuable information regarding the relevance of our IVs. In Table 1, we report highly significant *F* statistics of 28.1 and 42.3 for the first-stage regressions of *rule* and of *malrisk*. 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. However, even large first-stage *F* statistics can be misleading. For example, the two instruments *lnmort* and *maleco* may not carry sufficient independent information, which could make it difficult to identify distinct effects of *rule* and *malrisk*. To this end, we compute a statistic proposed by Cragg and Donald (1993) that, loosely speaking, represents the relevance of the weakest instrument. Using weak instrument asymptotic theory, Stock and Yogo (2004) 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 we obtain a Cragg-Donald statistic of 11.45 for our baseline specification, we conclude that our results are not affected by weak instrument problems.

As a further robustness check for specifications with potentially weak instruments, we follow Hausman et al. (2004) and apply the Fuller estimator. Our Fuller point estimates are almost identical to our 2SLS estimates, so we can maintain the estimates of 0.89 for β_2 and –1.05 for β_3. In addition, we compute 95 percent confidence intervals based on inverted conditional likelihood ratio (CLR) tests that take any weak instrument problem into account (Moreira, 2003).[9] It turns out that the CLR intervals are 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, turning the CLR confidence intervals into significance tests, we can conclude that both β_2 and β_3 are individually statistically significant because the confidence intervals do not include zero.

Before proceeding with a number of further robustness checks, it is worth summarizing the results obtained with our baseline specification. Our point estimates for the effects of institutional quality and malaria prevalence have the expected signs and are economically

[9] 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 means of the projection method put forward by Dufour (1997).

important. They do not appear to suffer from a weak instruments problem and hence appear to be statistically significant as well. By identifying a direct effect of a measure of disease ecology on the level of development, our intermediate results conflicts with he evidence presented by HJ (1999), AJR (2001), EL (2003), and RST (2004), and confirms the evidence presented by Sachs (2003).

5. Robustness

We perform a number of robustness checks of the results of our baseline specification. We introduce alternative dependent and explanatory variables, we include additional IVs, we reconsider the validity of our baseline instruments *lnmort* and *maleco*, and we examine whether our results change when we use a larger sample of countries.

The Effects of Alternative Variables

In columns (2) to (4) in Table 1, we present estimation results for specifications with alternative explanatory variables. In column (2), institutional quality is still measured by the rule-of-law index (*rule*) but malaria prevalence is now measured by the risk of infection with the non-fatal malaria pathogen (*malrisk*). The main difference to the baseline specification is a smaller weight of institutional quality and a larger weight of malaria prevalence. This difference may simply be due to estimation uncertainty, which has increased compared to the baseline specification as indicated by the larger confidence intervals. Moreover, the weak instrument problem is of slightly more 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, to save space, we henceforth only report specifications with *malfal* as our measure of malaria prevalence.[10]

In column (3), the rule-of-law 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 mainly due to the smaller variance of *gadp* compared to *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

[10] Further results of the specifications with *malrisk* are available upon request.

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 actual per capita GDPs differ by a factor of about 2.7 in our sample. 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 actual per capita GDPs differ by a factor of about 2.6 in our sample. The statistical significance of these estimates can be inferred both from the conventional and the CLR confidence intervals which do not include zero. In addition, there is no weak instrument problem as indicated by a large Cragg-Donald statistic.

In column (4), the rule-of-law index used in the baseline specification as a measure of institutional quality 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 AJR (2001, Table 7) who obtain an insignificant effect of malaria prevalence with their sample of countries (without instrumenting *malfal*). Here, our point estimate of β_2 is smaller than in our baseline specification but this may be explained by the estimation uncertainty and the larger variance of *exprop* compared to *rule*. The point estimate of β_3 is virtually unchanged. At first sight, both estimates appear to be statistically significant as indicated by low standard errors (2SLS and Fuller). However, the Cragg-Donald statistic of 5.74 indicates a weak instrument problem. Hence, the 2SLS estimator may be biased and the conventional confidence intervals are 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 economic performance but rather that we cannot identify an independent effect with sufficient precision. This view can be supported by three observations. First, the instruments *lnmort* and *maleco* are strongly correlated with a correlation coefficient of 0.6, which does not leave much information in one instrument that is independent from the other. While this information appears to be sufficient for the previous specifications, it turns out to be insufficient for the present specification as indicated by lower Shea partial R^2s than before. Second, the joint hypothesis $\beta_2 = \beta_3 = 0$ can be clearly rejected by a conditional likelihood ratio test that takes the weak instruments problem into account (Moreira, 2003). Hence, while

the individual importance of a single parameter may be difficult to extract from the data, there is obviously a highly significant joint effect.

Third, the power of the significance test for β_3 (which we have derived from its estimated confidence interval) is 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 the (in our setting overly optimistic) 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). To get an idea in which region of true parameter values $\beta_3 \neq 0$ the test can be expected to accept the wrong null hypothesis $\beta_3 = 0$, we calculate an interval with power below 0.5 which 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 change to be undetected than to be detected. Moreover, 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. Our point estimate of β_3 is only -1.03. Since we have to assume that the correct power of a significance test based on weak instrument asymptotic theory is overstated in this exercise, the lower interval limits of -0.67 for power regions of below 50 percent and of -1.23 for power regions of below 95 percent are only upper bounds for the correct but unknown interval limits. Given that parameter values for β_3 like our point estimate are economically important but statistically difficult to distinguish from zero, we have to conclude that the power of the significance test for β_3 is quite low. Hence finding the coefficient estimate of β_3 to be statistically insignificant is probably due to the low power of the significance test rather than due to the unimportance of malaria prevalence for economic development. This conclusion is also corroborated by the statistically significant point estimates of β_3 in specifications (1) and (3) in Table 1.

The Effects of Additional Instruments

So far, we have only used our baseline instrument set consisting of *lnmort* and *maleco*. To exclude the possibility that our results are driven by the specific choice of instruments, we replicate the analysis using the additional instrument sets on the climatic environment (*frost*,

humid, *latitude*), on the Western European influence (*eurfrac*, *engfrac*), and on openness (*coast*, *trade*). We restrict our analysis to the baseline specification, where institutional quality is measured by the rule-of-law index (*rule*) and malaria prevalence is measured by the risk of infection with malaria falciparum (*malfal*). The results are presented in Table 2.

Column (1) of Table 2 shows the results for including the additional climate instruments. Compared to our baseline specification, the effect of institutional quality is stronger whereas the effect of malaria prevalence is weaker. Both differences are small and can be explained by estimation uncertainty. The Hansen test accepts 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 are correctly imposed on the instruments and that there are no direct effects from the additional instruments on the level of economic development.

The result of the Hansen test should be taken 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. We find 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 also be applied for overidentified equations, does not reject the null hypothesis of strong instruments. Since we obtain a significant effects both of institutional quality and malaria prevalence[11] if we use the robust CLR confidence intervals, we conclude on balance that the results of our baseline specification are not rejected by adding additional climate instruments.

Adding the Western European instruments (column (2)) and the openness instruments (column (3)) to the baseline instruments yields also does not change our results by much. In these cases 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 accepted and the conclusions of the baseline model are confirmed as indicated by the CLR confidence intervals. This is not altered when all instruments are included togehther, as reported in column (4).. The parameters remain highly significant and the Hansen test still accepts the overidentifying

[11] The 95% CLR confidence interval includes zero but is a borderline case. For example, a 93% CLR confidence interval does not include zero.

restrictions. Moreover, a Hansen difference test cannot reject the additional overidentifying restrictions of column (4) 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), respectively. We can thus conclude that the inclusion of additional instruments based on climate, Western European influence, and openness does not change the conclusions obtained from our baseline specification.

We also consider which of the instrument sets is favored by formal instrument selection criteria. To this end, we compute the expected average mean squared error (MSE) criterion by Donald and Newey (2001) as well as the Bayesian information criterion (BIC) and the Hannan-Quin information criterion (HQIC) by Andrews (1999). All test results suggest to choose the full instrument set.[12] However, one has to take into account that these criteria are also not designed for specifications with weak instruments, where they may lead to the wrong conclusions. As a general tendency, we find that the weak instrument problem appears to increase with the number of instruments, at least if the Cragg-Donald statistic is taken as the benchmark. Since we also find that the point estimates are quite robust to the number of included instruments, we would conclude from this evidence that the baseline specification should be preferred to minimize the weak instruments problem.

The Validity of Lnmort and Maleco

So far we have not questioned our baseline assumption that *lnmort* and *maleco* are valid IVs for institutional quality and malaria prevalence. Both variables have been criticized as flawed IVs in the literature. For instance, new evidence by Albouy (2004) reveals that settler mortality (*lnmort*) is measured imprecisely by AJR (2001). He constructs a "high revision" variable (*lnmort2*) that exhibits, as he argues, improved geographical relevance, statistical precision, and cross-country comparability. When Albouy (2004) re-estimates the effect of institutional quality on economic development with the revised instrument (*lnmort2*), he obtains a severe weak instrument problem which results in a failure to measure any statistically significant effect of institutional quality.

When we replace *lnmort* by *lnmort2*, we find a similar problem, as reported in column (1) of Table 3. The Cragg-Donald statistic becomes quite small, indicating the presence of weak

[12] The values of BIC and HQIC equal zero for all just-identified specifications and are therefore only reported for overidentified specifications.

instruments. Therefore, we use the CLR confidence intervals for inference. They turn out to be very large and include zero in the case of *malfal*. Somewhat surprisingly, the point estimate of the coefficient on institutional quality is still statistically significant (2SLS and Fuller), even though the first-stage statistics indicates a severe drop in explanatory power for the first-stage regression of *rule* on the instruments relative to our baseline estimate.[13]

To improve the first-stage regression results, we augment the instrument set with the Western European influence variables (*eurfrac*, *engfrac*), which should be good instruments for institutional quality as argued by Hall and Jones (1999). In fact, this increases the explanatory power in the first-stage regressions considerably (column (2)). However, according to the Cragg-Donald statistic, there is still a weak instrument problem. But in in this specification, the CLR confidence intervals once again indicate statistically significant effects of both institutional quality and malaria prevalence. Moreover, the point estimates remain almost unchanged compared to our baseline model. We conclude from this exercise that the revised settler mortality variable constructed by Albouy (2004) is less useful as an instrument than the original variable used by AJR (2001), but even though it is still possible to come up with statistically significant and economically meaningful estimates of the effects of institutions and malaria prevalence on the level of development.

RST (2004) question the exogeneity of our second IV, *maleco*. Therefore, we estimate a specification that excludes this instrument and instead only uses (original) settler mortality and the Western European influence variables, as reported in column (3) of Table 3. Our results indicate that this specification creates a weak instrument problem according to the Cragg-Donald statistic. Nevertheless, the point estimates remain statistically significant[14] and quantitatively similar to the baseline specification. Hence, our finding of a significant influence of malaria prevalence on economic development does not hinge on the use of *maleco* as an instrument. Moreover, when we include *maleco* as an additional instrument, its validity is not rejected by a Hansen difference test.[15] That is, we do not find statistical evidence for the endogeneity of *maleco* presumed by by RST (2004). This assessment remains true when we use the revised settler mortality variable (*lnmort2*) in column (4),

---

[13] Note that in the first-stage regression of *rule*, the revised variable *lnmort2* is still highly significant with a *t*-value of -3.66 but much less so than in the baseline model.

[14] The 95% CLR confidence interval for institutional quality includes zero but is a borderline case. For example, a 93% CLR confidence interval does not include zero.

[15] Detailed results are available upon request.

where we report largely unchanged point and interval estimates relative to the results in column (3).

The Sample Size

As a final robustness check, we report results for a larger sample of countries in an Appendix Table. We first re-estimate our baseline specification (column (1)). Compared to our baseline sample of 45 countries, we find more or less the same point estimates and interval estimates as in column (1) of Table 1. Hence independent of the sample of countries, the estimated coefficients on both institutional quality and malaria prevalence appear to be economically important and statistically significant. Also as before, the Cragg-Donald statistic does not signal weak instruments for this specification. What we find is a reduced fit of the first-stage regressions that appears to reflect the presumed weak data quality for some of the countries in the larger sample, especially for the measure of institutional quality (*rule*).

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

Using additional IVs also does not change any previous insights for the larger sample of countries. In column (3), we present the results for the baseline specification augmented by all available instruments. Compared to our smaller sample of countries (Table 2, column (4)), the point and interval estimates change only slightly, both institutional quality and malaria prevalence remain statistically significant, and the overidentifying restrictions are accepted. Finally, using the revised settler mortality variable (*lnmort2*) and excluding malaria ecology (*maleco*) from the instrument set again leads to almost identical results in the large sample (column (4)) compared to the smaller sample (Table 3, column (4)).

Overall, we conclude from our robustness checks that the results of our baseline model are not sensitive to changes in the explanatory variables, the instrument set, and the sample of countries. Therefore, we maintain the hypothesis derived from our baseline model. Both

institutions and malaria prevalence appear to be economically important determinants of the level of development.

6. Conclusion

Our empirical results suggest that the Sachs hypothesis of direct performance effects of malaria prevalence cannot be dismissed as easily as claimed in recent studies by AJR (2001) and RST (2004). Different from AJR (2001), we estimate statistically significant effects of malaria prevalence once we control for its potential endogeneity, and we do not find empirical evidence that the specific instrumental variable used by Sachs (2003) is invalid, as presumed by RST (2004). For given effects of institutional quality, our results indicate quantitatively important direct negative effects of malaria prevalence on economic performance. This result appears to be robust to using alternative measures of institutions and malaria prevalence, alternative and additional IVs, and an alternative sample of countries.

Taken at face value, our results imply that institutions do not trump all other potential determinants of development. Emphasizing 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 also direct assistance from foreign donors for interventions against diseases may indeed be necessary 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 our results point out that good institutions and a favorable disease ecology appear to be necessary but not necessarily sufficient recipes for economic success.

References

Acemoglu, Daron, Simon Johnson, James A. Robinson (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. Quarterly Journal of Economics 117: 1231-1294.

Acemoglu, Daron, Simon Johnson, James A. Robinson (AJR) (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. American Economic Review 91 (5): 1369-1401.

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

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

Andrews, Donald W. K. (1989). Power in Econometric Applications. Econometrica 57: 1059-1090.

Andrews, Donald W. K. (1999). Consistent Moment Selection Procedures for Generalized Method of Moments Estimation. Econometrica 67: 543-564.

Arrow, Kenneth J. (2004). New Antimalarial Drugs: Biology and Economics Meet. Finance and Development, March: 20-21.

Bloom, David E., Jeffrey D. Sachs (1998). Geography, Demography, and Economic Growth in Africa. Brookings Papers on Economic Activity (2): 207-295.

Cragg, John G., Stephen G. Donald (1993). Testing Identifiability and Specification in Instrumental Variable Models. Econometric Theory 9: 222-240.

Diamond, Jared (1997). Guns, Germs, and Steel. New York: W. W. Norton.

Donald, Stephen G., Whitney K. Newey (2001). Choosing the Number of Instruments. Econometrica 69: 1161-1191.

Dufour, Jean-Marie (1997). Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models. Econometrica 65: 1365-1387.

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

Engerman, Stanley L., 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 CA: Stanford University Press, 260-304.

Frankel, Jeffrey A., David Romer (1999). Does Trade Cause Growth? American Economic Review 89 (3): 379-399.

Gallup, John Luke, Jeffrey D. Sachs (2001). The Economic Burden of Malaria. Supplement to: The American Journal of Tropical Medicine & Hygiene 64 (1/2).

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

Hahn, Jinyong, Jerry Hausman (2002). A New Specification Test for the Validity of Instrumental Variables. Econometrica 70:163-189.

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

Hausman, Jerry, James H. Stock, Motohiro Yogo (2004). Asymptotic Properties of the Hahn-Hausman Test for Weak Instruments. mimeo.

Kamarck, Andrew M. (1976). The Tropics and Economic Development. A Provocative Inquiry into the Poverty of Nations. Baltimore: The Johns Hopkins University Press.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2003). Governance Matters III. Governance Indicators for 1996-2002. The World Bank.

Kiszewski, Anthony, Andrew Mellinger, Pia Malaney, Andrew Spielman, Sonia Ehrlich Sachs, 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): 486-498.

Landes, David (1998). The Wealth and Poverty of Nations. New York: W. W. Norton.

Lee, Douglas H. K. (1957). Climate and Economic Development in the Tropics. New York: Harper and Collins for the Council on Foreign Relations.

Masters, William A., Margaret S. McMillan (2001). Climate and Scale in Economic Growth. Journal of Economic Growth 6 (3): 167-186.

McArthur, John W., Jeffrey D. Sachs (2001). Institutions and Geography: A Comment on Acemoglu, Johnson, and Robinson (2000). NBER Working Paper, 8114, February.

Moreira, Marcelo J. (2003). A Conditional Likelihood Ratio Test for Structural Models. Econometrica 71: 1027-1048.

Olsson, Ola and Douglas A. Hibbs Jr. (2005). Biogeography and Long-run Economic Development. European Economic Review 49: 909-938.

Parker, Philip M. (1997). National Cultures of the World. A Statistical Reference. Westport, Connecticut: Greenwood Press.

Rodrik, Dani, Arvind Subramanian, Francesco Trebbi (RST) (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. Journal of Economic Growth 9: 131-165.

Sachs, Jeffrey D. (2001). Tropical Underdevelopment. NBER Working Paper, 8119, February.

Sachs, Jeffrey D. (2003). Institutions Don't Rule: Direct Effects of Geography on Per Capita Income. NBER Working Paper, 9490, February.

Sachs, Jeffrey D., Pia Malaney (2002). The Economic and Social Burden of Malaria. Nature 415 (6872): 680-685.

Shea, John (1997). Instrument Relevance in Multivariate Linear Models: A Simple Measure. Review of Economics and Statistics 79: 348-352.

Staiger, Douglas, James H. Stock (1997). Instrumental Variables with Weak Instruments. Econometrica 65: 557-586.

Stock, James H., Motohiro Yogo (2004) Testing for Weak Instruments in Linear IV Regressions. mimeo.

Summers, Robert, Alan Heston (1991). The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. Quarterly Journal of Economics, 106: 327-368.

World Bank (2002). World Development Indicators. CD-ROM.

World Health Organisation (WHO) (2005). World Malaria Report 2005. Geneva: WHO and UNICEF, http://rbm.who.int/wmr2005/index.html

Appendix: Definitions and Sources of Variables

coast
Proportion of land area within 100 km of the sea cost
Source: Gallup, Sachs, Mellinger (1999), here taken from McArthur and Sachs (2001)

engfrac
Proportion of the population speaking English
Source: Hall and Jones (1999)

eurfrac
Proportion of the population speaking one of the major languages of Western Europe: English, French, German,Portugese, or Spanish
Source: Hall and Jones (1999)

exprop
Index of protection against expropriation in 1985-1995; limited to 64 countries but includes Bahamas and Vietnam, which are not included in *socinf*; measured on a [1,10] scale.
Source: Acemoglu et al AER (2001), p. 1398.

frost
Proportion of a country's land receiving five 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 [0,1] scale.
Source: Masters and McMillan (2001).

gadp
Index of government anti-diversion 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 [0, 1] scale
Source: Hall and Jones (1999)

humid
Highest temperature during the month when average afternoon humidity is at its highest; measured in degrees Celsius.
Source: Parker (1997).

latitude
Distance from the equator as measured by the absolute value of country-specific latitude in degrees.
Source: Hall and Jones (1999).

lngdpc
Real GDP per capita, adjusted for purchasing power parity (PPP), 1995; measured in international dollars.
Source: World Bank, Development Indicators CD-ROM, 2002.

lnmort
Settler mortality rates in colonies in the early 19th century, fourth mortality estimate (72 countries, excluding France and UK); measured as death rate among 1,000 settlers where each death is replaced with a new settler.
Source: Acemoglu, Johnson, Robinson (2001, p. 1398), and Acemoglu, Johnson, Robinson (2000).

lnmort2
Revised estimate of *lnmort*
Source: Albouy (2004)

maleco
Combines climatic factors and specific 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 sub-national 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 27 October 2003.
Source: Kiszewski and Sachs et al. (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 [0,1] scale. Revised version, dataset as of 27 October 2003.
Source: Sachs (2003), here taken from http://www.earth.columbia.edu/about/director/malaria/index.html#datasets.

malrisk
Proportion of each country's population that live with risk of malaria transmission, involving three largely non-fatal species of the malaria pathogen (Plasmodium vivax, P. malariae, P. ovale); measured on a [0,1] scale; dataset as of 27 October 2003.
Source: 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.
Source: Kaufmann et al. (2003)

trade
Natural log of the Frankel-Romer predicted trade share, based on measures of population size and geography.
Source: Hall and Jones (1999)

Figure 1 — Alternative Links Between Geography, Institutions, and Development

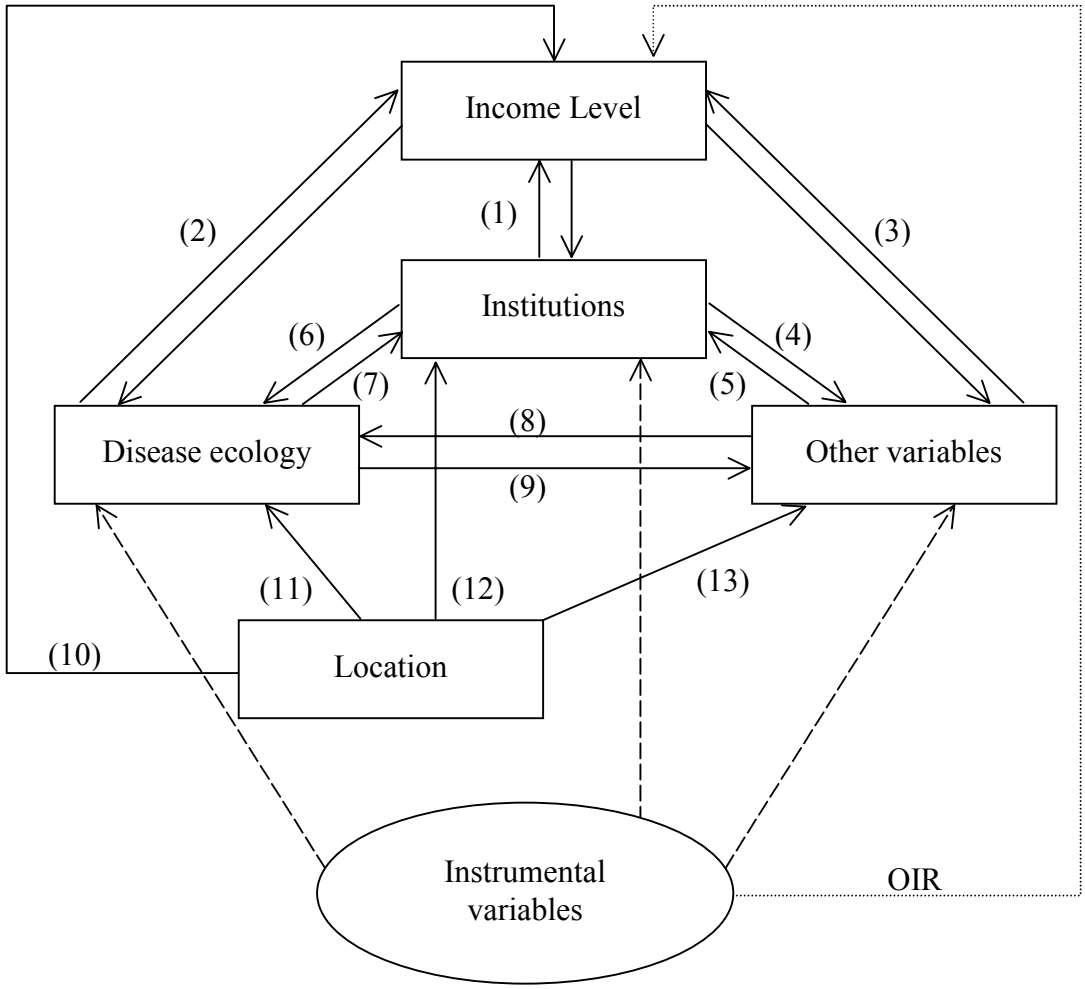

Table 1: Baseline Estimation Results

	(1)		(2)		(3)		(4)	
Explanatory variables	rule	malfal	rule	malrisk	gadp	malfal	exprop	malfal
Estimated coefficients (s.e. below)								
2SLS	0.89	-1.04	0.78	-1.31	3.31	-1.48	0.55	-1.03
	0.18	*0.30*	*0.22*	*0.42*	*0.60*	*0.22*	*0.12*	*0.34*
Fuller	0.89	-1.05	0.78	-1.29	3.31	-1.48	0.53	-1.07
	0.17	*0.30*	*0.22*	*0.40*	*0.59*	*0.21*	*0.12*	*0.33*
Bounds of 95% 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
CLR 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
Test of $\beta_2=\beta_3=0$								
CLR statistic	21.50		21.50		21.88		21.50	
(5% critical value)	2.97		2.97		2.97		2.97	
No of observations	45		45		45		45	
No 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^2	0.57	0.67	0.57	0.49	0.53	0.67	0.45	0.67
Shea partial R^2	0.37	0.43	0.33	0.28	0.37	0.46	0.23	0.34
Weak instrument test								
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 *lnmort* and *maleco* for the small sample of 45 countries. 2SLS denotes the two-step least squares estimator and Fuller denotes the Fuller estimator with correction parameter c=1 proposed by Hausman et al. (2004). The CLR confidence interval and the test for $\beta_2=\beta_3=0$ are based on the conditional likelihood ratio test proposed by Moreira (2003). The Cragg-Donald statistic is introduced by Cragg and Donald (1993) and used by Stock and Yogo (2004) for weak instrument tests. The critical value is interpreted as follows. If the Cragg-Donald statistic does not exceed the critical value, then a standard significance test with nominal size of 5% has maximal size of 10%.

Table 2: The Impact of Additional Instrumental Variables

	(1)		(2)		(3)		(4)	
Instruments	baseline + climate		baseline + Europe		baseline + openness		all instruments	
Explanatory variables	rule	malfal	rule	malfal	rule	malfal	rule	malfal
Estimated coefficients (s.e. below)								
2SLS	0.97	-0.90	0.86	-1.14	0.81	-1.09	0.84	-1.08
	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
	0.16	*0.29*	*0.16*	*0.27*	*0.16*	*0.29*	*0.14*	*0.26*
Bounds of 95% 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
CLR 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
No of observations	44		44		44		44	
No 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								
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
Instrument selection criteria								
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 instrument sets are baseline (*lnmort*, *maleco*), climate (*frost*, *humid*, *latitude*), Europe (*eurfrac*, *engfrac*), and openness (*coast*, *trade*). The weak instrument test of 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 left-hand side variable and the endogenous right-hand side variable are interchanged. The Donald-Newey instrument selection criterion suggested by 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 proposed by Andrews (1999) for the choice of instruments. See also the note to Table 1.

Table 3: The Validity of *Lnmort* and *Maleco*

	(1)		(2)		(3)		(4)	
Instruments	lnmort2 + maleco		lnmort2 + maleco + Europe		lnmort + Europe		lnmort2 + Europe	
Explanatory variables	rule	malfal	rule	malfal	rule	malfal	rule	malfal
Estimated coefficients (s.e. below)								
2SLS	1.12	-0.85	0.89	-1.15	0.75	-1.35	0.77	-1.39
	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
	0.29	*0.38*	*0.19*	*0.28*	*0.20*	*0.35*	*0.21*	*0.33*
Bounds of 95% 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
CLR 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
No of observations	45		44		44		44	
No 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
Weak instrument test								
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								
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	

Note: See Tables 1 and 2

Appendix Table: Results for a Larger Sample of Countries

	(1)		(2)		(3)		(4)	
Instruments	lnmort + maleco		lnmort + maleco		all instruments		lnmort2 + Europe	
Explanatory variables	rule	malfal	exprop	malfal	rule	malfal	rule	malfal
Estimated coefficients (s.e. below)								
2SLS	0.92	-0.95	0.61	-0.81	0.78	-1.09	0.77	-1.26
	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
	0.24	*0.30*	*0.17*	*0.36*	*0.15*	*0.24*	*0.23*	*0.31*
Bounds of 95% 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
CLR 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
No of observations	64		62		61		61	
No 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
Shea partial R^2	0.21	0.35	0.10	0.25	0.56	0.61	0.24	0.35
Weak instrument test								
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

Note: See Tables 1 and 2.