THE AUSTRALIAN NATIONAL UNIVERSITY
WORKING PAPERS IN ECONOMICS AND ECONOMETRICS

DISEASES DOMINATE*

Aaron Batten and Alan Martina

School of Economics
Australian National University

April 2006

Working Paper No. 470

ISBN 086831 470 6

* We thank the following for comments made on earlier versions of this paper: Steve Dowrick, and participants at seminars given in the Departments of Economics at The Australian National University and Oxford University respectively, The London School of Hygiene and Tropical Medicine and the symposium on Institutions and Economic Development, University of Otago. Considerable assistance was provided with James Stock, Motohiro Yogo and Heather Anderson with the provision of the empirical information set out in Appendix B. We remain, however, responsible for the contents of this paper.

#This paper is to appear in William R. Garside (ed.), Institutions and Market Economics. London: Palgrave, 2007

Diseases Dominate

I Introduction

What underlying, or deep factors determine the distribution of social well-being across the nations of the world?

Variations of this question have received the attention of many over the centuries. Clearly Adam Smith, in *An Inquiry into the Nature of the Causes of the Wealth of Nations* (1776), addressed this question. In far more recent times economic historians and others attempted to provide, implicitly at least, cogent answers to this sort of question – authors such as Rostow (1960), Reynolds (1983), North and Thomas (1973), North (1990, 1994, 1995), Diamond (1997) and Bernstein (2004). The most recent group of contributors to the search for answers to the sort of question raised has been economists (in particular, Hall and. Jones (1999), Acemoglu, Johnson and Robinson (2001, 2002, 2006), Easterly and Levine (2003), Sachs (2003) and Rodrik, Subramanian and Trebbi (2004)).

This amount of attention given recently to attempting to answer versions of the question asked is not surprising given the contemporary world as it is. For instance, if a broad measure of social welfare is adopted, such as the Human Development Index (HDI), then in 1994 the measure for the country with the highest level was some 5.6 times greater than that for the country with the smallest measure. (Canada had the highest level of 0.9801 and Sierra Leone the smallest with a measure of 0.176.)

In our attempt to answer the question posed we re-examine the empirical results, based on cross-country data sets, presented in Rodrik *et al*. (2004) and Acemoglu *et al*. (2001) respectively. Both sets of authors conclude that the variation across countries of a proxy variable that measures the quality of relevant economic institutions - that are solely concerned with ensuring the rule of law and the protection of individual property rights - is the dominant deep determinant that explains the variation in per capita gross domestic product (GDP) across countries in the mid-1990s. (The other deep determinants of economic development taken into account in Rodrik *et al*. (2004) were a measure of the level of integration of a country into international trade and a measure of the geography for this country.)

We depart from their general approach in two ways. First, we like others (for instance Sen (1985, 1999), Dasgupta (1993) and Ranis (2004)) believe that a measure of per capita GDP for any community inadequately captures the range of relevant considerations that may influence the level of social well-being in this community – and presumably it is the variation in the level of social well-being across countries that most needs to be explained. For this reason we concentrate our attention on attempting to explain the variation across countries of a measure of the HDI in 1994.[1] While this measure of social well-being has been criticized on a number of grounds[2] in our view it still is to be preferred to the use of per capita GDP. The second way that we depart from previous approaches is that we include one additional possible deep determinant of economic development besides the three considered by Rodrik *et al*. (2004). This additional deep determinant is the percentage of the population at risk of contracting malaria (hereon called malaria risk). This measure

[1] We adopt the year 1994 for consideration in order to avoid the influence that the pandemic of HIV/AIDS in sub-Saharan Africa is likely to have on the measure of the HDI after the mid-1990s.

[2] See, for example, Dasgupta (1993), Ravallion (1997) and Noorbakhsh (1998).

should be thought of as a proxy measure for the level of virulent diseases associated with malaria that are present in a community. (As pointed out in Section III.2.2, the reduction in the prevalence of malaria in a community results in a fall in the child mortality rate from all causes of death.)

We find that malaria risk is the dominant deep determinant (of the four determinants considered) that explains most of the variation across countries of the HDI in the mid-1990s. In particular we find that on applying the standard tests of statistical significance (i) the regressor representing economic institutions (*viz*. the rule of law) sometimes is not statistically significant whereas malaria risk is always statistically significant in all the estimated regression equations presented later in which both malaria risk and the rule of law (along with other relevant variables) are included as regressors in relevant regression equations. We also find that (ii) the coefficient estimates for relevant regression equations imply that a reduction in the level of malaria risk by one standard deviation (SD) always has a greater influence on increasing the level of the HDI than does a one SD improvement in the measure of the rule of law. Exactly how much greater seems to be somewhere within the range of 1.25 to 5.68 times greater. (iii) Within this range the most reliable values probably are towards the lower bound. This is indicated by an argument developed later (in Section IV. 4.1) that sets out a number of reasons why ordinary least squares estimates of relevant regression equations may be more reliable than those generated by using two-stage least squares estimation procedures. The ordinary least squares estimates, in turn, imply values of between 1.25 and 1.60 within the range of values indicated at the end of point (ii). (iv) These empirical results only apply to the average country in the late 20th Century.

When the full data set is split into two sub-sets – one made up of countries in which malaria risk is present and the other made up of countries in which this risk is absent - we find that the regressor economic institutions is not statistically significant for countries in which malaria risk is present. Naturally malaria risk is a highly statistically significant regressor for these countries. In contrast, in those countries in which there is no malaria risk the regressor economic institutions is highly statistically significant.

While these empirical results emphasize the importance of malaria and related diseases for explaining the variation in the level of well-being across countries, Rodrik *et al*. (2004: 151), in contrast, asserts that because malaria is a debilitating rather than a fatal disease it is unlikely to have a marked influence on the level of per capita GDP. The relevant medical literature suggests otherwise. For instance in 1994 some 46 per cent of the world's population was exposed to the risk of contracting malaria (Hay *et al*., 2004: 328). In addition, in 2001 in Africa alone near to three million people died from malaria and, in that year, this disease was responsible possibly for two thirds of the mortality rate from all causes for children under five years of age living in mildly to heavily malaria-infected regions on that continent.[3] Should a person survive a severe attack of this disease what, as

[3] See Breman *et al*. (2004: 3 and 5) and Snow *et al*. (2004). These estimates are much larger than those provided by the World Health Organization (WHO) and others. This is because WHO and others fail to take into account that an increase in the prevalence of malaria also increases the severity of a range of other childhood diseases. This matter is discussed in more detail later in the Section III.2.2 in the main text.

emphasized by Bremen (2001: 1), is '[p]articularly devastating to development are the long-duration effects of malaria, manifest in learning and behavioral disorders and decreased capacity to work.' For a country in which these long-term consequences are prevalent amongst the population, they are likely to have profound detrimental economic consequences.[4] (These matters will be explored a little further in Section III.2.2.)

For these reasons we take seriously a result presented by Sachs (2003); *viz*. that the risk of contracting malaria is a statistically significant regressor in explaining the variation in the level of per capita GDP across countries. Indeed, motivated by this result we reconsider this matter here in some detail by taking into account a range of issues not considered by Sachs (2003).[5] (i) To begin with we allow for a wider range of regressors in the estimation of relevant regression equations. (ii) Whereas in Sachs (2003) the dependent variable was per capita GDP we replace this dependent variable by considering a measure of social well-being (the HDI). (iii) We question the use by Sachs (and others such as Acemoglu *et al*. (2001) and Rodrik *et al*. (2004)) of instrumental variables within a two-stage least squares estimation procedure to estimate relevant

---

[4] See Breman (2001), Malaney and Sachs (2002), Breman *et al*. (2004), Malaney, Spielman and Sachs (2004) for surveys of the relevant medical and economic literature on this point.

[5] Sachs, along with a number of co-authors has published a range of stimulating papers that bear on the issue of the burden of disease on economic development (Bloom and Sachs (1998), Gallup, Mellinger and Sachs (1998), Gallup and Sachs (2001), McArthur and Sachs (2001), Malaney and Sachs (2002) and Malaney, Spielman and Sachs (2004)).

coefficients and suggest, instead, that coefficient estimates derived by the application of ordinary least squares may be more reliable. (iv) For the sake of completeness, however, we do generate some empirical results by applying instrumental variables within a two-stage least squares estimation procedure while also indicating how certain useful instrumental variables can be exploited by drawing upon the theory, developed by Stock and Yogo (2002), for identifying weak instrumental variables. (v) Making use of the full data set only provides empirical results for the average country in the mid-1990s. To determine the influence of malaria risk and economic institutions on sub-groups of countries we split the full data set into those countries that are faced with malaria risk and into those countries that are not.

Some may not be convinced by the econometric evidence to be presented later. We point out, however, that a brief perusal of the history of epidemics or pandemics of virulent infectious diseases (VIDs) suggests that these diseases have had a profound influence on how various societies have evolved over time.[6] One only needs to think of the pandemic of the Black Death that spread through Europe from 1347 through to 1350 and the waves of pandemics of various VIDs that struck the civilizations of the Americas through the 16th century. The Black Death caused mortality rates of 30 to 40 per cent (Hays, 2003: 37) while the total population of the Americas was reduced from between 50 to 100 million to below 10 million by the middle of the 17th century. Much of this reduction was due to pandemics of VIDs (Hays, 2003: 72 - 77) accidentally introduced from Europe.

[6] See, for instance, Cartwright (1972), Kiple (1993), Porter (1997) and Hays (2003).

The long-term consequences of these dramatic demographic changes for the societies concerned were profound.

This theme is taken up in the next section where we consider a case history from the late 19th Century. During this episode History conspired, accidentally, with Nature to create a near perfect controlled experiment to demonstrate the influence that an epidemic of VIDs alone can have on the economy of the community concerned. During this quasi-controlled experiment the level of economic institutions (from hereon institutions) were of a comparatively high level.

An inference that can be drawn from this quasi-controlled experiment is that if, in the mid-1990s, some set of VIDs were widespread and their virulence varied markedly in intensity across the regions of the globe then the presence of these VIDs must explain some proportion of the variation in the level of social well-being across countries. As indicated above (and as explained in more detail in Section III.2.2) malaria and its associated diseases in the mid-1990s was just such a set of diseases.

Naturally, this particular line of argument does not allow us to quantify how significantly important malaria and associated diseases were relative to other factors (such as institutions) in explaining the variation in the level of social well-being across countries in the late 20th century. To do so requires the application of econometric analysis.

The argument developed here is not meant to imply that we regard institutions as being of no importance in explaining the variation in the level of social well-being across countries. If possible, fatal contagious diseases need to be brought under control through the provision of the institution of Public Health. Once this has been done then, and most likely only then, does the provision of a reasonable level of the rule of law and the protection of individual property rights come into their own in facilitating further economic development. What is more, just containing VIDs, without subsequently also providing a reasonable level of the rule of law and the protection of individual property rights probably will do relatively little to facilitate further expansion of the level of the HDI for the country concerned. As indicated by the analysis developed below and culminating in Section VI there seems to be a particular sequence of steps that need to be taken, or a particular path that has to be traversed by a community before it can experience a sustained rise in the level of its HDI. First VIDs have to be contained, followed by the provision of the rule of law and the protection of individual property rights.[7]

While we recognize the particular role that economic institutions have to play in the attempt to expand the level of social well-being for a poor community, we cannot agree with the follow bold statement made by Bernstein (2004: 293):

> Institutions, not the bounty of nature or freedom from imperialist domination, separate the winners from the losers in the global economy. First and foremost, it is the degree of respect and reverence for the Rules of the Game – rule of law,

[7] The conjecture in the main text is similar to an argument presented in Ranis, Stewart and Ramirez (2000). They, however, were only concerned with the matter of economic growth and not with the variance of well-being across countries. Also see Ranis (2004) and Ranis and Stewart (2001).

> equality under the law, and respect for the civil liberties of all – that dominates the wealth of nations.

This emphasis on the over-arching importance of the institution of the Rules of the Game is not supported by the evidence to be presented here. Rather the relevant issues are rather more complicated than Bernstein suggests. The presence or absence of widespread virulent contagious disease in the various communities of the world has to be part of any cogent story that attempts to explain the separation of 'the winners from the losers in the global economy'.[8]

The plan of the paper is as follows. In the next section the History-plus-Nature case study is considered. The basic regression model and data used to estimate this model are both discussed in Section III. The core of the regression analysis is presented in Section IV. In Section V the data set used to estimate the regression model is extended to determine how robust the initial regression results are. In Section VI we split the data set into those countries that, in the mid-1990s were faced with malaria risk and those that were not. Concluding comments are made in Section VII.

---

[8] In the index for Berstein's book, *The Birth of Plenty* (2004), none of the following words appear: infectious disease, disease, malaria – or mention of any of the other major fatal infectious diseases that afflict humankind.

II Quasi-Controlled Experiment Generated by History-Plus-Nature

II.1 The Panama Canal case

The controlled experiment takes the following form: In the initial situation in the community concerned there is a comparatively high level of the rule of law and the protection of individual property rights (henceforth Rule) is in place. A VID is present in the environment of this community, however. Those who invest in the economy for this community are not aware of this fact or disregard it. The question is: Will this investment prove to be privately profitable? Or more generally, will it be possible, at reasonable cost, to develop the economy for this community? To answer this question we consider the attempt to build the Panama Canal between 1881 and 1914.

After Ferdinand de Lesseps created his company to raise financial capital to fund the building of the Panama Canal, construction began in 1881.[9] Clearly investors in this project perceived that the quality of the rule of law and the protection of individual property rights were adequate otherwise presumably they would not have invested in this project. Within a short time, however, deaths caused by either yellow fever or malaria began to occur amongst the workforce. By the end of 1888 the level of mortality had risen to an estimated thirty thousand. (At some locations along the building corridor for the canal it is likely that two out of every three workers died from these diseases.) At this point this civil engineering project was abandoned by the French entrepreneurs.

[9] The discussion in the main text draws on Hays (2003), McCollough (1997), Malaney and Sachs (2001), Porter (1997) and Spielman and D'Antonio (2001).

While there were engineering and geological difficulties that also needed to be overcome for this project to be completed, it seems reasonable to infer from these facts on mortality that the main constraint on this investment project becoming privately profitable was that created by the prevalent fatal contagious diseases. That the institutions of the rule of law and individual property rights were in place did not alter this situation. This argument was understood by senior members of the United States Federal government when it took over the task of building the canal in 1904.

To explain, in response to an outbreak of yellow fever that had occurred along the canal corridor in the first part of 1905, John Stevens, on being appointed the chief engineer for the canal project in that year, immediately called for building work on the canal to cease. Stevens then allocated all available resources to William Crawford Gorgas - who had gained considerable experience in controlling the breeding of mosquitoes (in particular *Aëdes aegypti*) in Cuba between 1900 and 1904 in the attempt to contain yellow fever[10] – to allow him to mount a vigorous campaign aimed at drastically reducing mosquito population along the canal corridor. Above all Stevens and Gorgas realized that until this task had been

[10] This mosquito-control program arose out the work by Carlos Juan Finlay, Jesse Lazaer, Aristedes Agramonte, James Carrol and Walter Reed. They all contributed in various ways to demonstrating in 1900 that *Aëdes aegypti* was the vector carrying yellow fever from person to person. Whereas Finlay in 1872 published his first paper on the possible link between the mosquito and yellow fever, others had hypothesized this link earlier - for instance Dr Josiah Clark Nott in 1848 (McCollough, 1977: 142 – 3).

performed successfully it would not be economically efficient to attempt to construct the canal. This decision received the support of President Theodore Roosevelt on the advice, in particular, of his friend Dr William David Lambert. In this advice Lambert stated the following:'…If you [President Roosevelt] fall back upon the old methods … you will fail, just as the French failed. If you back up Gorgas and his ideas and let him pursue his campaign against the mosquito, you will get your canal.' (McCollough: 77 and 467 – 8)) Rooseveldt heeded this advice and gave his support to Stevens and Gorgas. One consequence was that only two per cent of the workforce used in the building of the canal was hospitalized in the period 1904 and 1914 (when the canal was completed) compared to about a 33 per cent mortality rate during the period 1881 to 1888.

II.2 Inference and interpretation

This well-documented case history – and other case histories not discussed here[11] - indicate that controlling VIDs is a necessary condition that has to be satisfied before sustained economic development is possible in any community. In particular this case history indicates that having in place an adequate level of the rule of law is not the only institution

---

[11] Some other case histories that demonstrate the importance of controlling malaria before sustained economic development can take place, in an environment where an adequate level of economic institutions are in place, are: Batavia in the mid-18th century (van der Brug, 1997), Malaya from1901 through to the mid-1920s (Watson, 1943) and Northern Rhodesia in the 1930s (Watson, 1953). Also see Keiser *et al*. (2005), Utzinger *et al*. (2001, 2002) and Gallup and Sachs (2001: 90 – 91).

that a society needs to have in place to ensure that its economy functions reasonably effectively. It also is crucially important to have in place institutions that facilitate the provision of an adequate level of public health. In the present study, however, the reason for not allowing for Public Health in the econometric analysis is that the variable we attempted to use to represent Public Health - *viz*. the level of the effectiveness of government to provide public services – is so highly correlated with the rule of law that it is not possible to distinguish between these two measures of the quality of quite different types of institutions. (The simple correlation coefficient between the rule of law and government effectiveness is 0.9098.[12])

As noted in the introduction, this quasi-controlled experiment allows the inference to be drawn that if the prevalence of some VIDs are widespread and varies markedly in intensity across the regions of the globe then these diseases must explain a significant proportion of the variation in the level of social well-being across countries. What makes this inference pertinent here, and as noted in Section I, is that in the contemporary world such a set of diseases exists - namely malaria and its associated diseases.

III. Model to be Estimated and its Variables

In the previous section advantage was taken of quasi-controlled experimental evidence to draw the inference (bearing in mind relevant information presented later in Section III.2.2) that malaria and associated diseases must be of significant importance in explaining the variation in the level of social well-being across countries in the mid-1990s. Attention now

[12] The relevant data is taken from Kaufmann *et al*. (2002).

turns to attempting to confirm this inference and to quantifying the relative importance of malaria and associated diseases in explaining the variation in the level of social well-being across countries.

III.1 The basic model

To guide the selection of the data we draw on and extend the model sketched out by Rodrik, Subramanian and Trebbi (RST) (2004). This model is indicated in Figure 1. In this diagram it is indicated that a range of variables are assumed to interact with each other and with the level of social well-being for the representative country. The variables RST (2004) employed were GDP per capita, geography, institutions and the level of integration of the economy into international trade for the country concerned. We have added to this list the disease burden that the representative country faces. We also have replaced GDP per capita with the HDI – a variable that is used to represent the level of social well-being.

It can be seen in Figure 1 that some of the arrows have been numbered with these numbered arrows indicating the lines of causation. The lines of causation we are particularly interested in this and the next section are those associated with the arrows marked (1), (2), (3) and (4). In Section III.3 we also will consider the line causation associated with the arrow marked (8) (the influence of the disease burden on the level of institutions) and the arrow marked (5) (the influence of the level of institutions on the disease burden) while also allowing for other considerations (such as the influence of Geography on Institutions).

The main technical difficulty associated with estimating at least part of the system represented in Figure 1 is the endogeneity or reverse causation between certain sets of variables. (The only variable for which this issue does not arise, and as indicated in Figure 1, is Geography which is treated as an exogenous variable.) One method for handling this problem is the use of suitable Instrumental Variables (IVs) in a Two-Stage Least Squares (2SLS) estimation procedure. While we apply this estimation technique later, there are some practical difficulties associated with applying this technique within the present context. This matter is returned to in Section IV.4.1.

Figure 1 goes about here.

The regression equation system associated with the arrows marked (1), (2), (3) and (4) in Figure 1 has as its basic regression equation the following:

$$HDI_i = \beta_o + \beta_1 INS_i + \beta_2 INT_i + \beta_3 GEO_i + \beta_4 DIS_i + \varepsilon_i , \qquad (1)$$

where HDI_i denotes the Human Development Index for 1994 for country i, INS_i, INT_i, GEO_i and DIS_i are the chosen measures, yet to be discussed, for economic institutions, integration, geography and disease respectively, and ε_i is the random error term for this regression equation.

The first stage regression equations (in the 2-stage estimation process) are assumed to be of the following form:

$$INS_i = \alpha_o + \alpha_1 \text{ logem4}_i + \alpha_2 \text{ logfrankrom}_i + \alpha_3 \text{ me}_i + \alpha_4 \text{ Pop100km}_i + \varepsilon_{iINS}; \qquad (2)$$

$$INT_i = \gamma_0 + \gamma_1 \text{ logfrankrom}_i + \gamma_2 \text{logem4}_i + \gamma_3 \text{ me}_i + \gamma_4 \text{ Pop100km}_i + \varepsilon_{iINT}; \qquad (3)$$

$$DIS_i = \kappa_o + \kappa_1 \text{ me}_i + \kappa_2 \text{ logem4}_i + \kappa_3 \text{ logfrankrom}_i + \kappa_4 \text{ Pop100km}_i + \varepsilon_{iDIS}. \qquad (4).$$

The regressors in these last three regression equations are IVs that are listed and specified in Table 1. These IVs are discussed later. The term ε_{INS} etc refer to error terms in first stage regression equations.

The estimates of these first-stage regression equations provide the basis for generating the predicted values of INS, INT and DIS respectively that are used as instruments to estimate regression equations (1).

Table 1 goes about here.

III.2 The regressors

An interpretation of one aspect of what is attempted in this section is that we decompose the measure of geography (Geo) used in RST (2004: 162). There this measure was defined as the distance of the capital of the country concerned from the Equator (measured as the absolute latitude divided by 90). In our view this is too all-encompassing a measure. There are many dimensions of the geography of a country and these different dimensions capture quite different aspects of the given economic and climatic environment in which the country operates.[13] For this reason we essentially consider two variables that measure two quite different aspects of the geography of a country.

III.2.1 Distance from water transport

This variable draws on the observation that a nation's degree of access to navigable waterways has been shown to play an important role in its development since it will influence the extent to which a region can specialize in the production of goods in which it has a

[13] Sachs (2003) equally is critical of the use of distance from the equator as a measure of the geography of a country.

comparative advantage (Gallup *et al.* (1998, 1999) and Radelet and Sachs (1998)). For this reason we use the proportion of a nation's population within 100km of the coast or navigable water ways (Pop100km).

The other variable we employ to reflect another aspect of a country's geography is the risk of contracting malaria – a variable whose level for a country is determined to some extent by this country's location relative to the equator. (The risk of contracting malaria in a region also is influenced by the altitude of this region, the amount of rainfall and types of mosquitoes and malaria parasites to be found in this region.) Both this variable and the other geography variable (hereon Pop100km) were employed in Sachs (2003). He applied these variables separately, however, in different regression equations. Later we employ both variables in a single regression equation to estimate regression equation (1).

III.2.2 Malaria

On consulting estimates for the 1990s and beyond of the mortality rate and disease burden (measured by disability adjusted life years) imposed by malaria, relative to that for a range of other contagious diseases (to be found in. Murray and Lopez (1997a, Table 3; 1997b, Table 3) and WHO (2002)), the impression may be gained that malaria imposes a relatively small disease burden. In 1990 malaria ranked only sixth of all contagious diseases, with lower respiratory infections and diarrheal diseases appearing to have imposed far greater burdens on the world community. However, as Snow *et al*. (2004) point out, these estimates are based on the assumption that an individual death or morbidity event is caused by a single disease – an assumption at variance with the empirical facts. In contrast, empirical evidence presented by Snow *et al*. (2004) indicates that in sub-Saharan Africa malaria contributes in indirect ways towards the death of children, under the age of five, from a range of diseases. The upper limit of their estimates, based on statistical inference, suggests that on allowing for these indirect mechanisms about two-thirds of all deaths for this age group are caused by malaria.[14] Snow *et al*. (2004) qualify their estimates since they emphasize that these estimates have not been confirmed by satisfactory controlled experimental evidence – experimental evidence that would allow for all confounding factors and where, during the experiment, malaria is eliminated completely from the relevant population. Nevertheless their estimates are suggestive. [15,16]

---

[14] This predicted change is for a group of countries for which the average child mortality rate (CMR) in 1990 was 181 per 1000 live births. For the same year the CMR was 11 and below for the wealthy industrialized – and nearly all non-tropical - countries. See United Nations Children's Fund (2001: 3).

[15] Also see Molineaux (1997), Bremen (2001: 7) and Bremen *et al*. (2004: 2 – 3, 3 - 6) on this issue.

[16] Kremer and Miguel (2004:159) also illustrate the general point made in the main text; viz. that an externality is created by altering the prevalence of a particular parasite or vector in the environment. In their study they find that de-worming children in several villages across Kenya resulted in improved health outcomes more than twice as large as that estimated when considering only the direct health impacts of the de-worming program.

Two inferences follow from these observations. One is that clearly actual mortality and morbidity rates for malaria (where malaria is treated as the single cause of death or morbidity) would understate (and possibly by a large amount) the true burden imposed by this disease. For this reason the indicator used here for the disease burden imposed by malaria in a given country is the risk of contracting malaria in this country. The other inference is that since malaria interacts with various other diseases the risk of contracting malaria could be viewed as a proxy variable for the risk of contracting a range of infectious diseases besides just malaria.

A further reason for treating this variable as a proxy variable is that a range of less virulent tropical diseases (such as sleeping sickness (trypanosomiasis), hookworm (ancylostomiasis), shistosomiasis (bilharzia) and yellow fever) are to be found in geographic regions of the world in which the risk of contracting malaria also is comparatively high. (This point can be established by consulting the world map of the distribution of incidence of malaria to be found in Hay *et al*. (2004) and comparing it with the various relevant world maps (to be found in Kiple (1993)) of the geographic distribution of the incidence of various relevant tropical diseases.)

As for the various channels along which malaria (and its associated diseases) is thought to influence the level of the HDI, these are indicated in Figure 2. In this regard is noted that Sachs and others have argued that the prevalence of malaria in a country can have a significant direct, as distinct from an indirect, impact on the level of development of this country. (See the references cited in footnote 5.) For example, even when workers who are ill do attend their place of employment they are likely to function well below their normal levels of productivity due to reduced energy levels and an impaired ability to concentrate. In addition, a life plagued with sickness almost certainly will reduce an individual's life expectancy which, in turn, is likely to reduce the value of the human capital invested in this person. Consequently this will tend to reduce the incentive for employers to invest in the education of the workforce.

Figure 2 goes about here.

As for younger individuals who are sick regularly, it will be difficult for them to find the energy needed to attend school. In a study by Foster and Leighton (1993: 11) of the economic impacts of malaria in Kenya and Nigeria estimated that school children in Kenya miss, on average, 11 per cent of their schooling due to malaria and malaria related illness. (In this study malaria also is estimated to deprive the economy of anywhere up to 14 per cent of its workforce due to worker absenteeism and reduced productivity per year.) In addition, the comparatively high child mortality rate caused by malaria will reduce the incentive for parents to invest in their children's education.

The risk of contracting malaria also will tend to reduce the flow of information from low-malarious to the higher-malarious regions of the world (Malaney and Sachs (2002: 684) and Malaney, Speilman and Sachs (2004: 143 – 4)). For instance, since foreign direct investment is one of the most important channels along which new knowledge and technologies flows into a developing country, the increased level of risk associated with foreign investment in comparatively highly malarious areas will retard the flow of useful information into these areas. In addition, the relatively high risk of foreigners contracting malaria, as a result of visiting malarious countries, also has the potential to impede the flow of relevant knowledge to these countries.

Also indicated in Figure 2 is that the risk of contracting malaria will influence, directly, the well-being of parents who see their children dying in relatively large numbers. This is turn influences the level of social well-being in the community concerned. Mention also should be made here of the matter of immunity from malaria, and specifically the immunity that children inherit, in the form of the sickle cell trait, from one parent. Those children that inherit this immunity receive a haemoglobin mutation which effectively renders the carrier comparatively immune to the transmission of malaria. However, if this sickle cell trait is inherited from both parents then the carrier often contracts Sickle Cell Anaemia – a condition that cannot be cured. This illness in effect incapacitates the infected individual for extended periods over a life which usually ends prematurely.[17]

Finally, it also is indicated in Figure 2 that malaria has an indirect influence on the level of social well-being. In this instance malaria operates along two channels to influence the level of institutions which, in turn, influences the level of social well-being. Along one channel an

[17] Far more detailed medical information on sickle cell anaemia is provided by the United States Department of Energy, Biological and Environmental Research, Human Genome Project Information Program, [Online], Found at:

http://www.ornl.gov/sci/techresources/HumanGenome/posters/chromosome/sca.shtml (as of 12/06/04).

increase in the level of malaria risk directly reduces the level of institutions. (This line of causation was indicated by Acemoglu *et al*. (2001).) Along the other channel an increased level of malaria risk reduces the level of education which, in turn, reduces the level of institutions. (This line of causation is suggested by a point made by Glaeser *et al*. (2004); *viz*. the higher the level of education in a community the higher the level of institutions. Glaeser *et al*. did not mention, however, the influence of malaria – or other diseases – on the level of education.) The possible influence of malaria risk on the level of institutions is discussed further in Section III.3.

III.2.3 Rest of the regressors

The other two regressors in the basic model are the institution of the rule of law and the protection of individual property rights (Rule) and the integration of a country into the world economy as measured by the logarithm of the openness of the economy to international trade (Lcopen). The preferred measure of integration is an index adopted from RST (2004), which is the natural logarithm of a ratio of nominal imports plus nominal exports compared to nominal GDP, measured in nominal US dollars averaged over all available data for the period 1950-1998.

It might be argued that this last variable, Lcopen, overlaps with Pop100km since the ease of access to water transport would facilitate international trade. This does not seem to be the case, however, as the simple correlation coefficient between these two variables for the cross-county data sets we use for these two variables is just 0.1193.

III.3 Instrumental variables

As indicated in Figure 1, endogeneity and reverse causality is assumed to exist between all the variables (except Geography) referred to in that figure. It could be argued that to allow for this endogeneity relevant IVs should be applied within a 2SLS estimation procedure. While this point of view will be questioned later in Section IV.4.1 for the present the IVs that are applied later are considered, briefly, here.

The IV for Rule is that provided by Acemoglu *et al*. (2001: 1370) - the mortality rates for colonists from the 17th, 18th and 19th centuries (henceforth logem4). The hypothesis is that a high level of these mortality rates within a given region would have reduced the rate of adoption of desirable economic institutions in this region. In addition, since these colonial mortality rates are unlikely to be correlated with the current levels of economic activity, it might be thought that this variable serves as an appropriate instrument for contemporary measures of institutional quality. This belief is questioned in Section IV.4.1.

Frankel and Romer (1999) construct an instrument for international trade by estimating a bilateral trade equation and then aggregate the fitted values for this equation in order to estimate the geographic component of a nation's overall trade relations. The constructed trade share variable (logfrankrom) is used as the IV for international integration.

The IV chosen to control for the potential endogeneity of malaria risk is an index (me) devised from Kiszewski *et al*. (2004) that estimates a nation's malaria burden based upon its comparative ecologically-based transmission vectors. This index has been criticized by RST (2004, 151) amongst others as not being truly exogenous and, therefore, not a satisfactory IV for malaria risk. This view could be challenged since this index measures mosquito species, parasite type, regional ambient temperatures and rainfall conditions – factors that tend to be exogenous to any reasonable intervention by human action. Hence, this IV appears to possess the required desired properties. (Gundlach (2004, 10) makes a similar point.) Nevertheless, the use of this IV can be questioned for reasons other than mentioned by RST (2004). The relevant argument is sketched out later in Section IV.4.1.

IV Estimated Regression Results

IV.1 Ordinary least squares estimations

As a point of reference the influence of diseases, institutions, geography and integration on HDI is estimated using ordinary least squares (OLS).[18] In the regression results presented in Table 2 allowance has been made for heteroscedasticity within the context that the data set is of a limited or finite size.[19] (This implies that robust t-statistics are estimated.)

Table 2 comes about here.

---

[18] A list of data sources used to derive all the regression equations cited in the main text is to be found in Appendix A.

[19] See MacKinnon and White (1985). Also see Durlauf *et al*. (2006) and the references cited there on the matter of allowing for heteroscedasticity when the data set is of a finite size.

In this table a variation of the core OLS estimate of Equation 1 is set out in column 1. The estimated coefficients for Geography and Institutions are highly statistically significant and with their expected signs. The Integration coefficient is neither statistically significant, nor correctly signed. This result is consistent with the findings reported in RST (2004).

Next Disease (malaria risk) is added to the regression equation to be estimated. The estimated coefficients are set out in column 2 in Table 2. Again all estimated coefficients, except that for integration, have the desired properties. Hence malaria risk appears to matter in determining the level of HDI.

If these coefficient estimates are interpreted as accurately representing the causal links amongst each of these variables then, as a general rule, those nations with a low exposure to malaria risk, high quality institutions and a large proportion of their population residing near water transport could be expected to have substantially higher levels of social well-being (as defined) in the mid-1990s compared to those nations that were exposed to high levels of malaria risk, had low quality institutions and were landlocked without easy access to navigable rivers.

Next, to assess the importance of malaria risk in influencing the level of the HDI, the coefficient estimate for malaria risk (Mal94p) in Table 2, column 2, is multiplied by an improvement in the level of this variable (*viz*. the level of malaria risk is reduced in size) by one standard deviation (SD). (The value of the SD for this regressor is found in Table 1 and in the column 1 in Appendix C.) The resulting adjusted value (0.328 x 0.44) is 0.144. The same adjustment, *mutatis mutandis*, is made to the coefficient estimate for each of the other regressors listed in Table 2, column 2. These adjusted coefficient estimates are set out in

Appendix C column 2.[20] These calculations indicate that malaria risk has the greatest impact on the level of the HDI. Indeed this regressor has nearly twice as much impact on the level of the HDI than do Institutions (Rule), and over four times more impact than does the Geography (Pop100km). These calculations based on OLS estimates of relevant coefficients suggest, therefore, that diseases dominate in explaining the distribution of the HDI across all countries.

IV.2 Allowing for the sub-Saharan African dummy

Approximately 90 per cent of the morbidity due to malaria takes place in the tropical and sub-tropical regions of Africa (Hay *et al*. 2004: 333). This fact raises the following questions: is the disease burden created by malaria significant in its own right in explaining the cross-country variation in the level of the HDI, or is malaria risk simply acting as an alternative indicator for some other unidentified variable(s) closely related to Sub-Saharan Africa?

In the attempt to answer this question a dummy variable for sub-Saharan Africa – a variable that represents a range of factors unique to this area of the world – is included in the basic regression equation. The resulting coefficient estimates are set out in Table 2, column 3. The coefficient estimate for Disease remains statistically significant as do the coefficient estimates for all the other regressors (except that for Integration). The size of the coefficient estimate for Disease, however, now is smaller than previously.

[20] The reason for collecting these results in the table in Appendix C and not in a table in the main text is that this appendix will be referred to on a number of occasions throughout the main text.

These coefficient estimates next are adjusted as before to determine the influence of a particular variable on the level of the HDI. The results are set out in Appendix C column 3. Now Disease has 26 per cent more impact on the level of the HDI than does Institutions, and still over four times more impact than does Geography.

IV.3 Allowing for the indirect influence of disease on institutions

It is indicated in figures 1 and 2 that Diseases also may influence the level of Institutions which in turn will influence the level of the HDI. It also is indicated in figure 1 that Institutions may influence the level of Diseases. These matters are considered here.
The OLS regression results presented in Table 3, column 1 tend to support the conjecture that Diseases influence the level of Institutions. In that regression the levels of Disease (for 1994) and the average years of schooling (over the period 1970 through to 1985) have a lagged influence on the average level of the rule of law (for 2000 and 2002). All the coefficients have the expected sign and are highly statistically significant.[21]

---

[21] In estimating the regression results presented in Table 3 the variable logarithm GDP per capita for 1995 (GDPPC) was omitted from the regression equation. The reason is that there is a high level of multicollinearity between GDPPC and malaria risk, and GDPPC and the average years of schooling. (The simple correlation coefficient between GDPPC and malaria risk is 0.8205 and between GDPPC and schooling 0.8471.) In addition it is know that the level of malaria influences the level of GDPPC and not the other way round (Gallup and Sachs, 2001: 88 - 90). Also the comparatively long lag between the average years of schooling (for 1970 – 85) and GDPPC (for 1995) suggests

It follows that a one SD reduction in the level of Disease increases the level of the Institutions by (0.44 x 0.6862 =) 0.3019 units – which in turn would raise the level of the HDI. This calculation is taken into account in column 2 in Appendix C. Now Disease has 2.29 times more impact than do Institutions on the level of the HDI. A similar re-calculation is done for column 3 in that appendix (where the sub-Saharan African dummy variable is included in the relevant regression equation in Table 2) so that now Disease has 58 per cent more impact on the level of the HDI than do Institutions. These re-calculations are shown in brackets in columns 2 and 3 respectively in Appendix C.

The lagged structure present in the regression equation represented in Table 3, column 1 implies that in that equation there should be no reverse causation going from the average level of the rule of law (for 2000 and 2002) through to the level of the malaria risk (in 1994) and the average level of schooling (from 1970 through to 1985). An alternative test of the possible presence of this reverse causality was devised by making use of a two-stage least squares estimation procedure. In this case an instrumental variable (the mortality rate for settlers) was used as the instrument for the rule of law (for the 1990s). In this test it was found that the rule of law had no statistically significant influence on malaria risk. The relevant coefficient estimates are set out in Table 3 column 2.

---

that it is the level of schooling that influences the level of GDPPC and not the other way round. It follows that it seems reasonable to omit GDPPC from the relevant regression equation rather than the malaria risk and the average years of schooling.

IV.4 Two-stage least squares (2SLS) estimations

IV.4.1 Should instrumental variables be applied?

As indicated in Figure 1 and within the context of the basic regression model set out in equations 1 through to 4 above, it is assumed that there is reverse causality running from the level of the HDI through to the level of each of the regressors on the right-hand side of (1) - except for Geography. Does it follow, however, that the attempt should be made to correct for this assumed endogeneity by making use of IVs within a 2SLS estimation procedure when estimating relevant regression equations? There are grounds for arguing that no such correction should be made.[22]

First, while the presence of endogeneity will bias the relevant coefficient estimate away from zero, there are other factors at work that may bias these estimates in the opposite direction. One such factor is measurement errors that are present in the estimation of the coefficients presented in Table 2. This is so since most of the regressors in (1) are only proxy variables for the true values of the variables we wish to be able to measure. (For instance, the variable used to measure the influence of malaria - the risk of contracting malaria in a particular country - does not measure directly the level of mortality and morbidity caused by this disease in the country concerned. Similarly, the regressor, Rule, is only a proxy variable for the quality of economic institutions.) That being the case the coefficient estimates for regression equations 3 and 4 in Table 2 will suffer from attenuation bias. It is not possible to determine the direction of this bias, however, given

[22] The discussion in the main text has been influence by that to be found in Durlauf *et al.* (2006).

that a number of the regressors in those regression equations suffer from measurement errors (Green, 2003: 83 – 6).

Another bias probably is caused by the presence of omitted variables. This is so since there is a wide range of regressors that possibly should have been included in the regression equations set out in Table 2. Because of the presence of multicollinearity, however, these variables could not be identified. Consequently they were omitted from those regression equations. The omission of these variables also will tend to result in relevant coefficient estimates being biased.

It follows that the influence of the simultaneous presence of endogeneity, measurement errors and omitted variables within a relevant regression equation - estimated by using OLS - implies that it is not at all clear in which direction the coefficient estimates, set out in Table 2, are biased. The overall consequence possibly is that these estimates are not unduly biased. If that is the case then IVs should not be used (along with 2SLS) to re-estimate the regression equations set out in Table 2. This is especially so if the IVs that are available do not possess satisfactory properties. This matter is turned to next.

A satisfactory IV must possess the following two properties: (i) it is adequately correlated with the regressor for which it is the instrument and (ii) it is uncorrelated with the error term in the regression equation being estimated. While the IVs listed in Table 1 may satisfy property (i) some do not satisfy (ii). For instance, the IV for the rule of law - setter mortality rates for the 17^{th} to 19^{th} century - will be correlated with contemporary variables such as tropical climate, regional dummies and malaria risk in a given country. In addition, these latter variables probably suffer from measurement errors (since they do not measure what is actually required). Consequently setter mortality rates if used as an IV in a 2SLS estimation of a relevant regression equation will tend to be correlated with the error term in this equation. (Similar remarks apply to the IV (malaria ecology) for malaria risk.) The resulting coefficient estimates consequently will tend to be biased. On combining these two lines of argument it follows that the use of the IVs listed in Table 1 in a 2SLS estimation of relevant regression equations may result in coefficient estimates that are less satisfactory (*viz*. more biased) than those derived by just using OLS to estimate these equations. For the sake of completeness, however, some 2SLS estimates of relevant regression equations will be provided here by making use of the IVs listed in Table 1.

IV.4.2 Applying the instrumental variables

In estimating regression equation (1), based on 2SLS procedures that make use of regression equations (2), (3) and (4), it needs to be established whether the proposed instruments are sufficiently strongly identified with each of the endogenous regressors (in equation (1)). To do so use is made of the test proposed by Stock and Yogo (2002).[23] This test requires that if a particular instrument is sufficiently strongly identified with its endogenous regressor then the estimated F-statistic, for the first stage estimation of the 2SLS estimation procedure, will

[23] The basic reason for applying this test is that if the instrumental variables are comparatively weakly identified with the relevant endogenous regressor then the second stage estimation may generate a bias in the estimated coefficients towards those estimated by using OLS (Stock (2001: 7580) and Stock and Watson, (2003: Ch. 10)).

exceed a certain critical value. (A table of critical values is found in Stock and Yogo (2002, Table 2).) The sizes of these critical values of the F-statistic are conditional on the number of endogenous regressors within the regression equation and on the number of IVs that are being applied. (In the present case there are three endogenous explanatory variables and three instruments.)

A difficulty with applying this test in practice is that the critical values of the F-statistic are only provided in Stock and Yogo (2002, 59) up to a maximum of two endogenous regressors. Fortunately, however, it has been possible to extend this table up to three endogenous explanatory variables.[24] That being the case, the critical value in the present case - three endogenous regressors and three instruments - is 4.23.[25] This value is well below the smallest F-statistic (12.94) for the first-stage regression equations. The F-statistic for these equations is reported at the bottom of column 1 in Table 4. Hence the IVs are non-weak for each of the endogenous variables.

---

[24] These calculations were made possible through the generosity of two people: Motohiro Yogo who provided the Gauss program used in the calculations reported in Table 2 in Stock and Yogo (2002, 59), and Heather Anderson who extended this program and carried out the time-consuming calculations required for the three-endogenous-variable case. The relevant calculations are available upon request.

[25] In Appendix B is set out a truncated version of Table 2 in Stock and Yogo (2002) plus the additional critical values for the case where there are three endogenous variables.

Also set out in Table 4, column 1 are the 2SLS coefficient estimates and the descriptive statistics for the core regression equation. As can be seen there Institutions are no longer statistically significant whereas Disease is. (While the relevant regression results are not shown, the inclusion of the sub-Sahara Africa dummy variable does not improve the situation since the coefficient estimate for this regressor is not statistically significant.)

Table 4 goes about here.

The regression estimates presented in Table 4 column 1, when multiplied by the SD for the relevant regressor, imply that Disease has a greater impact on the level of the HDI than any of the other regressors. As reported in Appendix C, Disease has some 3.41 times greater impact on the level of the HDI than do Institutions. But clearly this calculation is biased strongly in favor of Institutions since this regressor is not nearly statistically significantly different from zero according to the standard two tailed test of significance. Disease also has 2.20 times more impact on the level of the HDI than does Geography.

V Expanding the Limited Sample Size

A reasonable criticism that can be leveled at the coefficient estimates presented in Tables 4 column 1 is that they are based upon a comparatively small data set. This is the consequence of the use of the mortality rates for colonial settlers in the 17^{th} through to the 19^{th} centuries (logem4) as an IV for the rule of law. The question is, however, what IV, or IVs can be used in place of logem4. Here we turn to Hall and Jones (1999: 100) where they argue that the influence that Western culture has had on the institutions that have evolved throughout the world can be measured by the percentage of a nation's current population that speak either

English (engfrac) or a Western European (eurfrac) language as their first language.[26] The use of both or one of these alternative IVs will allow the data set to be expanded from 71 to around 120 sets of observations – about a 70 per cent increase in the sample size. This increase should assist in increasing the accuracy of the regression coefficient estimates by reducing the variance of these estimates. That said these alternative IVs may be weak (according to the Stock-Yogo test) and, if so, should not be used.

To test if this is the case to begin with we experimented with combinations of the alternative IVs (engfrac and eurfrac). The most satisfactory results were derived by applying enfrac as the IV for the rule of law. The F-statistics for each of the first-stage regression equations on applying this IV are set out at the bottom of column 2 in Table 4. Each is greater than the critical value of 4.23 and hence engfrac can be regarded as a non-weak instrument, along with the other IVs listed in Table 4, column 2.

Table 4 goes about here.

The coefficient estimates that are derived by applying this alternative set of IVs also are set out in Table 4 column 2. On comparing these results with those in column 1 in that table it is clear that Disease and Geography now are more statistically significant than when the settler mortality rate (*viz*. logem4) is used as the IV for the rule of law. Also Institutions and Integration now are even more statistically insignificant than before. (While not shown, introducing the dummy variable for sub-Saharan Africa into the regression equation does not

[26] These IVs also have been used by Dollar and Kraay (2003).

alter these results since the coefficient for this regressor is not statistically significantly different from zero on applying to the standard two tailed test of significance.)

The coefficient estimates in Table 4 column 2 also are adjusted by multiplying them with the SD for the relevant regressor to determine the influence that a particular regressor has on the level of the HDI. The consequent results set out in column 5 in Appendix C indicate that in this instance Diseases has 5.68 times more impact on the level of the HDI than do Institutions. This is so even though this calculation is biased strongly in favor of Institutions since (as just noted) the coefficient estimate for this regressor is not statistically significantly different from zero. Disease also has 5.43 times more impact than does Geography on the level of the HDI.

VI The Influence of Disease and Institutions on Different Groups of Countries

VI.1 Econometric results

The econometric results reported in Tables 2 and 4 and in Appendix C indicate that Disease dominates in explaining the level of the HDI for the average country in the late 20th Century. Is this still the case for sub-groups of countries? Clearly for one sub-group of countries this cannot be so for the citizens of these countries, in the mid-1990s, faced no risk of contracting malaria. Indeed, for the average country in this group Institutions now completely dominates in determining the level of the HDI for this country. This is indicated by the relevant coefficient estimates for this group of countries reported in Table 5, column 3.

As for the group of countries that faced a non-zero malaria risk in the mid-1990s, Disease dominates again in determining the level of the HDI for the average country in this group. If the dummy variable for sub-Saharan Africa is taken into consideration then Disease is the only other regressor (besides the sub-Saharan Africa dummy) that is statistically significant. The relevant coefficient estimates are reported in Table 5, columns 1 and 2.

(OLS was used to derive these various regression results. Attempts made to use 2SLS to derive relevant regression results had to be abandoned. This was because the relatively low values for the F-statistics for the estimated first stage regression equations indicated (on applying the Stock-Yogo test) that the IVs being applied were weak. To save on space these results are not presented here.)

VI.2 Conjecture regarding development policy

The results presented in Table 5 along with those in Tables 2 and 4 suggest a conjecture regarding the path that a country, with a low level of the HDI and facing malaria risk, should adopt in the attempt to achieve for its citizens a sustained increase in the level of social well-being. Specifically, first malaria and associated infectious diseases have to be contained. Once this has been accomplished then it becomes essential to provide the rule of law and the protection of individual property rights. The first part of this line of argument is suggested by the coefficient estimates in Tables 2 and 4 (when taken in conjunction with the results reported in Appendix C), and those in Table 5 columns 1 and 2 – results that indicate that while malaria and associated diseases are virulent within the community the level of well-being for this country is severely constrained by these diseases. Once these diseases have been contained the coefficient estimates in Table 5, column 3 suggest that economic institutions come into their own and make a fundamental contribution towards increasing the level of social well-being for the average country in that group of countries that is free of the scourge of malaria and related diseases.

VII Conclusions

The quasi-controlled experimental evidence derived from History-plus-Nature cited in Section II indicates that even with the institution of the rule of law and the protection of property rights of a comparatively high quality in place in a community, the presence of an epidemic of a virulent infectious disease in this community will severely adversely affect the level of economic activity and well-being in this community. This quasi-controlled experimental evidence is then combined with the fact (noted in Section III.2.2) that in the mid-1990s the risk of contracting malaria (where malaria risk is a proxy variable for the morbidity and mortality caused by malaria and a number of associated diseases) was widespread but varied in intensity across the regions of the globe. The implication of these two broad sets of observations is that it allows the conjecture that the risk of contracting malaria must explain a significant amount of the variation in the level of the HDI across countries in the mid-1990s.

This conjecture is not contradicted by the empirical results presented in Tables 2, 4 and 5 and the summary of the appropriate adjustments of these results in Appendix C. These results - especially those in Appendix C - indicate that Disease (as measured by the risk of contracting malaria) dominates over all other possible deep determinants that might contribute to the level of the HDI for the average country in the late 20th Century.

It is open to doubt, however, whether Disease is as dominant as that indicated by the summary results set out in columns 4 and 5 in Appendix C. This is so since these results are based upon coefficient estimates (reported in Table 4) that were derived by applying 2SLS estimation techniques along with certain instrumental variables. And as was pointed out in Section IV.4.1, the use of these techniques along with applying questionable IVs may not produce results that are to be preferred to those derived by applying OLS to generate relevant results (that are reported in Tables 2 and 5).

Nevertheless, whichever results are regarded as being the more reliable, the results reported in Appendix C indicate that Disease dominates in the sense that the coefficient estimates for this variable are always statistically significant (when malaria risk is included in the relevant regression equation to be estimated) and this variable (when appropriately adjusted) always has the greatest impact on the level of the HDI for the average country in the late 20th Century. Economic institutions, however, appear to have a role to play in the development process once malaria risk has been brought under control.

There are a number of implications that follow on from these conclusions. First, since 58 per cent of all malaria cases occur in the poorest 28 per cent of the world's population (Barat, 2004) it follows that the countries that could be expected to contribute most to any increased effort aimed at containing this disease would be those wealthier countries of the world that are not faced with the risk of contracting malaria. (This additional effort also would need to be combined with an increased effort to contain the other diseases associated with malaria.) The theory of the supply of global public goods suggests, however, that comparatively little would be contributed by these countries as they would gain comparatively little economically from any contribution they may make.[27] Yet while this theory may suggest this, since about the mid-1990s on improvements have been made in the containment and treatment of malaria. For instance, Bremen *et al*. (2004: 1) note: 'There has been rapidly increasing optimism for greatly improved malaria control, particularly in Africa where the vast majority of people affected by the disease reside … . The new positive attitude is based on major advances in malaria research and public health practices, and careful epidemiological and economic analyses of malaria's burden and cost-effective ways to lessen it.'

If malaria and related diseases indeed are contained throughout the world in the next few decades naturally this should result in a narrowing of the variation in the HDI across countries. This narrowing would be even greater, as emphasized in the previous section, if the reduction in the level of diseases in the countries concerned was also accompanied by an improvement in the level of economic institutions such as the rule of law and the protection of individual property rights.

[27] See Kaul, Gruber and Stern (1999), and Smith *et a*l. (2003).

REFERENCES

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

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

___________, 'Institutions as the Fundamental Cause of Long-Run Growth', in P. Aghion and S. Durlauf (eds) (2006).

P. Aghion and S. Durlauf (eds), *Handbook of Economic Growth* (Amsterdam: Elsevier, 2006).

L. M. Barat, 'Do Malaria Control Interventions Reach the Poor? A View Through the Equity Lens', *American Journal of Tropical Medicine and Hygiene*, 71(Suppl. 2) (2004) 174 – 178.

R. J Barro and J.-W. Lee, 'International data on Educational Attainment: Updates and Implications', *Center for International Development*, Working Paper No. 42 (2000).

W. J. Bernstein, *The Birth of Plenty: How the Prosperity of the Modern World was Created.* (New York: McGraw-Hill, 2004).

D. E. Bloom and J. Sachs, 'Geography, Demography, and Economic Growth in Africa', *Brookings Papers on Economic Activity*, 1998 (2), 207 - 96.

J. Breman, 'The Ears of the Hippopotamus: Manifestations, Determinants, and Estimates of the Malaria Burden', *American Journal of Tropical Medicine and Hygiene,* 64 (2001) 1 – 11.

J. Breman, G. J. Alilio, S. Martin and A. Mills, 'Conquering the Intolerable Burden of Malaria: What's New, What's Needed: A Summary', *American Journal of Tropical Medicine and Hygiene*, 71 (suppl. 2) (2004) 1 – 15.

F. F. Cartwright, *Disease and History* (London: Rupert-Davis, 1972).

P. Dasgupta, *An Inquiry into Well-Being and Destitution* (Oxford: Clarendon Press, 1993).

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

D. Dollar and A. Kraay, 'Institutions, Trade and Growth', *Journal of Monetary Economics*, 50 (2003) 133 – 162.

S. N. Durlauf, P. A. Johnson, P.A. and J. R. W. Temple, 'Growth Econometrics', in Aghion and Durlauf (2006).

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

R. Foster and C. Leighton, 'Economic Impacts of Malaria in Kenya and Nigeria', Research Paper 6, *Health Financing and Sustainability Project*, Agency for International Development, U.S.A. (1993).

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

J. L. Gallup, A. Mellinger and J. Sachs, 'Geography and Economic Development', *National Bureau of Economic Research*, (Cambridge, MA) Working Paper No. 6849 (1998).

_________, 'Climate, Water Navigability and Development', *Centre for International Development*, Harvard University, Working Paper 24, (1999).

J. L. Gallup and J. D. Sachs, 'The Economic Burden of Malaria', *American Journal of Tropical Health and Hygiene*, 64 (2001) 85 - 96.

E. L. Glaeser, R. La Porta, F., Lopez-De-Sianes and A. Schleifer, 'Do Institutions Cause Growth?', *Journal of Economic Growth,* 9 (2004) 271 – 303.

W. Greene, *Econometric Analysis*, Fifth Edn (Upper Saddle River, New Jersey: Prentice Hall, 2003).

E. Gundlach, 'The Primacy of Institutions Reconsidered: The Effects of Malaria Prevalence in the Empirics of Development', *Kiel Institute for World Economics* (Germany) Working Paper No. 1210 (May 2004).

R. E. Hall and C. I. Jones, 'Why Do Some Countries Produce So Much More Output per Worker than Others?', *The Quarterly Journal of Economics*, 114 (1999) 83 - 116.

S. I. Hay, C. A. Guerra, A. J. Tatem, A. M. Noor and R. W. Snow, 'The Global Distribution and Population at Risk of Malaria: Past, Present, and Future', *The Lancet: Infectious Diseases*, 4 (2004) 327 – 336.

J.N. Hays, *The Burdens of Disease* (New Brunswick, N.J.: Rutgers University Press, 2003)

A. Heston, R. Summers and B. Aten, Penn World Tables Version 6.0, *Centre for International Comparisons at the University of Pennsylvania* (CICUP) (2001, December). Data Website: http://pwt.econ.upenn.edu/.

D. Kaufmann, A. Kraay, and P. Zoido-Lobatón, 'Governance Matters II—Updated Indicators for 2000/01', *World Bank Policy Research Department*, Working Paper 2772 (2002).

D. Kaufmann, A. Kraay, and P. Zoido-Lobatón, 'Governance Matters IV—Governance Indicators for 1996 – 2004', *World Bank Policy Research Department*, Working Paper 3630 (2005).

I. Kaul, I. Gruberg and M. A. Stern (eds), *Global Public Goods* (New York: Oxford University Press, 1999).

J. Keiser, B. Singer, and J. Utzinger, 'Reducing the Burden of Malaria in Different Eco-Epidemiological Settings with Environmental Management: A Systematic Review', *Lancet Infectious Diseases*, 5 (2005) 695 – 708.

K. F. Kiple (ed.), *Cambridge World History of Human Diseases* (Cambridge: Cambridge University Press, 1993).

A. Kiszewski, *et al*., 'A Global Index Representing the Stability of Malaria Transmission', *American Journal of Tropical Medicine and Hygiene*, 70 (2004) 486 - 498.

M. Kremer and E. Miguel, 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities', *Econometrica*, 72 (2004) 159 - 217.

J. MacKinnon and H. White, 'Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties', *Journal of Econometrics*, 29 (1985), 305 – 325.

J. McArthur and J. Sachs, ‘Institutions and Geography: Comment on Acemoglu, Johnson and Robinson 2000’, *National Bureau of Economic Research* (Cambridge, MA) Working Paper No. 8114 (2001).

D. McCullough, *The Path Between the Seas: The Creation of the Panama Canal 1870 – 1914* (New York: Simon and Schuster, 1977).

P. Malaney, and J. D. Sachs, ‘The Economic and Social Burden of Malaria’, *Nature*, 415 (6872) (2002), 680-85.

P. Malaney, A. Speilman and J. Sachs, ‘The Malaria Gap’, *American Journal of Tropical Medicine and Hygiene*, 71 (Supp. 2) (2004)141 – 146.

L. Molineaux, ‘Malaria and Mortality: Some Epidemiological Considerations’, *Annals of Tropical Medicine and Parasitology*, 91(1997) 811 - 825.

C. J. L. Murray and A. D. Lopez, ‘Mortality by Cause for Eight Regions of the World: Global Burden of Disease Study’, *The Lancet*, 349 (1997a) 1269 – 76.

_________________, ‘Global Mortality, Disability, and the Contribution of Risk Factors: Global Burden of Disease Study’, *The Lancet*, 349 (1997b) 1436 – 42.

F. Noorbakhsh, ‘The Human Development Index: Some Technical Issues and Alternative Indices’, *Journal of International Development*, 10 (1998) 589 – 605.

D. C. North and R. P. Thomas, *Rise of the Western World* (Cambridge: Cambridge University Press, 1973).

D. G. North, *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990).

________ 'Economic Performance Through Time', *The American Economic Review*, 84 (1994) 359 - 368.

_________ 'The New Institutional Economics and Third World Development', in J. Harriss, J. Hunter and L. M. Lewis (eds), *The New Institutional Economics and Third World Development* (London: Routledge, 1995).

R. Porter, *The Greatest Benefit to Mankind: a Medical History of Humanity from Antiquity to the Present* (London: Harper Collins, 1997).

S. Radelet and J. Sachs, 'Shipping Costs, Manufactured Exports, and Economic Growth', Paper presented at the annual meetings of the American Economics Association (January, 1998).

G. Ranis, F. Stewart and A. Ramirez, 'Economic Growth and Human Development', *World Development*, 28 (2000), 197 – 219.

G. Ranis and F. Stewart, 'Growth and Human Development: Comparative Latin American Experience', *Economic Growth Centre*, Yale University, Discussion Paper No. 826 (2001).

G. Ranis, 'Human Development and Economic Growth', *Economic Growth Centre,* Yale University, Discussion Paper No. 887 (2004).

M. Ravallion, 'Good and Bad Growth: The Human Development Reports', *World Development*, 25 (1997), 631 - 638.

L. G. Reynolds, 'The Spread of Economic Growth to the Third World: 1850 – 1980', J*ournal of Economic Literature*, 21 (1983) 941 – 980.

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

W. W. Rostow, *Stages of Economic Growth: a Non-Communist Manifesto* (Cambridge: Cambridge University Press, 1960).

J. D. Sachs, 'Institutions Don't Rule: Direct Effects of Geography on Per Capita Income', *National Bureau of Economic Research* (Cambridge MA) Working Paper No. 9490 (2003).

A. K. Sen, *Commodities and Capabilities* (Oxford: Oxford University Press, 1985).

________, *Development as Freedom,* 1st edn (New York: Knopf Press, 1999).

A. Smith, *An Inquiry into the Nature of the Causes of the Wealth of Nations* (Oxford: Oxford University Press, 1976 (1776)).

R. Smith, *et al*. (eds), *Global Public Goods for Health* (Oxford: Oxford University Press, 2003).

R. W. Snow, E. L. Korenromp and E. Gouws, 'Pediatric Mortality in Africa: *Plasmodium Falciparum* Malaria as a Cause or Risk?', *American Journal of Tropical Medicine and Hygiene*, 71(Suppl 2) (2004)16 – 24.

A. Spielman and M. D'Antonio, *Mosquito: The Story of Man's Deadliest Fo*e (London: Faber and Faber, 2001).

J. H. Stock, 'Instrumental Variables in Statistics and Econometrics', *International Encyclopaedia of the Behavioural Sciences* (Amsterdam: Elsevier Publishing, 2001), pp. 7577 - 7582.

J. H. Stock, and M. Yogo, 'Testing for Weak Instruments in Linear IV Regression', Paper Prepared for the Festschrift in Honor of Thomas Rothenberg, *National Bureau of Economic Research*, (Cambridge, MA), Technical Working Paper No. 284 (2002).

J. H. Stock, and M. W. Watson, *Introduction to Econometrics* (Boston: Addison Wesley, 2003).

United Nations Children's Fund, *Progress Since the World Summit for Children: Statistical Review*. (Paris: United Nations Children's Fund, 2001).

United Nations. *United Nations Human Development Report,* 1997 (New York: Oxford University Press, 1997).

J. Utzinger, Y. Tozan and B. H. Singer, 'Efficacy and Cost-Effectiveness of Environmental Management for Malaria Control', *Tropical Medicine and International Health*, 6 (2001) 677 – 687.

J. Utzinger, Y. Tozan, F. Doumani and B. H. Singer, 'The Economic Payoffs of Integrated Malaria Control in the Zambian Copperbelt Between 1930a and 1950', *Tropical Medicine and International Health*, 7 (2002) 657 – 677.

P. H. van der Brug, 'Malaria in Batavia in the 18th Century', *Tropical Medicine and International Health*, 2 (1997) 982-902.

M. Watson, *The Geographical Aspects of Malaria* (Washington, D.C: United States Printing Office 1943).

M. Watson, *African Highway: The Battle for Health in Central Africa* (London, John Murray, 1953).

World Health Organization *World Health Report. Reducing Risks, Promoting Health Life* (Geneva: World Health Organization, 2002).

Figure 1. Channels of Inter-Causality between the Deep Determinants:
Extended Model Separating Geography into Disease Burden and Ease of Access to Water Transport

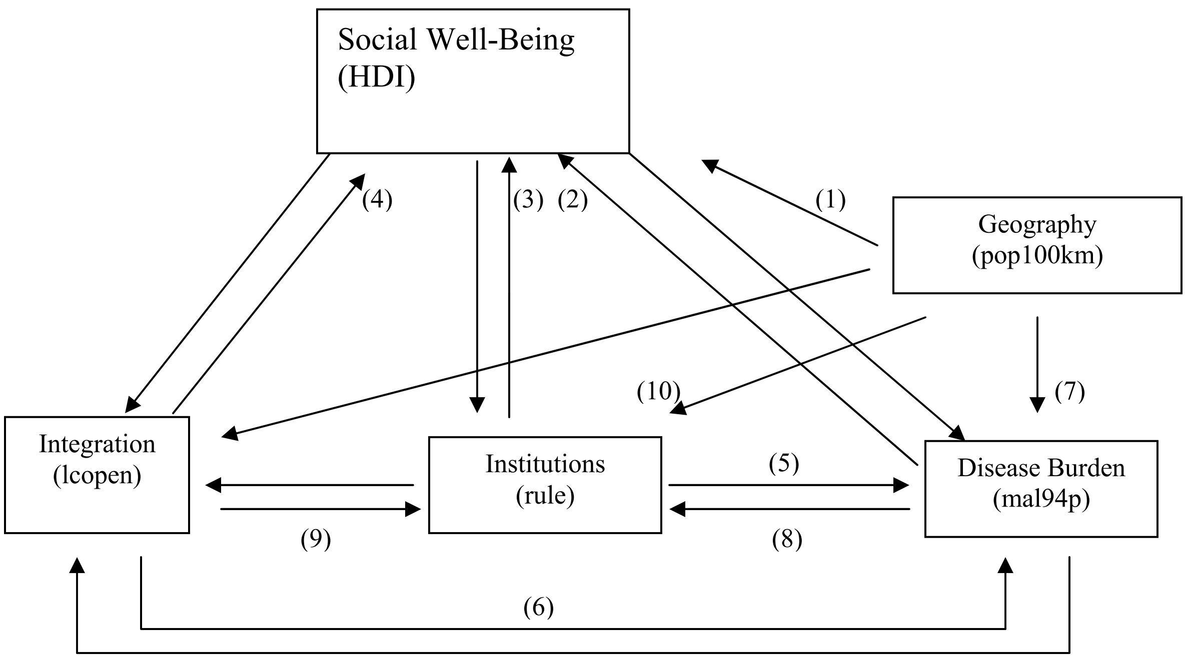

Figure 2. Flow Chart of Hypothesized Channels of Influence of Disease Burden on Levels of Social Well-Being

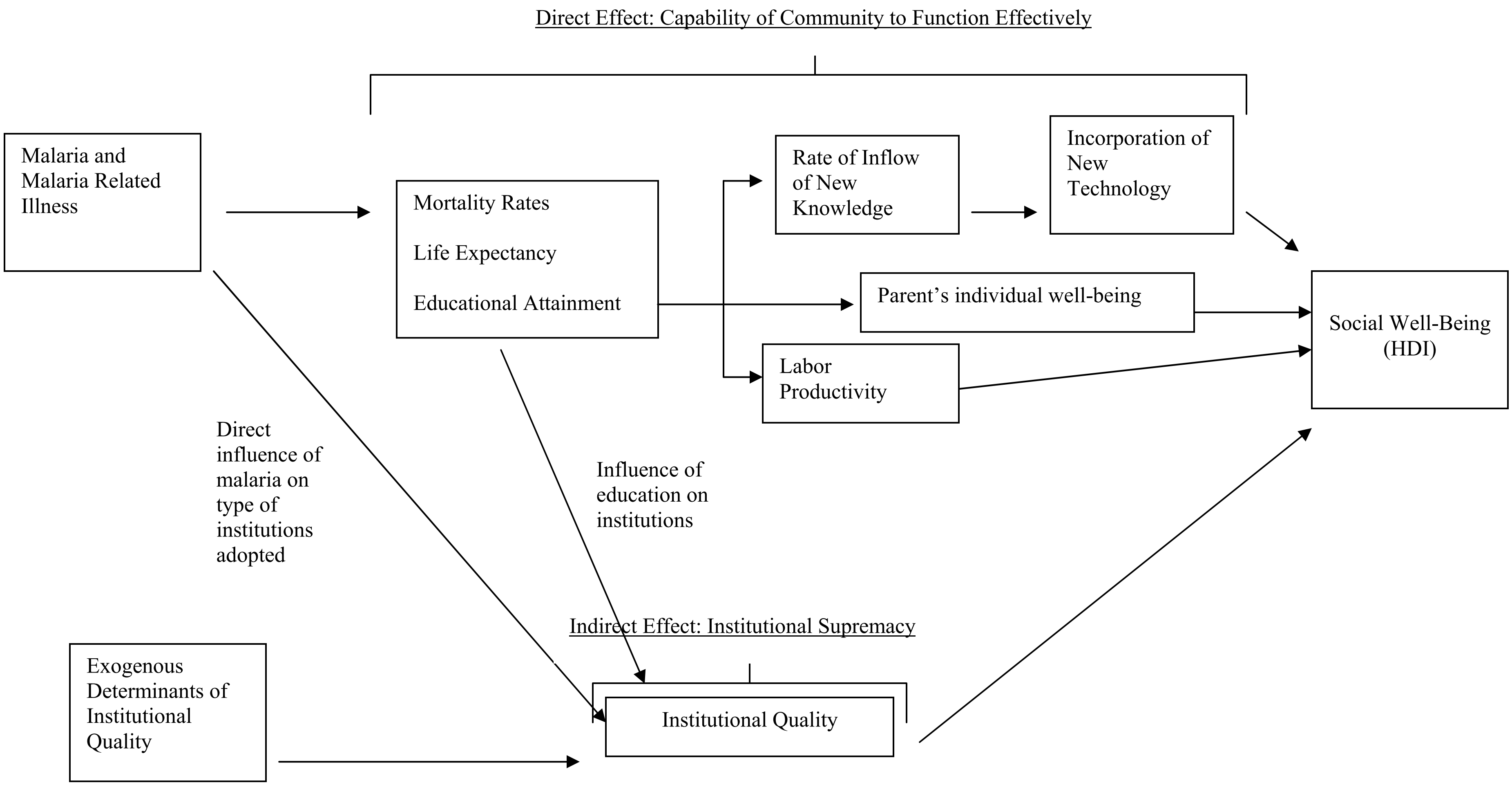

Table 1. Descriptive Statistics of Key Variables.

	(1) Mean	(2) Standard Deviation	(3) Max Observations	(4) Min	(5) Max
Human development Index 1994 (HDI94)	0.6643	0.23	169	0.176	0.98
OLS ESTIMATORS					
Rule of Law (Rule)	0.003	0.95	168	-2.16	1.91
Log Nominal Openness (Lcopen)	4.12	0.59	170	2.55	5.78
Population within 100km of Navigable Waterways 1995 (Pop100km)	432	0.37	150	0	1.00
Percentage of Population at Risk of Contracting Malaria (Mal94p)	0.36	0.44	165	0	1.00
INSTRUMENTAL VARIABLES					
Log European Settler Mortality during 17th, 18th and 19th Centuries (logem4)	4.66	1.21	81	2.14	7.99
Percentage Population Speaking English (engfrac)	0.07	0.24	182	0	1.00
Percentage Population Speaking European Languages (eurfrac)	0.22	0.38	185	0	1.00
Log Constructed Trade Share (logfrankrom)	3.01	0.80	163	0.83	5.64
Malaria Ecology (me)	3.66	6.42	173	0	31.55

Note: Sources for all data provided in Appendix A

Table 2 Determinants of Development: The Human Development Index for 1994 is the Dependant Variable using Ordinary Least Squares Estimations

Regressors	(1)	(2)	(3)
Integration (Lcopen)	0.0165 (0.64)	-0.0108 (-0.61)	-0.0035 (-0.20)
Geography (Pop100km)	0.0985*** (2.70)	0.0953*** (3.35)	0.0649** (2.42)
Institutions (Rule)	0.1606*** (12.96)	0.0766*** (6.95)	0.0811*** (7.25)
Disease (Mal94p)	-	-0.328*** (-12.19)	-0.2197*** (-6.51)
Sub-Saharan Africa Dummy Variable			-0.1465*** (-4.51)
Adj. R-Squared	0.5066	0.7685	0.8060
No. Obs.	138	138	128
F-Statistic	83.46	184.03	145.09

Notes: Robust *t*-statistics in Parenthesis: according to the two tailed test significant at the: 5 per cent level *: 2 per cent level**: 1 per cent level ***.

Table 3 Factors Determining the Level of the Rule of Law 2000 – 2002 and the Level of Malaria Risk (1994)

	1	2
Regressors	Rule of Law 2000 – 2 OLS estimate	Disease (mal94p) 2SLS estimate
Disease (mal94p)	-0.6862*** (-3.48)	
Geography (Pop100km)	0.6204*** (3.60)	
Average Years of Schooling 1970 – 1985	0.1835*** (6.41)	0.0711 (0.63)
Institutions (Rule 1990s)		0.5538 (-1.35)
Latin American Dummy Variable	-0.7422*** (-4.63)	
Sub-Saharan Dummy Variable		-0.3657*** (2.86)
Adj R-squares	0.6785	0.1164
No. Obs.	103	59

Notes: Robust *t*-statistics in Parenthesis: according to the two tailed test significant at the: 5 per cent level *: 2 per cent level**: 1 per cent level ***.

Table 4 Deep Determinants of Development: The Human Development Index for 1994 is Dependant Variable in Two Stage Least Squares Estimations

Regressors	(1)	(2)
Integration (Lcopen)	-0.1184 (-1.65)	-0.0435 (-0.94)
Geography (Pop100km)	0.2086* (2.06)	0.0983** (2.50)
Institutions (Rule)	0.0485 (0.69)	0.2530 (0.48)
Disease (Mal94p)	-0.3572** (-2.60)	-0.4495*** (-4.23)
Instruments	(logfrankrom, logem4, me)	(logfrankrom, engfrac, me)
Adj.R-Squared	0.6308	0.7283
No. Obs.	71	120
F-Statistic	30.49	53.92
F-Statistic for each of the first stage regression equations	20.71 12.94 22.91	25.57 10.92 25.14

Notes: Robust *t*-statistics in Parenthesis: according to the two tailed test significant at the: 5 per cent level *: 2 per cent level**: 1 per cent level ***.

Table 5: The Influence of the Deep Determinants of Development on the HDI When the Risk of Contracting Malaria and Associated Diseases in a Group of Countries is non-Zero or Zero

	Dependent Variable		
	HDI for countries for which the risk of contracting malaria is **non- zero**		HDI for countries for which the risk of contracting malaria is **zero**
Regressors	OLS Regression	OLS Regression	OLS Regression
	1	2	3
Integration (Lcopen)	0.0283 (0.85)	0.0599 (1.80)	-0.0258 (-1.95)
Geography (Pop100km)	0.1830*** (3.69)	0.0872 (1.52)	-.0150 (0.64)
Institutions (Rule)	0.0488 (1.65)	0.0482 (1.82)	0.1034*** (12.05)
Disease (Mal94p)	-0.4105*** (-9.05)	-0.3177*** (-7.08)	
Sub-Saharan Africa Dummy Variable		-0.1393*** (-3.61)	
Adj. R-squared	0.6145	0.6761	0.6398
No. of observations	71	71	67

Notes: Robust *t*-statistics in Parenthesis: according to the two tailed test significant at the: 5 per cent level *: 2 per cent level**: 1 per cent level ***.

Appendix A for Diseases Dominate: Data and Sources

(asiae) = Dummy variable taking value of one if a country belongs in South East and East Asia and zero if otherwise. Source: Own determination

(engfrac) = Fraction of the population which speak English. Source: Hall and Jones (1999).

(eurfrac) = Fraction of the population which speak a major Western European language, including, English, French, German, Portuguese or Spanish. Source: Hall and Jones (1999).

(hdi94) = United Nations Human Development Index for 1994. Source: United Nations Human Development Report (1997).

(laam) = Dummy variable taking value of one if a country belongs to Latin America or the Caribbean and zero if otherwise. Source: Own determination

(lcopen) = Natural logarithm of nominal openness averaged between 1950 and 1996. Source: Heston; Summers and Aten, (2001).

(logem4) = Natural logarithm of estimated European settlers' mortality rates. Source: Acemoglu *et al.* (2001).

(logfrankrom) = Natural logarithm of predicted trade shares based on a bilateral trade equation. Source: Jeffrey A. Frankel and David Romer (1999).

(mal94p) = Proportion of a population at risk of contracting malaria in 1994. Source: Gallup and Sachs (2003).

(me) = Estimate of a nation's malaria burden based upon its comparative ecologically based transmission vectors in 1995. Source: Kiszewski *et al.* (2004).

(pop100km) = Proportion of the population within 100km of a sea navigable waterway in 1995. Source: Gallup *et al.* (1999).

(rule) = Composite index for rule of law and the protection afforded to private property rights during the 1990s. Source: Kaufmann, Kraay and Zoido-Lobatón (2002).

(rule of law for 2000 - 2002) = Composite index for rule of law and the protection afforded to private property rights during the 2000 - 2002. Source: Kaufmann, Kraay and Zoido-Lobatón (2005).

(safrica) = Dummy variable taking the value of one if a country belongs to Sub-Saharan Africa and zero if otherwise. Source: RST (2004).

(schooling 1970 – 85) = Average years of schooling 1970 – 85. Source: Barro and Lee (2000). Data found at http://www.cid.harvard.edu/ciddata/ciddata.html.

(schooling 1970 – 85) = Average years of schooling 1990 – 2000. Source: Barro and Lee (2000). Data found at http://www.cid.harvard.edu/ciddata/ciddata.html.

Appendix B:

A Truncated Table of Critical F-Statistics to Test for Weak Instruments in Linear IV Regression. Based on Stock and Yogo (2002)

The Stock and Yogo (2002) weak instruments test rejects an instrument for a 2SLS estimation, if its estimated first stage F-Statistic is smaller than the critical value indicated within the table below. This critical value, g_{min}, is calculated as a function of the number of included endogenous regressors (n), the number of instrumental variables (K_2) and the desired maximal size (r) of a Wald test of $\beta = \beta_0$, when the significance level is 5 percent. In each of the justifications made in this paper, the minimum value of (r) is adopted to derive the critical rejection values to maximize the likelihood of the instruments passing the test for the maximal size r = 0.10. The table below, which in part is taken from Stock and Yogo (2002:59) and which is then extended to the n = 3 case (using the Gauss computer program provided by Motohiro Yogo and extended and run by Heather Anderson to generate the relevant critical values), reports these critical g_{min} values. (The full table, covering many more values of K_2 and two more values of r, is available upon request.)

Critical Values for the 2SLS Weak Instrument Test Based on Size of First Stage F-Statistic.

	n = 1, r =		n = 2, r =		n = 3, r =	
K_2	**0.10**	**0.15**	**0.10**	**0.15**	**0.10**	**0.15**
1	16.38	8.96				
2	19.93	11.59	7.03	4.58		
3	22.30	12.83	13.43	8.18	4.23	3.34
4	24.58	13.96	16.87	9.93	10.01	6.48
5	26.87	15.09	19.45	11.22	13.41	8.23
6	29.18	16.23	21.68	12.33	15.99	9.54
7	31.50	17.38	23.72	13.34	18.16	10.64
8	33.84	18.54	25.64	14.31	20.11	11.61

Appendix C. Standard Deviation Adjustments to the Size of the Estimates of Core Coefficients

	Standard Deviation	Table 2 (Col. 2) Dep. variable: HDI	Table 2 (Col. 3) Dep. Variable HDI	Table 4 (Col. 1) Dep. variable: HDI	Table 4 (Col. 2) Dep. variable: HDI
	1	2	3	4	5
Integration (Lcopen) x S.D.	**0.59**	0.006*	0.002*	0.07*	0.0257*
Geography (Pop100km) x S.D.	**0.37**	0.035	0.0.24	0.077	0.0364
Institutions (Rule) x S.D.	**0.95**	0.073	0.077	0.046*	0.0324*
Disease (Mal94p) x S.D.	**0.44**	0.144	0.097	0.157	0.1978
IV for (rule)				logem4	engfrac
Disease/ Institutions		**1.98 (2.29)****	**1.26 (1.58)****	**3.41**	**5.68**
Disease/Geography		**4.11**	**4.04**	**2.02**	**5.43**

* Coefficient estimate for variable concerned not statistically significantly different from zero at the 5 per cent level according to the two tailed test.

** Number in brackets is that calculated after allowing for the indirect influence of Disease on the level of Institutions.