UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

THE DETERMINANTS OF ECONOMIC DEVELOPMENT: INSTITUTIONS OR GEOGRAPHY?

ANDREA PRESBITERO

QUADERNO DI RICERCA n. 217

Settembre 2004

Comitato scientifico:

Renato Balducci
Marco Crivellini
Marco Gallegati
Alberto Niccoli
Alberto Zazzaro
Collana curata da:
Massimo Tamberi

The Determinants of Economic Development: Institutions or Geography?

di Andrea Presbitero[1]

Sintesi

This work investigates the roots of economic development. The debate about the predominance of institutions over geography has not yet reached a firm conclusion: this analysis wants to highlight the main difficulties that one should address in order to find which are the real determinants of long-run economic growth. I argue that the institutional view is not so strong as it may appear: different specifications and different institutional indicators undermine the exclusive importance of institutions. The results of Acemoglu, Johnson and Robinson in favour of the institutional approach are no more valid if other institutional indicators are used instead of the risk of expropriation. Geographical factors related to the health environment and the physical integration in the world markets play a role in the process of economic growth, beyond their effect on institutional development. Geography seems to be a factor even using different specification of the model. However, in this case, the Instrumental Variable procedure is far from being perfect: lack of sensible indicators and strong problem of endogeneity are the main difficulties. A closest look at the economic history and ecology should always be part of this sort of analysis.

JEL classification: C31, O10, O11, P16
Keywords: Economic development, Institutions, Geography, Instrumental variables

---

[1] Università Politecnica delle Marche – Dipartimento di Economia, Piazza Martelli 8, 60121 Ancona - Italy. E-mail: presbitero@dea.unian.it. I am very grateful to Prof. Robert K. Eastwood and Prof. M. Tamberi for the useful comments. All remaining errors are my own responsibility.

Indice

1. Introduction

The quality of institutions is widely considered one of the most important source of economic growth and development. In the last years, a lot of emphasis was given to the creation and to the development of good institution as a necessary condition for economic growth. Both the theory and the policy recommendations stress the importance of good institutions. Nonetheless, the differences in per capita GDP across countries do not seem to be explained exclusively by institutions. The economic literature recognizes three main sources of economic development: geography, institutions and market integration. The main debate, nowadays, is about the relative importance of these factors and, in particular, about geography versus institutions. The latter are generally believed to affect considerably the GDP per capita, while geography is sometimes perceived as a factor that do not have any further explanatory power beyond its effect on institutions. Following this approach (the Institutional hypothesis), the geographical variables only affect the making of good or bad institutions. This work, instead, argues that climate and health environment, as well as other geographical measures related to the integration in the world markets, have a direct effect on the level of income across countries. This evidence supports the Geography hypothesis and shows that, once geography is taken into account, the importance of institution is reduced. In a simple specification, with only institutions, they explain half of the variation between the four rich and the 24 Sub-Saharan countries of the sample. A more detailed model, that includes other geographical variables, reveals that institutions predict roughly one sixth of the cross country variation in per capita GDP. The importance of geographical factors is a little smaller or even larger when the health environment (malaria risk) is taken into account.

This work underlines how some empirical evidence in favour of the institutional hypothesis is strictly connected with the choice of a particular proxy for institutional quality. There are a lot of different measures of institutions and some of them do not confirm the irrelevance of the geographical variables as direct determinants of the level of income. I look at the model by Acemoglu, Johnson and Robinson [2001] and I find out that, changing the institutional measure, some geographical variables turn out to be significant in the level regression of GDP per capita. The presence of malaria, the coastal proximity, the natural resources availability and being landlocked are the variables that result to affect directly the economic development, other than institutions.

The first part briefly reviews the literature on this topic and presents some of the methodological problems that arise in this field: the endogeneity of the institutional variables and the subsequent choice of the instruments are the main econometric difficulties to be addressed.

The central part is about the choice and the availability of the data and presents the sample. There are many other possible institutional indicators, other than the ones used by Acemoglu, Johnson and Robinson. Furthermore, it seems that there are no sound reasons in favour of one rather than another: this is why this work addresses some robustness controls and this section reviews other possible institutional and geographical indicators, trying to identify their relationship with the level of income.

The last part presents the empirical results. I firstly replicate the results obtained by Acemoglu, Johnson and Robinson using updated data, and then I estimate the same model using different indicators for institutional quality and geography. I finally expand the analysis, estimating different models and using different instruments and variables. The possibility of having a degree of overidentification supports the goodness of the specification and malaria turns out to be another important determinant of the variability of income across countries. The preferred specification shows that both geography and institutions affect the level of per capita GDP: the health environment, which include a lot of climatic factors and being landlocked have a direct effect on income, beyond the effect that geography has on institutional quality.

Finally, the conclusions summarize the most important results and try to quantify the real impact of institutions on GDP per capita. A very simple experiment divides the sample in rich and Sub-Saharan countries in order to evaluate how much of the difference in income across countries is explained by institutions, climate and geography.

The list of all the variables is reported in the Appendix A. All the tables with the summary statistics and the regression outputs are in the Appendix B.

2. Review of the literature

Large part of the literature – see Easterly and Levine [2002] and Rodrick et al. [2002] – has identified three main factors as the determinants of the differences in the economic development across countries: institutions, geography and trade policies. Cultural and religious factors and the classical determinants of economic growth [Barro, 1997] are important as well, but recently the debate has focused on the relative importance of geography and institutions. Therefore, I will briefly present these two approaches, even if the other factors should not be neglected (see, in particular, the Frankel and Romer [1999] gravity model, Dollar and Kraay [2002] and Winters [2004] about the relevance of trade policies and integration).

2.1 The Institutional Hypothesis

This approach is strongly linked with the ideas of Douglas North [1990] about the difference between British and Spanish institutions, with the former that are believed to be more favorable to economic growth. This explains the relative success of the former British colonies in North America, with respect to the Latin American countries which were influenced by the Spanish and by other European institutions. The colonial legacy is embedded in the work of Acemoglu, Johnson and Robinson [2001: 1395] (thereafter AJR), who use data about the mortality of the European settlers to demonstrate that "differences in colonial experiences could be a source of exogenous differences in institutions". The conditions found by the early settlers affected their colonization policies. In other words, the natural conditions, the endowments and the diseases have an effect on the current level of GDP per capita only through the way in which they shaped the institutions. Geography does not have any other direct effect on the level of income. The same authors, in a recent paper [Acemoglu et al. 2004], restate the importance of institutions as a fundamental cause of long-run economic growth. One of the argument that Acemoglu et al. [2004: 18-20] use to support the predominance of economic institutions over geography and culture is related to the Korean experience: they treat Korea as a natural experiment and, because both North and South Korea share the same culture, the same disease environment and the same geography, they consider the difference in economic institutions (the North followed the Soviet socialism while the South maintained a system of private property) as the "primary factor shaping cross-country differences in economic prosperity".

The relevance of colonial legacy is confirmed also by Lange [2004], who underlines also the importance of the actual form of colonial rule: the disease environment and the precolonial population affected not only European settlement, but also the extent of indirect rule, indirectly shaping the postcolonial governance. Engerman and Sokoloff [2002] argue that the specific factor endowments that the colonialist found in the Americas determined the kind of institutions that originated: a country rich of natural resource that favors economies of scale is more likely to develop high level of inequality and an elite incline to rent-seeking (look at *encomiendas* in Latin American countries). On the other hand, countries like North America have endowments that encourage a more egalitarian distribution of land and, therefore, a less powerful elite and better institutions. Sokoloff and Engerman [2000] identify in the institutions the element that contributed to the persistence of the initial inequality, due to difference in factor endowments.

Easterly and Levine [2002] confirm that endowments explain both economic and institutional development, but natural conditions do not have any explanatory power on economic development other than their effect on institutions, which are the key determinant of income. Besides, Easterly and Levine do not find any positive relationship between macro policies and GDP, after controlling for endowments.

Rodrick, Subramanian and Trebbi [2002] (thereafter RST) control not only for geography but also for integration and find evidence that institutions are the real source of growth. Eventually, Hall and Jones [1999] argue that part of the cross-country differences in output per worker can be attributed to what they call social infrastructure: one of its component is related to institutional quality, that they instrument with the distance from equator and language spoken.

2.2 The Geographical Hypothesis

This second approach stresses the importance that many geographical factors have directly on the economic development. This theory is generally thought to derive from the 18th century contribution of Montesquieu, who proposed a climate theory of underdevelopment that was dismissed because of the racist interpretation that received. Nonetheless, in the last decade, many work by Sachs and other authors pointed out the importance of tropical and geographical variables in explaining the poor performance of some countries: the ones situated in the tropics are harmed by many disadvantages related to soil fertility, diseases, difficult access to markets, high presence of crop pests and parasites and water availability. McArthur and Sachs [2001: 4] review the results obtained by AJR and find that "*both institutions and geographically-related variables* (such as malaria incidence or other health indicators) play a role in determining GDP per capita".

A recent study [Hibbs and Olsson, 2004] demonstrates the importance of the biogeographic initial conditions 12,000 years ago – which allowed for the transition from hunting-gathering to agriculture – as a nearly ultimate source of contemporary prosperity. Even if institutional conditions are considered, biogeography and geography remain significant explanatory variables for the differences in the level of economic development across the world. Hibbs and Olsson [2004: 3715] argue that "the richer the biogeographic endowment in broad regions of the world, the earlier was the transition to settled agriculture and, thus, the earlier was the onset of accelerated technological development and economic growth"[1]. Sachs [2001: 28] underlines how some technologies in critical fields – agriculture, health and manufacturing – are ecologically specific and claims that "at the core of long-term economic growth has been the continued development of technology, a complex social activity that has benefited the temperate zones more than the tropics".

Gallup, Sachs and Mellinger [1999] state that geography plays a direct and fundamental role in economic productivity through four main channels: human health, agricultural productivity, physical location, and proximity and ownership of natural resources.

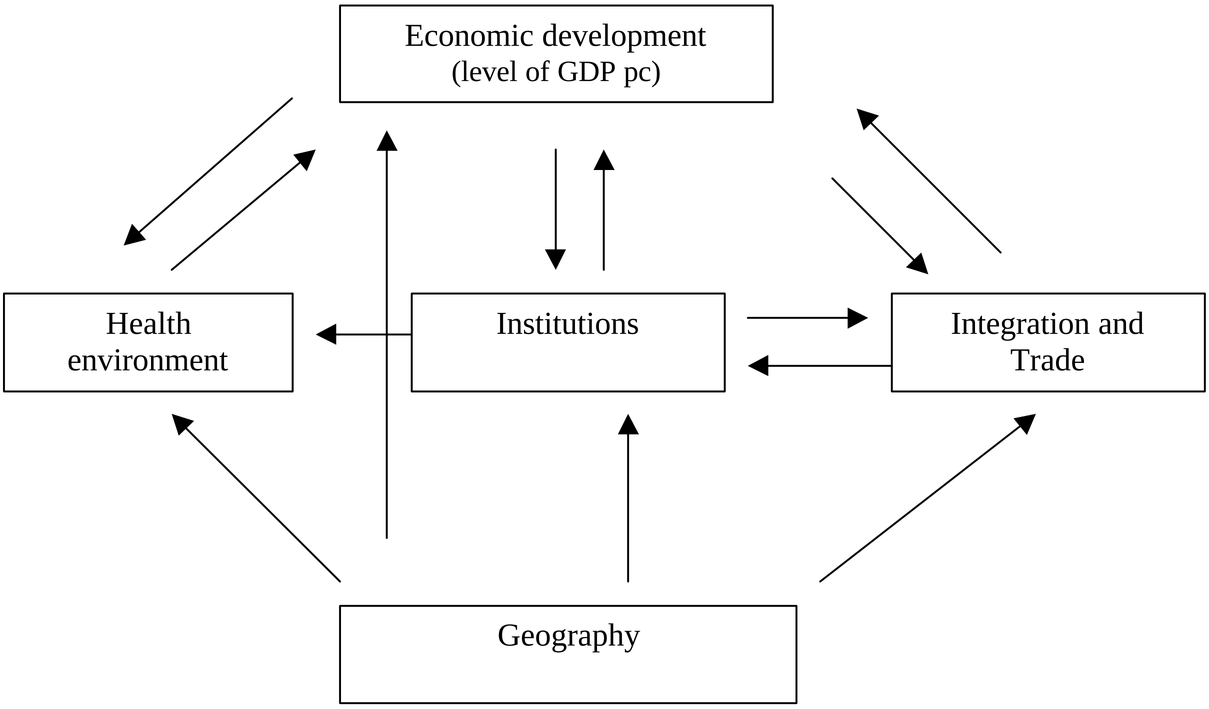
Although this evidence in favour of the importance of geographical factors, the institutional hypothesis seemed to be predominant and it was confirmed by the findings of Easterly and Levine [2002], AJR [2001] and RST [2002]. However, Sachs [2003] replies to those conclusions and reasserts the importance of geography. He argues that the health environment, represented by the ecological conditions which affect malaria transmission, is strongly related to poverty and to the differences in GDP across the world. The development of an exogenous index predictive of malaria transmission allow

---

[1] This is the view reflected by the work of Diamond [1999], who develops a long term view of the human development and brings geography and ecology into the analysis of this topic: the initial conditions, especially in terms of crops and animals, helped the Old World to generate surpluses. These allow for the development of more stratified societies, where the division of labour and the technological progress enhanced the process of economic development

for showing the strong and negative effect that malaria exerts on the level of income[2]. The attention given to the health environment and a more accurate treatment of this problem make the Geography hypothesis current again.

Figure 1: The determinant of economic development.



The *Figure 1* summarizes the variables that are generally considered the real determinants of the economic development. At the bottom, there is the only exogenous factor – geography – while in the middle there are the variables which are endogenous: they affect the degree of economic development but, at the same time, the relation of causality could go in the opposite direction. Richer economies could afford to have better institutions and they also tend to trade and be more linked to the rest of the world. Besides, Institution and Integration influence each other and the same problems of endogeneity is present whether one would control for macro policies. Malaria and health environment are no more treated as a simple exogenous factor. The presence of malaria risk can not be considered truly exogenous, because the level of wealth, as well

[2] For a more detailed explanation of the causality of the relationship between malaria and income and of the exogenous index (Malaria Ecology), see also Gallup and Sachs [2001] and

as the quality of institutions, certainly affect the possibility of improving the sanitary conditions and of reducing the diseases in a country. Some climatic and biologic variables are the exogenous determinants of the presence of malaria.

The Geography hypothesis stresses the direct effect that geography itself and also the health environment have on GDP while, following the institutional approach, the effect of geography is limited at the influence over institutions.

3. Methodological Problems

In order to identify the reasons that originate different level of economic development across different countries, one of the main problems is regressor endogeneity. The literature has identified some factors that can be considered exogenous – like the ones related to geography – and that do not present econometric difficulties, and other factors which, instead, are undoubtedly endogenous – the ones linked to the quality of institutions.

The presence of both these kind of explanatory variables, in a general cross-country linear regression, makes the OLS estimates biased and inconsistent. Because of these reasons, we need to use the Generalized Instrumental Variable Estimator (GIVE)[3] technique and the 2SLS procedure. Under certain assumptions[4], the GIVEs are generally consistent, but not unbiased and it is important to underline that this method presents some significant drawbacks.

The general shortcoming of this procedure is that the IV estimates have larger standard errors than the OLS ones, and the magnitude of this difference is inversely related to the degree of correlation between the instrumented variable and the instrument. More formally, it can be shown that the asymptotic variance of $\hat{a}_{GIVE}$ is equal to:

(1) $$VAR(b_{GIVE}) = s^{2}(X'P_{Z}X)^{-1}$$

Kiszewski et al. [2004].

[3] The Generalized IV estimator is used when the number of instruments (h) is greater than the number of endogenous variables (k). In the simplest case of exact identification h is equal to k, and this is the classical IV regression. However, overidentification is a desirable property for two reasons: the possibility of testing the validity of the instruments, and the achievement of good finite-sample properties.

[4] See Davidson and MacKinnon [1993: 216].

where ó^2 is the variance of the error term (estimated dividing the RSS by n instead than n – k), X is the k column matrix of explanatory variables, both endogenous and exogenous, P_Z is the projection matrix (equal to $Z(Z'Z)^{-1}Z'$), and Z is an h column matrix of instruments. All the exogenous variable of X should be also included in Z and the identification condition requires that h must be greater or equal to k. So, the IV variance is always larger than the OLS one, the difference depending on the correlation between the endogenous variable and the instrument. Reducing the sample variance requires an increase in the R-squared of the first stage (the reduced form) regression.

As explained by Bound, Jaeger and Baker [1995] (thereafter BJB), the IV estimates can lead to inconsistency and finite sample bias. They argue that the common way of searching candidate instruments that are weakly correlated with the endogenous variable leads not only to problems of large standard errors but could also originate large inconsistency in IV estimates. Moreover, BJB claim that, in finite sample, the â$_{IV}$ is biased in the same direction as â$_{OLS}$, as the R^2 of the reduced form equation tends to zero.

It is possible to show that even a small correlation between x and z can lead to inconsistency problems if the instrument is weak. Considering the simple case of one instrument, the inconsistency of IV relative to OLS can be expressed as:

(2)
$$\frac{p\lim b_{IV}-b}{p\lim b_{OLS}-b}=\frac{r_{z,v}/r_{x,v}}{r_{x,z}}$$

where ñ$_{i,j}$ is the correlation between i and j, and v is the error term of the structural equation. A weak instrument means that the denominator will be small, so that even a small correlation between the instrument z and the error term will originate larger inconsistency in the IV estimates than in the OLS ones.

The second shortcoming highlighted by BJB is the finite-sample bias. This arises because of the use of the estimated coefficients of the first stage regression, instead of the real ones. It can be showed that the direction of the bias is the same of the OLS case and its magnitude approaches the OLS bias as long as the R^2 of the first stage regression approaches zero. If the correlation between the instrument and the endogenous variable is weak enough, "even enormous samples do not eliminate the possibility of quantitatively important finite-sample biases" [BJB, 1995: 446]. One solution to this problem, as pointed out by Angrist and Krueger [2001], involves a fewer use of instruments because the bias is proportional to the degree of overidentification. They claim that, if the number of the instruments is equal to the one of the endogenous variables, the bias will be approximately zero. However, Buse [1992: 178] proves that the bias will increase with the number of instruments “only if the proportional increase in the instruments is faster than the rate of increase in R2 measured relative to the fit of Y on X” (the first stage equation, in Buse’s notation). Even if this cannot be determined a priori, a good rule of thumb is that adding further instrumental variables beyond the ones for which one has strong information about their explanatory power, is likely to add little to the R2 and to increase the bias. On the other hand, if one lacks of good a priori information, could use weak instruments: in this situation the addition of new instruments will increase the R2 and reduce the bias.

These findings should be considered in the selection of the instruments because it seems that this choice is not as simple as one could imagine. IV estimates are not unbiased and their consistency is not guaranteed if the instrument is weak enough: they are an asymptotic technique and applied works should be related with large sample. The two objectives that are desirable to achieve – estimates asymptotically as efficient as possible and a small finite-sample bias – are unfortunately contrasting. Asymptotical efficiency is generally improved increasing the number of instruments, but this is not generally a good idea when dealing with small sample. In this case, the more the instrument the more the bias of the IV estimates. Davidson and MacKinnon [1993: 220-4] show this trade-off: when the sample size is small, adding instruments increases the bias and make the distribution of $\hat{a}_{IV}$ closer to the one of $\hat{a}_{OLS}$. On the other hand, the simple IV estimator (h = k) is less biased but it is highly dispersed and it results to be very inefficient. However, increasing the sample size makes this problems negligible because as n increases both the bias and the variance of the IV estimator decrease. Staiger and Stock [1997] and Staiger [2002] underline the importance of the first-stage F-statistic, whose value should be at least greater than 10, as well as the number of instruments and the amount of the OLS bias as key parameters on which depend the IV properties. In fact, many instruments and/or a first stage F-statistic under 10 make the 2SLS confidence interval too short while the bias of the estimate towards the OLS is generally well approximated by 1/(first stage F). So it is important to not have too many instruments and to have high values of the first stage F, in order to avoid substantial biases in 2SLS estimates.

Another important result about 2SLS estimator is that its m^{th} moment exists only if $m < h - k + 1$. This means that the simple IV estimator ($h = k$) does not even have the mean and so its small-sample properties are very poor. It would be desirable to have at least $h - k = 2$ (the number of instruments should be equal to the number of explanatory variables plus two), in order to have an estimator with mean and variance [Davidson and MacKinnon: 221-2].

In the 2SLS framework is crucial to test the overidentifying restrictions (OIR). The number of OIR is just the difference between the instruments used and the endogenous variables and the null hypothesis is that all the zi are uncorrelated with the error term. If the null is not accepted, it means that at least one of the instruments is endogenous.

At the end, there is a sort of trade-off between using more instruments – and so increasing the efficiency of the 2SLS estimates – and the increase in the bias due to overidentification. BJB show that this is an issue even with very large sample: too many instrument are likely to bias the results. In this case, where the sample size is in a range between 60 and 150, this problem can not be ignored.

4. How to measure Institutions and Geography

One of the main difficulties in evaluating the relative importance of institutions and geography for the economic development relies in the measurement issue. AJR uses the risk of expropriation of private investment (the average value for the period 1985-95) as a proxy for institutional quality. Even if they take the risk of expropriation as preferred variable, they argue that their findings “hold also for a variety of other measures of institutions” [2001: 1378]. However, this claim is shown not to be true because, whether different institutional indicators are used instead of the risk of expropriation, some geographically-related variables turn out to have a significant effect on the level of GDP per capita. Geographical factors as well are difficult to capture in a single variable, thus some indicators could be better than others.

4.1 Institutions

Institutional quality can be measured in many different ways other than the risk of expropriation and the other indicators reported by AJR. A first general classification of all these measures separates them between indexes that describe the features of political and economic institutions and others concerned about the evaluation of their performance [Aron, 2000]. The second group is the more suitable for this work since it describes whether the existing institutional norms are implemented or not. However, all the variables used as proxies present advantages and drawbacks because they are generally focused on a particular aspect of governance. Aggregate index might solve this problem but, at the same time, it is not clear how to weight the single components to create a global institutional index.

Easterly and Levine [2002] use, actually, the average of the six measures of institutional development from "Governance Matters" by Kaufman, Kraay and Zoido-Lobaton, even if they assert that their findings hold for each single indicator. On the other hand, RST [2002] focus their work on one of those indicators, the rule of law index for 2001. A lot of institutions and research centres provide measures of institutional quality and it is very difficult to disentangle which are the best[5]. For my purposes, three datasets are good candidates because of the time extension of the data and thanks to their relevance acknowledged in most of the studies on this subject. These datasets are IRIS-3, which covers at least 20 years, Polity IV, a broader project that covers the institutional development from its origins in 1800 and the Economic Freedom index, provided by the Fraser Institute. The chosen datasets give the possibility of averaging the values over 15-20 years, at least, providing more stable and reliable indicators of institutional quality. Because of the difficulty of determining which single variable represents better the notion of institutions quality, I believe that the best way to address this problem should be a deep analysis, that should not be limited to test the institutional hypothesis using a single measure, together with some robustness checks,

---

[5] See Aron [2000] and the World Bank (http://www1.worldbank.org/publicsector/indicators.htm) for a detailed review of the empirical measures of institutions. The data from *Business Environmental Risk Intelligence (BERI)* and from *Business International* are based on a small sample, while other datasets – as *Heritage Foundation* and *Transparency International* – provide observation only for the last years. Eventually, the data from the *Freedom House* about political rights are available online only starting from 2000, even if the dataset starts from 1973 and covers 165 countries. A brief

made with a different specifications of the model. However, for reasons of space, I present the results obtained using the IRIS-3 dataset, even if the data from Polity IV confirms the geography hypothesis and allow for enlarging the sample size and for changing the instruments.

The IRIS-3 dataset is an updated version of IRIS-2 and its major changes concern the inclusion of data for 1996 and 1997, as well as an update of the data for the period 1982-1995 contained in the previous version. Moreover, the annual values are means calculated from all 12 monthly values for each indicator, while the annual values in IRIS-2 were calculated using only data for April and October. IRIS-3 contains five different indicators of quality of governments: (1) Corruption in government, (2) Rule of law, (3) Quality of bureaucracy, (4) Risk of repudiation of contracts by government, and (5) Risk of expropriation of private investment[6].

The AJR argument, in favour of the choice of the risk of expropriation, is related to the fact that "This measure is appropriate for our purposes since the focus here is on differences in institutions originating from different types of states and state policies" [2001: 1378]. So, the notion of extractive state should correspond to low values of the index. However, I think that also the other indicators could be a good proxy for institutional quality. AJR claim that they expect high values of the expropriation index to be associated with tradition of rule of law and well enforced property right. Therefore, also the other four indexes are good candidates to be indicators of institutional quality, especially rule of law and the risk of repudiation. All this variables are averaged over the period 1982-1997.

The summary statistics of these indicators – considering the mean for each country over the period considered in the dataset – are reported in the *Table 1* (The number of observations is referred to the number of observations that match with my data, so they are less than in the original datasets). The *Figure 2* shows the relationship between the logarithm of GDP pc and the average expropriation risk: it confirms the findings of AJR and it is very similar to their Figure 2 [Acemoglu et al., 2001: 1380], even if I use the full sample and not the small one considered by AJR, limited to the country for which there are data available on mortality rates[7]. The *Table 3*, which shows the correlation

illustration of the links between institutional quality (and its measurement problems) and development is online at: http://www1.worldbank.org/publicsector/legal/institutional.htm.

[6] For a more detailed description of these and of all the other variables, see the Appendix A.

[7] The same relationship is present also for the other institutional indicators.

between the logarithm of GDP and each of the indicators, confirms the strong positive relationship between institutions and GDP pc illustrated graphically. All the pairwise correlations are significant at 1% level of confidence.

Figure 2: Relationships between per capita GDP and different institutional indicators.

4.2 Geography

The bulk of AJR work is focused on the insignificance of the latitude variable, which measures the absolute distance from equator and which is used as a geography proxy every time they check for other control variables. It is possible to argue, however, that latitude is not necessarily a good indicator of geography, especially if the aim is to quantify how a place is unpleasant. The absolute distance from equator says something about the climate, even if it is not the most precise indicator, but much less about the endowments, the health environment and other characteristic that make a place more or less attractive for a potential settler: latitude alone is not a good indicator neither for the environmental and climatic features of a country, nor as a proxy of European institutions penetration.

Besides, AJR take the data from Parker [1997] about temperature and humidity variables. They consider the average temperature, the minimum and the maximum monthly high and monthly low temperature, and the maximum and minimum humidity in the morning and in the afternoon. As well as for latitude, even these indicators do not seem to catch properly the attractiveness of a country for a potential settler. Therefore, it is not surprisingly that these indicators are not significant. The measurement of the effect of humidity and temperature on economic development should be linked to the distance of the actual from an "ideal" value that fosters economic development and, moreover, humidity and temperature are linked together and their different combinations could originate different climates. Therefore, the available data about highest and lowest values are not sufficient for achieving a precise description of the environment and the attractiveness of a country.

As done for institutions, I try to find out more and better geography indicators in order to test the validity of AJR conclusions. Following Mellinger, Sachs and Gallup [1999], Gallup, Sachs and Mellinger [1999 thereafter GSM] and Sachs [1999, 2003], I use the data about the Köppen and Geiger climate zones, which report the percentage of land area or population (in 1995) that is located in a determined climate environment. Mellinger, Sachs and Gallup [1999] and Sachs [1999] argue that climate and coastal proximity are two key determinants of economic development. They find that temperate ecozones within 100 km of sea navigable waterway represent the 8% of world's landmass, they are inhabited by the 22% of world's population, but they produce more than 50% of economic output.

The main variables from Köppen and Geiger classification are (1) the percentage of land area in tropics, in tropics and subtropics, and in temperate zones; (2) the percentage of population in 1995 living in tropics, in tropics and subtropics, and in temperate zones. Both the population and the land indicators are far from being perfect, even if they are an improvement with respect to the distance from equator. The population percentage might be partially endogenous, at least at some degree, because the level of economic development could affect the decision of settling, but, on the other hand, it has the advantage of telling something about the attractiveness of that country, which is exactly what is needed for this work. Even the percentage of land in temperate zones – although truly exogenous – is not a good measure of geography for our purposes because quite all the economic activity of a country could be located in the small and attractive fraction of land (think at the Scandinavian countries).

Other indicators from the GSM dataset [1999] are related to the proportion of a country's population or land area within 100 km of the coastline or a navigable river; the coastal density in 1965 and 1995; the distance from the closest major market; the landlocked dummies – all of them stress the importance of geography thought as possibility of being connected to the markets, not as in the climate approach – and the proportion of country's land area (and population in 1995) within the geographical tropics.

Another part of the geographical literature stresses the importance of health environment[8] which is generally quantified by a variable that measures the percentage of population at risk of malaria transmission. A large part of the empirical works[9] uses indicators of malaria which are likely to be endogenous and subject to measurement errors. The first problem is due to the fact that the level of economic development has an effect on the health environment, through an improvement of general sanitary conditions. The second point is linked mostly to the fact that the percentage of population affected by malaria is a poor indicator because the number of cases officially reported are only a tiny fraction of the real number of infections. Sachs [2003: 5-7] provides an index – Malaria Ecology (ME) – which is defined as exogenous because it is "an ecologically-based variable that is predictive of malaria risk [… that] is therefore exogenous to public health intervention and economic conditions") and provides an instrument for malaria risk[10].

Because of their large use in the literature, I look also at other malaria indicators: the Falciparum Malaria Index in 1966 and in 1994, the proportion of population in 1995 living in area with malaria and malfal, which is the previous index multiplied "by an estimate of the proportion of national malaria cases that involve the fatal species, Plasmodium falciparum" [Sachs, 2003: 5]. However, ME seems to be a more complete

[8] Gallup and Sachs [2001] underline the strong connection between poverty and malaria and show that malaria is negatively associated with income levels, after controlling for geography, history and policy.

[9] See the works of Acemoglu, Johnson and Robinson, RST [2002], and GSM [1999].

[10] More precisely, this is a spatial index of the stability of malaria transmission, which is based on "published reports of anopheline bionomics, vegetation maps (defining suitable, unsuitable habitat), altitude (maxima or minima), monthly precipitation threshold (minima), and monthly temperature thresholds (minima, isotherms, length of frost-free season)" [Kiszewski et al., 2004: 487].

measure, since it takes into account also climatic and geographical aspects and it has the great advantage of being exogenous.

In the *Table 3*, I have reported the correlation coefficients between the log of GDP pc and many geographical indicators. The variables that I take into account as geography indicators are the proportion of land area (and population) in temperate zones, the tropics definitions from GSM, the landlocked dummies, the Malaria indexes, the coastal proximity indicators, and the distance from the closest port, whose statistics are described in *Table 4*. Eventually, another variable could capture the effect of the wealth of natural resources on the level of development. The log of hydrocarbons per capita in 1993 is taken from the GSM [1999] dataset and it is positively correlated with the logarithm of per capita GDP (the correlation coefficient is 0.30, significant at 5% level of confidence). Its summary statistics are in *Table 2* as well.

5. Replication of the results by Acemoglu et al.

As first thing, I replicate the findings of AJR [2001] using updated data to check whether their conclusions are confirmed. AJR want to estimate the level of GDP per capita as a function of Institutions and Geography. I look in particular at the very simple results of Table 4 [AJR. 2001: 1386], which are the core of their work and support the institutional view.

They use the logarithm of per capita GDP, at Purchasing Power Parity, in 1995, from the World Bank WDI. The most recent data now are from WDI 2003 and are referred to 2001. As proxy for institutions I use the risk of expropriation from IRIS-3, which is the update of the variable used in the original work. The instrument, the logarithm of mortality rates of early settlers (logM), is directly from the paper of AJR (Table A2, p. 1398). The other variable – latitude (absolute distance from equator) – is the same.

The formal specification of the model is based on a first stage regression in which the institutional variable is regressed on the instruments and on latitude, when it is considered as additional explanatory variable:

(3) $$INST = a + b\log M + cLAT + u$$

and on a second stage equation given by:

(4) $$\log GDP = a + bINST + dLAT + e$$

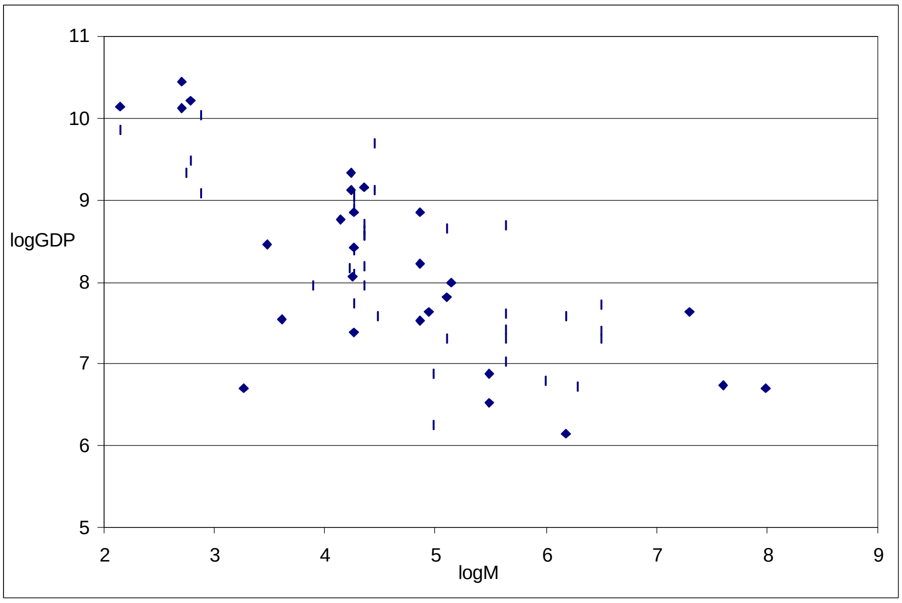
in which INST are the fitted value from the first stage regression.

The data confirm the strong relationship between logGDP and logM, as clear from the *Figure 3*, which replicates the Figure 1 in AJR [2001: 1371]

The *Tables 5* and *6* report the original estimates and, for each of them, the updated results both for the small sample (the one for which there are the data about settler mortality) and the full sample. The correlation between LogM and the risk of expropriation is equal to –0.58, which seems to avoid the problem of weak instruments and of inconsistency, because the denominator of (2) is not so small to cause larger inconsistency in the IV than in the OLS estimates. The R-squared from the first stage equation is 0.30, not too small, and also the F-statistic does not provide any evidence of weak instrument problems.

The new estimates confirm the results obtained by AJR: in particular, there is a clear difference between the OLS and IV estimates and latitude is not significant and enters in the regression with the “wrong sign”. These results are strongly in favour of the institutional hypothesis: geography matters only through its effect on institutions. More precisely, the 2SLS coefficient of the risk of expropriation means that a one point increase in the index is translated in a raise in GDP per capita of roughly 170%, on average. If I compare the four rich countries in the sample (Australia, New Zealand, US and Canada) with the 24 SSA countries using their mean GDP and their mean Institutions Index, it turns out that the difference of 3.86 points in the risk of expropriation index is consistent with a GDP per capita in the rich countries 6.4 times greater than in Sub-Saharan Africa. The real data, instead, say that the mean rich country is quite 13 times richer than the SSA mean country, in terms of GDP per capita. These results support the idea that institutions play a big role in economic development (they count for half the difference between rich and Sub-Saharan countries), even if a lot of the variation in GDP per capita still remains unexplained. AJR argue for the insignificance of latitude in their specification, however other more precise indicators could show a significant effect and explicate part of the variation in GDP per capita that still remains unexplained.

Figure 3: Relationship between income and mortality rates.



Moreover, the sample size is quite small, limited to the 60 former colonies for which there are data on mortality rates. The standard errors are quite large and so also the confidence intervals are pretty wide (the coefficient on expropriation risk varies in the range [0.63 – 1.33] for the second column specification in *Table 6*). Other specifications and further instruments could provide smaller sample variances and an increased R^2 for the first stage regression.

6. Different institutional and geographical indicators

Having defined alternative measures of institutions and geography, I test the robustness of AJR finding keeping on using the mortality rates of early settlers as instrument for institutional quality. The aim of this exercise is to find out if geography – whatever specified – affect or not the level of per capita GDP, beyond its effect on institutions. Moreover, different definitions of institutional quality are used in order to check the robustness of the results.

Formally, the first stage and the second stage regressions are similar to (3) and (4), even if LAT is substituted by GEO, another exogenous measure of geography. The mortality rates capture the effect of geography on institutions, while the other geographical variables are added for checking if some climatic factor could affect the GDP per capita as well in the second stage regression. The second stage equation is:

(5) $$LGDP = a + bINST + gGEO + e$$

where LGDP is the logarithm of per capita GDP in 2001. The coefficients to look at are â and ã, and the hypothesis to test is H_0: ã = 0 against the alternative H_A: ã ≠ 0 that means that geography has a direct impact on the level of GDP pc. â measures, instead, the impact that geography – measured as mortality rates of early settler – has on GDP pc, through its effect on the development of institutions (indirect effect). Also the R^2 from the first stage regression is of interest (it has not to be too small in order to avoid the problems related to weak instruments), as well as the F-statistic that should not be less than 10.

- Risk of expropriation as institutional proxy

The results in *Table 7a-b*[11] confirm the institutional hypothesis. Using the risk of expropriation as a proxy for institutional quality, all the geography indicators except one (I controlled also for the significance of the SSA dummy) are not significant in the second stage regression. Moreover, almost all the different geographical variables are insignificant also in the first stage regression, except the percentage of land area and population in temperate zone, the proportion of population living in tropics and within 100 km from the coast or a navigable river, and the dummy for landlocked countries (not W&C Europe)[12]. However, one indicator – malfal – is significant and has a strong negative effect on log GDP per capita. While risk of expropriation has generally a coefficient varying between 0.7 and 1.17 (in the most cases close to 1), adding malfal as explanatory variable reduces the effect of institutions to 0.54 (this result, however, is quite weak since malfal is not a good indicator of malaria risk, because of its endogeneity). An increase of one point in the expropriation risk is generally translated in an increase of one point in the logarithm of GDP, which means that GDP pc will rise

[11] I have not reported the results from the landlocked dummy and the falciparum malaria index in 1966 and 1994: they are not significant both in the first and second stage regressions.

[12] These variables are therefore candidates as instrument for the risk of expropriation. A first stage regression in which these variables are added altogether with LM (using once the percentage of population and once the percentage of land area in the temperate zones because of their high correlation) shows that only the KG climate zones variable and LM remain significant. This is the reason why, in the next sections, the ecozones variables are taken as additional instruments.

proportionally by 1.71 (i.e. by 171%), on average and ceteris paribus. Now, the effect is much more smaller because GDP per capita will rise only by 0.71.

- Other institutional proxies

However, the other indicators of institutional quality do not confirm the institutional approach. Taking the variables listed in *Table 1* as proxies for institution instead of risk of expropriation makes some geography indicator significant. More precisely, repudiation of government contract is the best candidate, confirming the impression given by the correlations showed in *Table 2*. Without any geographical control – using a specification equal to (1) in *Table 7* – the use of repudiation of government contract (thereafter, repcontract) as institutional quality indicator improves considerably the fit of the first stage regression and reduces the point estimate in the second stage equation. Furthermore, the confidence interval is narrower and the estimates result more efficient (see *Tables 8*). Besides, the effect of institutions on the level of per capita GDP is smaller when repcontract is used as institutional quality indicator. Comparing the basic specifications (1) in Tables 7a and 8a, one can see that institutions, measured as risk of expropriation, explain half of the difference in GDP pc between the four rich countries and the SSA countries in the sample – that is almost 13 to 1. On the other hand, the repudiation of contract measure predicts only a GDP per capita in the rich countries 4.6 times higher than in SSA: the effect of institutions in explaining the variation of GDP across countries is pretty reduced just changing the way for measuring institutional quality.

Also all the other institutional indicators are significant, even if the R^2 from the first stage regression is lower than using risk of expropriation or repcontract. For this reason, I will show only the results using repcontract as institutional quality measure.

From the *Tables 8 a-b* it is clear that the statement about the irrelevance of geography as causal factor of economic development is weak: a different way of measuring institutional quality leave space for geographical factors to enter significantly as determinants of per capita GDP. The percentage of land within 100km from the coast (or river) is significant at 10% level of confidence, as well as the proportion of population living close to the coast and in the tropics; the Malaria Ecology index (and three of the four malaria indicators, not shown) and the SSA dummy are instead strongly significant. The other geographical variables, instead, do not have any further explanatory power even with different definitions of institutions.

Very similar results are obtained even using the quality of bureaucracy index. ME, as well as the other malaria indicators and the SSA dummy, remains significant using all the other IRIS-3 indicators as proxy for institutional quality. They have a negative and a strong effect on the level of per capita GDP. However, the R-squared from the first stage regressions are quite smaller whether one uses rule of law, corruption and, even if the effect in this case is less important, bureaucratic quality (data not shown for reasons of space).

- OLS estimates

The OLS estimates (*Tables B1* and *B2*) confirms all the expectations about the necessity of using the IV technique. The point OLS estimates on the institutions variables, whether represented by risk of expropriation or by risk of repudiation of contracts, are much smaller than the 2SLS estimates, while the standard errors in the latter case are larger. The coefficient of institutions on GDP per capita is generally two folds larger than suggested by OLS estimates. For the simplest case of specification (1) in *Table 7a*, an increase of one point in the expropriation risk index is translated in an increase of 68% of GDP pc using the OLS estimate, while the effect computed by IV implies an increase of about 166%, on average. Furthermore, the OLS specification supports the geographical hypothesis, because almost all the geographical variables included in the regression are significant and have the right sign.

- The effect of institutions and geography on GDP per capita

The relevance of the malaria indicators and SSA dummy, in most of the IV specifications, suggests the relevance of health environment and of the speciality of Sub-Saharan Africa. Looking at the results from *Table 8a*, where Malaria Ecology and the Africa dummy are significant, it is possible to see how these two variables reduce the effect of institutions on GDP. In the basic model (1) a one point increase in the repudiation of contracts (repcontract) score will be translated in an increase of per capita GDP of 112%. However, if ME is added in the model, the effect of one point increase in the repcontract score on GDP pc becomes 70%. Now institutions alone are able to explain a difference in per capita GDP between rich and African countries in the sample of only 2.88 times, instead of the real difference of 13 times. The inclusion of the SSA dummy has a smaller effect on the institutional coefficient, that becomes 0.61. Now if the repcontract index raises by one point, the GDP per capita will increase by 84%. Furthermore, being a Sub-Saharan country means that the GDP per capita is, on average, 60% smaller than in the other countries. The two geographical variables are positively correlated (ñ = 0.69) and their effect vanishes if they are added simultaneously in the regression (data not shown). This result substantiates the intuition that Malaria is geographically specific and that it is a problem especially in SSA [Gallup and Sachs, 2001]. To evaluate better the impact of the presence of malaria on economic development, however, ME is not the best measure at all: it is quite abstract, while a more common index of the proportion of population living in areas with malaria could provide a better insight, even if it presents the endogeneity problem (see next section).

The significance of the indicator of proportion of land close to the coast or to a river, re-enforces the idea that institutions are not the only explanation of GDP pc. The results from the column (6) in the *Table 8a* show that a country with 10% more of land within 100km the coast has a per capita GDP 5.5% higher that the mean country. These results also say that a country with a complete access to the sea (all the land is close to the sea) has, ceteris paribus, a per capita GDP which is 71% higher than the one of a state that does not have any part of its land close to the sea. The importance of having access to the sea or to a river is a geographical indicator that reflects, at some extent, the importance of openness and trade for economic development.

From this first analysis, it seems quite clear that looking just at institutions is not sufficient to determine the level of economic development of a country: geography matters in a sense close to health environment in four out of five IRIS-3 indicators of institutional quality, as well as the SSA dummy, which captures the particularity of the African continent that is not explained simply by the institutional quality and, even for not all the specifications, the measure of coastal proximity.

However, this simple specification, with exact identification and with mortality rate as instrument, seems not to be a very powerful tool to understand the relative effect of institutions and geography on economic development. In the next sections, I will focus on these two aspects: the research of a better set of instruments and the overidentification tests.

- <u>Overidentification</u>

Until now, I have limited the analysis at the case of exact identification: one endogenous variable and one instrument. However, overidentification is a desirable property because it increases the efficiency of the estimates and the finite sample property of the 2SLS estimator and it allows also for overidentifying restriction tests (OIR). To define a better specification, I need other instruments in order to improve the fitness of the first stage regression. Following Sachs [2003], I take one of the climate zones from the Köppen and Geiger classification as additional instrument[13]: hence, the degree of overidentification is one, which guarantees that the IV estimator has a mean, even if it has not a variance. As institutional quality indicator, I keep on using the repudiation of contracts by government, the index of bureaucratic quality and the risk of expropriation. As additional geographical variables, I use the coastal proximity and the SSA dummy which, in the previous models, were significantly correlated with GDP pc. The results from these regressions are reported in the following *Tables 9a, b and c*, in which there are only the results of the second stage regression.

As expected, the point estimates are smaller than in the exact identification case: adding an instrument has the effect of biasing the 2SLS estimates toward the OLS ones. Furthermore, the IV estimates are more efficient and the goodness of fit in the first stage regression is definitely improved – the R-squared is 0.51 instead of 0.39, when repcontract is used as institutional proxy. In the basic formulation (1), a one point increase in the institutional score will raise the GDP per capita by 104%, on average.

The effect of institutions varies very little when using the proportion of population or of land area in temperate zones, suggesting that the choice between the two instruments does not have any practical effect. In the model (3) a one point increase in repcontract origins a 99% increase in per capita GDP, while a 10% increase in the proportion of land area within 100km the coast is translated in a 5.6% raise in per capita GDP, on average and ceteris paribus. A country with all the land area close to the sea, has a GDP pc that is 0.73 times higher that a landlocked state, other things equal. Being

[13] As seen before (Table 5 and 6), the KG ecozones variables enter significantly in the first stage regression, suggesting that they satisfy the condition of being partially correlated with the endogenous variable. Here again the choice between the proportion of population and the percentage of land area in the temperate zones is without a clear solution, so I look at both cases.

a SSA country has still a strong negative effect on per capita GDP, that is between 57% and 63% smaller than the GDP pc of the average country (see column (11) and (12)).

The same conclusions are roughly valid even considering the *Tables 9b and c*, where the risk of expropriation and the bureaucracy index are used. Furthermore, these results confirm the findings of the previous section about the lack of robustness of the institutional hypothesis. All the specifications pass the OIR test, giving further confidence about the validity of the instruments considered.

7. A more defined model: some results

As already stated, a more correct consideration of the effect of malaria should take into account the endogeneity of the indicators considered – Falciparum malaria index in 1994 and the percentage of 1995 population living in area with malaria. The basic formulation is given by this structural equation:

$$LGDP = a + bINST + gMAL + dGEO + e \quad (6)$$

where INST is the risk of repudiation of contracts by government, MAL is one of the two malaria indexes and GEO is an additional exogenous geographical variable (in the basic specifications it is not included, while in other cases the additional geographical controls could be more than one). Because both INST and MAL are endogenous, the two first stage regressions are:

$$INST = a_1 + b_1 LM + c_1 KG + d_1 ME + f_1 GEO + u_1 \quad (7)$$

$$MAL = a_2 + b_2 LM + c_2 KG + d_2 ME + f_2 GEO + u_2 \quad (8)$$

where LM is the log of mortality rates, KG is one of the two (land area or population in temperate zones) Köppen and Geiger ecozones and ME is the Malaria Ecology index. The results in the *Table 10* show that the choice between the two ecozones indicators does not affect the estimates and the goodness of fit of the first stage equations. The inclusion of malaria originates a very sharp reduction in the coefficient on repcontract, which is halved in the simple formulations (1) and (2), with respect to the previous models. The coefficient on institutions is pretty stable across the different specifications, even if it increases when Falciparum malaria index is used instead of the percentage of population living in area subject to malaria risk. Besides, its magnitude increases slightly when the landlocked dummy is included, because of the negative effect of the latter on GDP. One of the main results concerns the not significance of the SSA dummy, whose effect is likely embedded by the malaria indicator. Looking at the columns (3) and (4), a raise in the repcontract index by one point means that GDP pc will increase by 48% (on average and ceteris paribus); a decrease of 10 percentage point in the proportion of population subject to malaria risk will raise per capita GDP by 11% (on average and ceteris paribus); and being landlocked (not considering Western and Central Europe) means having a GDP pc smaller by 61% with respect to a country with access to the sea. The effect of not having access to the sea is very strong and it remains significant even using the other malaria indicator (the p-value is 0.08 and 0.09 respectively for column (9) and (10)).

These findings seems to be more interesting than the previous ones because they drop the significance of the Sub-Saharan Africa dummy: there is no more a need for a special explanation about the Africa's lack of development: the health environment and the geographical conditions, through institutions and also on their own are able to explain part of the variation of per capita GDP across countries. Being an African country do not add anything to the explanatory power of the model. The other geographical variables turn to be not significant. This means that, except that for being landlocked, that embeds the idea of integration in the trades, all the other geographical factors related to the climate and to the attractiveness of a country are likely to be already reflected in the institutions quality and in the health environment[14].

If one wants to evaluate how much institutions and geography contribute to the real differences in GDP across countries, the results from column (3) and (4) say that the difference between rich and SSA countries – whose per capita GDP ratio is 13 to 1 – is due in part to institutions, which count for a two folds difference; to the health environment, because the malaria indicator explain a difference in GDP pc of 1.6; while being landlocked is responsible only for a 7% differential in GDP pc, as expected because only 3 out of 24 SSA countries in the sample are landlocked.

This is just a very simple calculation of the relative magnitude of geography and institutions in the process of development. Even if the choice of looking at two sample is arbitrary, these numbers provide evidence of the importance of institutions but, at the same time, confirm that geography matters beyond its effect on government and underline that a lot of the variability in GDP remains unexplained. Furthermore, the magnitude of the institutional effect is now mitigated with respect to the AJR original model, in which institutions explain half of the income difference between rich and SSA countries.

One of the drawbacks of the model is the small sample size – roughly 60 observations of former colonies. As seen, one way for increasing the properties of the 2SLS estimator is to increase the sample size. This could be done dropping the mortality rates of early settlers as instrument. Instead of LM, I take a variable from the GSM dataset, newstate, that embraces the timing of national independence, and I keep on using the percentage of population living in temperate zones and the Malaria Ecology index. The degree of overidentification is still one. The first stage equations, as well as the OIR tests, do not present problems, and the number of observation now is 106 or 105. The state independence variable enters significantly in both the first stage regressions[15].

The risk of repudiation and the malaria indicator are still significant at 5% level of confidence and their magnitude is pretty similar to the one derived using the small sample. With this larger sample and with a different set of instruments, the indicators of coastal proximity, the average temperature, the log of hydrocarbons per capita and the landlocked dummy, taken singularly, are always significant, at least at 10% level of confidence (*Table 11*). If they are altogether included in the regression, only the landlocked dummy turn to be significant. Being landlocked means having a per capita GDP 77% smaller than the one of a country that has access to the sea, while the magnitude of the coastal proximity indicators is quite halved with respect to the previous specification (see *Table 9a*). The last two column (8) and (9) shows that the log of hydrocarbons per capita – an indicator of natural resources wealth – has a positive and significant effect on per capita GDP, even if it is considered together with the proportion of population living close to the sea. In the last specification the coastal

---

[14] This is not surprising at all because, as seen in the previous section about the data, Malaria Ecology already takes into account many geographical variables – humidity, altitude, temperature – which now do not have any further explanatory power.

[15] Data not shown for reason of space. Even if the specification seems to perform quite well, some doubts about the exogeneity of the independency variable remain. McArthur and Sachs [2001] consider newstate a candidate for being a valid instrument, but they acknowledge that it is endogenous to some extent, even if the correlation with the error term is likely to be small.

proximity counts for a 11% differential in GDP per capita between rich and SSA countries in the sample, as well as the natural resources wealth. Much more importance have the health environment – responsible for a difference of 2.13 times – and the institutions, which count for 1.61. Therefore, these last results claim that geography has a stronger impact on GDP pc directly (intended as health environment, as well as integration) than indirectly (thought its effect on institutions).

The main conclusion that could be drawn from this exercise is that the malaria indicator is an explanatory variable as well as institutions: across different formulations, malaria remains significant and negative correlated to GDP per capita. Moreover, geography seems to affect economic development not only through institutions and health environment, but also on its own. In particular, measures related to integration in the world markets and natural resources availability turn to be significantly related to GDP pc.

8. Concluding remarks

The search for the real sources of long-run economic development has identified institutions, geography and trade and macro policies as the main important factors that affect the differences in the current level of income across countries. Different points of view stress the relative importance of a component with respect to the others. In particular, the Institutional hypothesis asserts that geography and climate has no other effect on the level of GDP per capita beyond their influence over the shape of the current institutions. Nonetheless, different authors have acknowledged the relevance of climate and geography, which affect the level of development through many channels. The health environment and diseases; the soil quality and the presence of crop pests and parasites; the possibility of having access to the main markets and of trading with other countries, are some ways in which the natural location of a country can distress its development.

The findings of this simple exercise support the Geography hypothesis, because the health environment, the coastal proximity, the accessibility to the markets and the natural resources have some explanatory power in the regression of the current level of GDP per capita across countries. Geographic factors, in other words, exhibit a direct influence on the economic development, beyond their effect on institutions.

This work does not confirm the great reliance on the relevance of institutions as the only factor that matters: the results obtained by AJR about the primacy of institutions seem to be in some way influenced by a particular choice of the variable used as a measure of institutions and geography. It is true that, whether institutions are measured using the risk of expropriation index, geography results to be not significant. Nonetheless, different institutional indicators tell another story: whether repudiation of contracts by government is used as a proxy for institutions quality, i.e., the magnitude of the effect is reduced and, more important, geography is now a significant determinant of the variation in GDP across countries. The choice of a measure instead of another is quite difficult and, hence, many of them seem to have the same right to be considered. Following this way of thinking, the institutional approach is not very robust, because coastal proximity, diseases and market accessibility indicators enter significantly in a level regression of per capita GDP.

Moreover, more accurate models, that address also the problem of the endogeneity of the health environment, show that the effect of institutions is further reduced and that the risk of malaria transmission is another important obstacle to the economic development, as well as the being landlocked. The results of the small sample, due to the necessity of using the mortality rates instrument, are confirmed by different specifications of the model that enlarge the sample size to more than 100 countries.

Whether one wants to analyze how much of the differences in per capita GDP is explained by institutions and geography, the data show that the basic model (that support the institutional view) states that half of the difference between the rich and the Sub-Saharan countries is explained by difference in institutions quality. On the contrary, the preferred specification (see *Table 10*) reduces the magnitude of the institutions effect to less than one sixth, a little more than the effect of health environment. Other specifications are consistent even with a major effect of geography than of institutions.

However, a large part of difference in income across countries still remains unexplained, suggesting that these models need to be improved. In particular, the analysis has highlighted many difficulties concerning the way in which the institutions and the attractiveness of a country could be measured. Also the empirical part, related to 2SLS and IV procedure, is not without problems, that could undermine the estimates. The identification of the causal relationship and the choice of the instruments are difficult to address, and the dependence on the mortality rates presents some drawbacks related to the small sample size.

Eventually, there is space to deepen the analysis of this issue, that seems far from having reached a firm conclusion. Better indicators of geography and institutions, as well as better instruments, are necessary in order to confirm these findings, which, however, seem to be consistent with a theoretical model that can not leave out of consideration the geographic factors.

RIFERIMENTI BIBLIOGRAFICI

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). "The Colonial Origins of Comparative Development: an Empirical Investigation", American Economic Review, Vol. 91, No. 5, pp. 1369-1401, December 2001.

Acemoglu, D., S. Johnson, and J. A. Robinson (2004). "Institutions as the Fundamental Cause of Long-Run Growth", NBER Working Paper, No. 10481, May 2004.

Angrist, J. D., and A. B. Krueger (2001). "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments", Journal of Economic Perspectives, Vol. 15, No. 4, pp. 69-85, Fall 2001.

Aron, J. (2000). "Growth and Institutions: A Review of the Evidence", The World Bank Research Observer, Vol. 15, No. 1, pp. 99-135, February 2000.

Barro, R. J. (1997). Determinants of Economic Growth – A Cross-Country Empirical Study, The MIT Press.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak", Journal of the American Statistical Association, Vol. 90, No. 430, pp.443-450, June 1995.

Buse, A. (1992). "The Bias of Instrumental Variable Estimators", Econometrica, Vol. 60, No. 1, pp. 173-180, June 1992.

Davidson, R., and J. G. MacKinnon (1993). Estimation and Inference in Econometrics, Oxford University Press.

Diamond, J. M. (1997). Guns, germs and steel: The fate of human societies, W.W. Norton & Co., New York.

Dollar, D., and A. Kraay (2002). "Institutions, Trade and Growth", Paper prepared for the Carnagie-Rochester Conference Series on Public Policy.

Easterly, W., and R. Levine (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions", The Quarterly Journal of Economics, Vol. 112, No. 4.

Easterly, W., and R. Levine (2002). "Tropics, Germs and Crops: How Endowments Influence Economic Development", Center for Global Development Working Paper, No. 15, October 2002.

Engerman, S. L., and K. L. Sokoloff (2002). "Factor Endowments, Inequality, and Path of Development among New World Economies", NBER Working Paper, No. 9259, October 2002.

Englebert, P. (2000). "Solving the Mystery of AFRICA Dummy", World Development, Vol. 28, No. 10, pp. 1821-1835, 2000.

Frankel, J., and D. Romer (1999). "Does Trade cause growth?", American Economic Review, Vol. 89, pp. 379-399.

Gallup, J. L., and J. D. Sachs (2001). "The Economic Burden of Malaria", The Supplement to The American Journal of Tropical Medicine & Hygiene, Vol. 64, No. 1, pp. 85-96, January/February 2001.

Gallup, J. L., J. D. Sachs, and A. Mellinger (1999). "Geography and Economic Development", CID Working Paper, No. 1, March 1999.

Hall, R. E., and C. I. Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others", Quarterly Journal of Economics, Vol. 114, No. 1, pp. 83-116, February 1999.

Hibbs, D. A., and O. Olsson (2004). "Geography, biogeography, and why some countries are rich and others are poor", PNAS, Vol. 101, No. 10, pp. 3715-3720, 9 March 2004.

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

Lange, M. K. (2004). "British Colonial Legacies and Political Development", World Development, Vol.32, No. 6, pp. 905-22, 2004.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny (1999). "The Quality of Government", Journal of Law, Economics, and Organization, Vol. 15, No. 1, pp. 222-79, April 1999.

Mellinger, A. D., J. D. Sachs, and J. L. Gallup (1999). "Climate, Water Navigability and Economic Development", CID Working Paper, No. 24, September 1999.

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

Parker, P. M. (1997). National cultures of the world: A statistical reference, Cross-Cultural Statistical Encyclopedia of the World, Vol. 4, Greenwood Press, 1997.

Rodrick, D., A. Subramanian, and F. Trebbi (2002). "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development", CID Working Paper, No. 97, October 2002, Harvard University.

Sachs, J. D. (2001). “Tropical Underdevelopment”, NBER Working Paper, No. 8119, February 2001.

Sachs, J. D. (2003). “Institutions Don’t Rule: Direct Effects of Geography on Per Capita Income”, NBER Working Paper, No. 9490, February 2003.

Sachs, J. D., and A. M. Warner (1997). “Sources of Slow Growth in African Economies”, Journal of African Economies, Vol. 6, No. 3.

Sachs, J. D., and J. W. McArthur (2001). “Institution and Geography: Comments on Acemoglu, Johnson and Robinson (2000)”, NBER Working Paper, No. 8114, February 2001.

Sokoloff, K. L., and S. L. Engerman (2000). “History Lessons: Institutions, Factor Endowments, and Path of Development in the New World”, Journal of Economic Perspectives, Vol. 14, No. 3, pp. 217-32, Summer 2000.

Staiger, D. O. (2002), “Instrumental Variables”, Handout, Academy for Health Services Research and Health Policy, March 2002.

Staiger, D., and J. H. Stock (1997). “Instrumental Variable Regression with weak instruments”, Econometrica, Vol. 65, No. 3, pp. 557-586, May 1997.

Winters, L. A. (2004). “Trade Liberalization and Economic Performance: An Overview”, The Economic Journal, Vol. 114, February 2004.

APPENDIX A: LIST OF VARIABLES

Log GDP: It is the logarithm of GDP pc, measured in Purchasing Power Parity, for the 2001. Source: World Bank, WDI 2003.

Population: Total population in millions. Source: Englebert [2000].

Log Mortality rates: It is the logarithm of the European settler mortality rates. Source: Acemoglu et al. [2001].

Log of hydrocarbons per capita: It is the logarithm of hydrocarbon deposits per person. Source: Gallup et al. [1999].

Legitimacy dummy: It is a "dummy variable [that] takes the value 1 when a state is historically determined, 0 otherwise. It tries to capture either the historical continuity of state institutions (noncolonized areas), the embeddedness of the post colonial state into precolonial relations of authority" [Englebert, 2000: 1827]. Source: Englebert [2000].

Elf. Ethno-linguistic fragmentation. Source: Easterly and Levine [1997].

SSA dummy: Takes value 1 if the country is in Sub-Saharan Africa. It is the same for the other continental dummies. Source: Gallup et al. [1999].

Ethnic tensions: Ranges in value from 0-6, with higher values indicating fewer ethnic tensions. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

Life expectancy: It is the life expectancy at birth, in 1965. Source: Gallup et al. [1999].

School attainment: It is the logarithm of the years of secondary schooling, in 1965. Source: Gallup et al. [1999].

Openness: It is the Sachs-Warner openness indicator. Source: Gallup et al. [1999].

Newstate: It is a discrete variable that represent the timing of national independence. It takes the value 0 if independence was achieved before 1914; 1 if between 1914 and 1945; 2 if between 1946 and 1989; and 3 if after 1989 Source: Gallup et al. [1999].

<u>Geographical variables</u>

Latitude: It measures the absolute value of the latitude of a country, scaled to an index ranging between 0 and 1 where 0 is the equator. Source: La Porta et al. [1999].

Landlocked dummy: Equal to 1 if country do not adjoin the sea. Source: Gallup et al. [1999].

Landlocked dummy (not W&C Europe): Equal to 1 if country do not adjoin the sea, excluding countries in Western and Central Europe. Source: Gallup et al. [1999].

Tropics: It measures the percentage of country area with a tropical climate. Source: Englebert, website: www.politics.pomona.edu/penglebert. He expanded deductively the coverage from Sachs and Warner [1997].

Tropics (% land area): It is the proportion of the country's land area within the geographical tropics. Source: Gallup et al. [1999].

Tropics (% population): It is the proportion of the country's population living within the geographical tropics. Source: Gallup et al. [1999].

% land within 100km coast: It is the proportion of the country's land area within 100 km of the ocean coastline. Source: Gallup et al. [1999].

% land within 100km coast or river: It is the proportion of the country's land area within 100 km of the ocean or ocean navigable river. Source: Gallup et al. [1999].

% population within 100km coast: It is the proportion of the population in 1994 within 100 km of the coastline. Source: Gallup et al. [1999].

% population within 100km coast or river: It is the proportion of the population in 1994 within 100 km of the coastline or ocean navigable river. Source: Gallup et al. [1999].

Coastal Density: Coastal population/Coastal km2, in 1965 and 1995. Source: Gallup et al. [1999].

Inland Density: Interior population/Coastal km2, in 1965 and 1995. Source: Gallup et al. [1999].

Km to closet major market: It is the log of the minimum air distance (in km.) to one of the three main markets: New York, Rotterdam and Tokio. Source: Gallup et al. [1999].

% land area (population) in tropics: It measures the percentage of land area (or population in 1995) in Koeppen-Geiger tropics ecozones. Source: http://www.cid.harvard.edu/ciddata/ciddata.html

% land area (population) in tropics and subtropics: It measures the percentage of land area (or population in 1995) in Koeppen-Geiger tropics and subtropics ecozones. Source: http://www.cid.harvard.edu/

% land area (population) in temperate zones: It measures the percentage of land area (or population in 1995) in Koeppen-Geiger temperate ecozones. Source: http://www.cid.harvard.edu/

Temperature: These are the mean, monthly high (min and max), monthly low (min and max) values. The temperature range is just the difference between the monthly high maximum and the monthly low minimum. Source: La Porta et al. [1999].

Humidity: These are the morning and afternoon minimum and maximum values. Source: La Porta et al. [1999].

Malaria variables

Malaria Ecology: It is the Malaria Ecology index, that is an instrument for malaria risk. It is weighted for the population and updated to September 2003. Source: The Earth Institute at Columbia University, http://www.earthinstitute.columbia.edu and Kiszewski et al. [2004].

Falciparum Malaria Index: It is the index of malaria prevalence, computed as the product of the fraction of land area subject to malaria times the fraction of falciparum malaria cases, for the years 1966 and 1994. Source: Gallup et al. [1999].

% of 1995 population living in areas with malaria, 1994: It is an indicator about the proportion of population in 1995 living in area with malaria. Source: Sachs [2003].

Malfal: It is the % of 1995 population living in areas with malaria multiplied by an estimate of the proportion of national malaria cases that involve the fatal species, Plasmodium falciparum, as opposed to three largely non-fatal species of the malaria pathogen. Source: Sachs [2003].

Institutional Indicators

Corruption in government: Ranges in value from 0-6, with higher values indicating lower corruption. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

Rule of law: Ranges in value from 0-6, with high values indicating sound political institutions. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

Quality of bureaucracy: Ranges in value from 0-6, with high values indicating autonomy from political pressures. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

Risk of repudiation of contracts by government: Ranges in value from 0-10, with higher values indicating a lower likelihood that a country will modify or repudiate a contract with a foreign business. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

Risk of expropriation of private investment: Ranges in value from 0-10, with high values indicating a low probability of outright confiscation and forced nationalization. It is the averaged value over the period 1982-1997. Source: IRIS-3 File of International Country Risk Guide (ICRG) Data.

APPENDIX B: TABLES

Summary statistics

Table 1: Summary statistics for the IRIS-3 institution quality indicators.

Variable	Observations	Mean	Std. Dev.	Min	Max
Corruption	122	3.363686	1.327877	0.11875	6
Rule of law	122	3.432632	1.43061	0.91875	6
Bureaucracy	122	3.223064	1.494048	0.90625	6
Repudiation	122	6.416255	1.900521	2.05	10
Expropriation risk	122	7.108193	1.758161	2.4	10

Notes: All the indicators are average values for each countries over the period specified for each dataset.

Table 2: Summary statistics for the geography indicators.

Variable	Obs.	Mean	Std. Dev.	Min	Max
% land area in temperate zones	159	0.2930	0.4171	0	1
% population in temperate zones	159	0.3289	0.4310	0	1
Tropics (% land area)	146	0.4828	0.4784	0	1
Tropics (% population)	135	0.2825	0.3628	0	1
Km to closest major market	146	4034.9	2399.7	140	9320
Landlocked dummy (not W&C Europe)	146	0.1986	0.4003	0	1
Landlocked dummy	159	0.2264	0.4198	0	1
Falciparam Malaria Index, 1966	141	0.3265	0.4262	0	1
Falciparam Malaria Index, 1994	145	0.2975	0.4132	0	1
% of 1995 population living in areas with malaria, 1994	159	0.3168	0.4240	0	1
Malfal [Sachs, 2003]	159	0.3738	0.4387	0	1
Malaria Ecology	157	3.7151	6.6557	0	31.5479
% land within 100km coast	146	0.3406	0.3443	0	1
% land within 100km coast or river	146	0.4471	0.3726	0	1
% population within 100km coast	146	0.4231	0.3603	0	1
% population within 100km coast or river	146	0.5390	0.3715	0	1
log of hydrocarbons per capita	144	0.7800	4.5754	-4.6052	10.5947

Correlations

Table 3: Pairwise correlations between log GDP pc and the IRIS-3 institution quality indicators.

	Log GDP per capita	Corruption	Rule of law	Bureaucracy	Repudiation	Expropria-tion risk
Log GDP per capita	1.000 (144)					
Corruption	0.7020* (111)	1.000 (122)				
Rule of law	0.7575* (111)	0.8240* (122)	1.000 (122)			
Bureaucracy	0.7594* (111)	0.8411* (122)	0.8217* (122)	1.000 (122)		
Repudiation	0.8382* (111)	0.7737* (122)	0.8040* (122)	0.8615* (122)	1.000 (122)	
Expropria-tion risk	0.7921* (111)	0.7569* (122)	0.8515* (122)	0.8305* (122)	0.9244* (122)	1.000 (122)

Notes: The number of observation is in brackets. Log GDP per capita is PPP in 2001. The star means that the correlation coefficient is significant at 1% level of confidence.

Table 4: Correlations between log GDP per capita and geography indicators.

% land area in tropics	-0.3832* (144)	% population in tropics	-0.3996* (144)
% land area in tropics and subtropics	-0.4659* (144)	% population in tropics and subtropics	-0.4800* (144)
% land area in temperate zones	0.6636* (144)	% population in temperate zones	0.6941* (144)
Tropics [Englebert, 2000]	-0.7342* (123)	Temperature: range	0.4661* (144)
Tropics (% land area) [GSM, 1998]	-0.6211* (136)	: mean	-0.5194* (144)
Tropics (% population) [GSM, 1998]	-0.6420* (128)	: monthly high (min)	-0.3637 (143)
Latitude	0.6538* (144)	: monthly high (max)	-0.0033 (143)
Landlocked dummy	-0.1858* (144)	: monthly low (min)	-0.3249* (143)
Landlocked dummy (not W&C Europe)	-0.3953* (136)	: monthly low (max)	-0.2837* (143)
% land within 100km coast	0.3283* (136)	Humidity: morning min	0.0687 (143)
% land within 100km coast or river	0.4604* (136)	: morning max	-0.0673 (143)
% population within 100km coast	0.3459* (136)	: afternoon min	0.1341 (143)
% population within 100km coast or river	0.5212* (136)	: afternoon max	0.0979 (143)
Coastal Density, 1965	-0.0933 (129)	Malaria Ecology	-0.5614* (143)
Coastal Density, 1995	-0.1290 (136)	Falciparum Malaria Index, 1966 [GSM, 1999]	-0.6252* (133)
Inland Density, 1965	0.2131* (128)	Falciparum Malaria Index, 1994 [GSM, 1999]	-0.6726* (135)
Inland Density, 1995	0.0852 (136)	% of 1995 population living in areas with malaria, 1994	-0.7370* (144)
Km to closet major market	-0.5331* (136)	Malfal [Sachs, 2003]	-0.7277* (144)

Notes: The number of observations is in brackets and the star means that the correlation is significant at 5% level of confidence.

Regression tables

Table 5: OLS regressions of log GDP per capita.

Dependent variables	AJR	update	Update full sample	AJR	update	Update full sample
Risk of expropriation	0.52* (0.06)	0.52* (0.07)	0.57* (0.04)	0.47* (0.06)	0.43* (0.07)	0.40* (0.05)
Latitude				Yes	2.35* (0.72)	1.97* (0.47)
Obs.	64	60	111	64	60	111
R^2	§	0.50	0.63	§	0.58	0.68

Notes: standard errors in brackets. A star means the 5% significance level. §: data not reported in the original paper.

Table 6: 2SLS regressions of log GDP per capita.

Different specifications:	AJR	update	AJR	update
First stage regression. Dependent variable: risk of expropriation				
Dependent variables:				
Log mortality rates	-0.61* (0.13)	-0.58* (0.13)	-0.51* (0.14)	-0.47* (0.14)
Latitude			2.00 (1.34)	2.35 (1.29)
Obs.	64	60	64	60
R-squared	0.27	0.27	0.30	0.31
F-statistic	§	21.74	§	12.96
Second stage regression. Dependent variable: log of GDP per capita				
Dependent variables:				
Expropriation risk	0.94* (0.16)	0.98* (0.17)	1.00* (0.22)	1.00* (0.25)
Latitude			-0.65 (1.34)	-0.14 (1.47)

Notes: standard errors in brackets. A star means the 5% significance level. §: data not reported in the original paper.

Table 7a: 2SLS regression of log GDP per capita. Risk of Expropriation as proxy for institutions quality.

Different specifications:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage regression. Dependent variable: risk of expropriation								
Dependent variables:								
Log mortality rates	-0.58* (0.13)	-0.46* (0.14)	-0.41* (0.13)	-0.56* (0.12)	-0.59* (0.19)	-0.60* (0.13)	-0.59* (0.13)	-0.52* (0.15)
Latitude		2.35 (1.29)						
% land area in temperate zones			2.24* (0.74)					
Landlocked dummy (not W&C Europe)				-0.94* (0.47)				
Malaria Ecology					0.00 (0.03)			
% land within 100km coast						0.19 (0.48)		
% land within 100km coast or river							0.46 (0.44)	
SSA dummy								-0.34 (0.38)
Obs.	60	60	60	59	60	59	59	59
R-squared	0.27	0.31	0.37	0.33	0.27	0.29	0.30	0.30
F-statistic	21.74	12.96	16.95	14.06	10.68	11.36	12.01	11.80
Second stage regression. Dependent variable: log of GDP per capita								
Dependent variables:								
Expropriation risk	0.98* (0.17)	0.99* (0.25)	1.13* (0.31)	0.98* (0.19)	0.77* (0.21)	0.95* (0.17)	0.96* (0.17)	0.85* (0.21)
Latitude		-0.14 (1.46)						
% land area in temperate zones			-1.12 (1.22)					
Landlocked dummy (not W&C Europe)				0.33 (0.45)				
Malaria Ecology					-0.023 (0.018)			
% land within 100km coast						0.23 (0.36)		
% land within 100km coast or river							0.01 (0.37)	
SSA dummy								-0.27 (0.32)

Notes: standard errors in brackets. A star means the 5% significance level.

Table 7b: 2SLS regression of log GDP per capita. Risk of Expropriation as proxy for institutions quality.

Different specifications:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage regression. Dependent variable: risk of expropriation								
Dependent variables:								
Log mortality rates	-0.35* (0.14)	-0.45* (0.15)	-0.43* (0.14)	-0.56* (0.13)	-0.45* (0.18)	-0.57* (0.13)	-0.54* (0.13)	-0.39* (0.16)
% population in temperate zones	1.77* (0.54)							
Tropics (% land area)		-0.83 (0.46)						
Tropics (% population)			1.12* (0.46)					
Km to closest major port (in log)				-0.36 (0.23)				
malfal					-0.50 (0.48)			
% population within 100km coast						0.66 (0.45)		
% population within 100km coast or river							1.01* (0.44)	
% of 1995 pop. living in areas with malaria								-0.81 (0.45)
Obs.	60	59	59	59	60	59	59	60
R-squared	0.39	0.32	0.35	0.32	0.28	0.31	0.35	0.31
F-statistic	18.12	13.47	15.40	12.99	11.42	12.74	15.00	12.96
Second stage regression. Dependent variable: log of GDP per capita								
Dependent variables:								
Expropriation risk	1.17* (0.40)	1.03* (0.29)	0.97* (0.27)	0.96* (0.18)	0.54* (0.19)	0.95* (0.18)	0.98* (0.20)	0.70* (0.25)
% population in temperate zones	-0.83 (1.10)							
Tropics (% land area)		0.24 (0.57)						
Tropics (% population)			0.06 (0.61)					
Km to closest major market (in log)				0.002 (0.21)				
malfal					-0.97* (0.31)			
% population within 100km coast						0.08 (0.41)		
% population within 100km coast or river							-0.19 (0.47)	
% of 1995 pop. living in areas with malaria								-0.69 (0.43)

Notes: standard errors in brackets. A star means the 5% significance level.

Table 8a: 2SLS regression of log GDP per capita. Risk of repudiation of contracts by government as proxy for institutions quality.

Different specifications:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage regression. Dependent variable: risk of repudiation of contracts by government								
Dependent variables:								
Log mortality rates	-0.76* (0.12)	-0.65* (0.14)	-0.55* (0.13)	-0.76* (0.13)	-0.85* (0.19)	-0.78* (0.13)	-0.77* (0.13)	-0.73* (0.15)
Latitude		2.15 (1.28)						
% land area in temperate zones			2.58* (0.71)					
Landlocked dummy (not W&C Europe)				-0.30 (0.48)				
Malaria Ecology					0.02 (0.03)			
% land within 100km coast						-0.19 (0.45)		
% land within 100km coast or river							-0.01 (0.44)	
SSA dummy								-0.16 (0.38)
Obs.	60	60	60	59	60	59	59	59
R-squared	0.39	0.42	0.51	0.41	0.40	0.40	0.40	0.40
Second stage regression. Dependent variable: log of GDP per capita								
Dependent variables:								
Risk of repudiation of contracts by government	0.75* (0.11)	0.71* (0.13)	0.83* (0.17)	0.73* (0.11)	0.53* (0.12)	0.73* (0.10)	0.73* (0.10)	0.61* (0.12)
Latitude		0.66 (0.99)						
% land area in temperate zones			-0.73 (0.85)					
Landlocked dummy (not W&C Europe)				-0.37 (0.31)				
Malaria Ecology					-0.032* (0.012)			
% land within 100km coast						0.54** (0.28)		
% land within 100km coast or river							0.47** (0.28)	
SSA dummy								-0.47* (0.22)

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double star that the coefficient is significant at 10% level of confidence.

Table 8b: 2SLS regression of log GDP per capita. Risk of Expropriation as proxy for institutions quality.

Different specifications:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First stage regression. Dependent variable: risk of repudiation of contracts by government								
Dependent variables:								
Log mortality rates	-0.51* (0.13)	-0.59* (0.15)	-0.67* (0.14)	-0.74* (0.13)	-0.56* (0.17)	-0.76* (0.13)	-0.73* (0.13)	-0.53* (0.16)
% population in temperate zones	1.88* (0.52)							
Tropics (% land area)		-1.00* (0.45)						
Tropics (% population)			-0.66* (0.47)					
Km to closest major market (in log)				-0.26 (0.23)				
malfal					-0.74 (0.47)			
% population within 100km coast						0.27 (0.46)		
% population within 100km coast or river							0.58 (0.45)	
% of 1995 pop. living in areas with malaria								-0.94* (0.43)
Obs.	60	59	59	59	60	59	59	60
R-squared	0.51	0.45	0.42	0.41	0.42	0.41	0.42	0.44
Second stage regression. Dependent variable: log of GDP per capita								
Dependent variables:								
Risk of repudiation of contracts by government	0.81* (0.20)	0.79* (0.17)	0.63* (0.12)	0.72* (0.11)	0.44* (0.14)	0.71* (0.10)	0.71* (0.11)	0.51* (0.15)
% population in temperate zones	-0.27 (0.67)							
Tropics (% land area)		0.18 (0.44)						
Tropics (% population)			-0.62** (0.32)					
Km to closest major market (in log)				-0.16 (0.16)				
malfal					-0.92* (0.30)			
% population within 100km coast						0.53** (0.29)		
% population within 100km coast or river							0.39 (0.30)	
% of 1995 pop. living in areas with malaria								-0.77* (0.33)

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double star that the coefficient is significant at 10% level of confidence.

Table B1: OLS estimates for the specifications reported in Tables 7a-b.

Dependent variable: log of GDP per capita (Table 5a)								
Dependent variables:								
Expropriation risk	0.52* (0.06)	0.43* (0.07)	0.42* (0.08)	0.51* (0.07)	0.41* (0.06)	0.51* (0.07)	0.51* (0.07)	0.42* (0.07)
Latitude		2.35* (0.72)						
% land area in temperate zones			1.20* (0.48)					
Landlocked dummy (not W&C Europe)				-0.27 (0.30)				
Malaria Ecology					-0.05* (0.01)			
% land within 100km coast						0.43 (0.27)		
% land within 100km coast or river							0.32 (0.27)	
SSA dummy								-0.75* (0.18)
Obs.	60	60	60	59	60	59	59	59
R-squared	0.51	0.58	0.55	0.52	0.63	0.53	0.52	0.63
Dependent variable: log of GDP per capita (Table 5b)								
Dependent variables:								
Expropriation risk	0.38* (0.08)	0.43* (0.07)	0.39* (0.07)	0.49* (0.07)	0.33* (0.05)	0.49* (0.07)	0.47* (0.07)	0.32* (0.06)
% population in temperate zones	1.15* (0.33)							
Tropics (% land area)		-0.74* (0.25)						
Tropics (% population)			-1.02* (0.26)					
Km to closest major market (in log)				- 0.26** (0.14)				
malfal					-1.26* (0.16)			
% population within 100km coast						0.58* (0.27)		
% population within 100km coast or river							0.57* (0.28)	
% of 1995 pop. living in areas with malaria								-1.26* (0.17)
Obs.	60	59	59	59	60	59	59	60
R-squared	0.59	0.58	0.62	0.54	0.76	0.55	0.55	0.75

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double star that the coefficient is significant at 10% level of confidence.

Table B2: OLS estimates for the specifications reported in Tables 8a-b.

Dependent variable: log of GDP per capita (Table 5a)								
Dependent variables:								
Risk of repudiation of contracts by government	0.51* (0.06)	0.43* (0.06)	0.44* (0.07)	0.49* (0.06)	0.41* (0.06)	0.51* (0.06)	0.51* (0.06)	0.42* (0.06)
Latitude		2.08* (0.69)						
% land area in temperate zones			0.83** (0.48)					
Landlocked dummy (not W&C Europe)				-0.55* (0.27)				
Malaria Ecology					-0.04* (0.01)			
% land within 100km coast						0.57* (0.25)		
% land within 100km coast or river							0.53* (0.24)	
SSA dummy								-0.70* (0.17)
Obs.	60	60	60	59	60	59	59	59
R-squared	0.57	0.62	0.59	0.68	0.67	0.61	0.60	0.67
Dependent variable: log of GDP per capita (Table 5b)								
Dependent variables:								
Risk of repudiation of contracts by government	0.39* (0.07)	0.44* (0.07)	0.40* (0.06)	0.48* (0.06)	0.32* (0.05)	0.48* (0.06)	0.47* (0.06)	0.31* (0.05)
% population in temperate zones	0.95* (0.34)							
Tropics (% land area)		-0.55* (0.26)						
Tropics (% population)			-1.03* (0.23)					
Km to closest major market (in log)				-0.29* (0.13)				
malfal					-1.14* (0.18)			
% population within 100km coast						0.71* (0.24)		
% population within 100km coast or river							0.69* (0.25)	
% of 1995 pop. living in areas with malaria								-1.15* (0.18)
Obs.	60	59	59	59	60	59	59	60
R-squared	0.62	0.60	0.69	0.61	0.75	0.63	0.62	0.75

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double star that the coefficient is significant at 10% level of confidence.

Table 9a: Overidentification and risk of repudiation as institutional indicator.

Dependent variable: log of GDP per capita												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Risk of repudiation of contracts by government	0.71* (0.09)	0.74* (0.09)	0.69* (0.09)	0.73* (0.09)	0.70* (0.09)	0.73* (0.09)	0.68* (0.09)	0.70* (0.09)	0.67* (0.09)	0.70* (0.10)	0.59* (0.10)	0.62* (0.10)
% land within 100km coast			0.55* (0.27)	0.54** (0.28)								
% land within 100km coast or river					0.48** (0.27)	0.47** (0.28)						
% population within 100km coast							0.55* (0.27)	0.53** (0.28)				
% population within 100km coast or river									0.45 (0.28)	0.41 (0.29)		
SSA dummy											-0.49* (0.21)	-0.45* (0.21)
Instruments												
Log mortality rates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
% land area in temperate zones	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
% population in temperate zones	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Statistics from the first stage regression												
R-squared	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.52	0.52	0.51	0.51
F-statistic	29.44	29.23	19.15	18.90	19.09	19.03	19.33	19.08	19.94	19.79	19.20	18.97
Obs.	60	60	59	59	59	59	59	59	59	59	59	59
Overidentifying restriction tests (chi-squared critical value at 5%: 3.84)												
Value of the chi-squared	0.958	0.195	0.554	0.000	0.667	0.004	0.623	0.059	0.764	0.079	0.082	0.070
p-value	0.3277	0.6591	0.4567	0.9945	0.4141	0.9506	0.4299	0.8076	0.3820	0.7787	0.7741	0.7910

Notes: *Notes*: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double at 10% level of confidence.

Table 9b: Overidentification and risk of expropriation as institutional indicator.

Dependent variable: log of GDP per capita												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Expropriation risk	0.89* (0.14)	0.90* (0.13)	0.87* (0.14)	0.89* (0.13)	0.88* (0.14)	0.89* (0.13)	0.87* (0.14)	0.88* (0.14)	0.89* (0.16)	0.90* (0.15)	0.77* (0.15)	0.80* (0.15)
% land within 100km coast			0.26 (0.34)	0.26 (0.34)								
% land within 100km coast or river					0.07 (0.34)	0.06 (0.34)						
% population within 100km coast							0.17 (0.36)	0.16 (0.37)				
% population within 100km coast or river									-0.05 (0.40)	-0.07 (0.40)		
SSA dummy											-0.36* (0.27)	-0.33 (0.27)
Instruments												
Log mortality rates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
% land area in temperate zones	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
% population in temperate zones	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Statistics from the first stage regression												
R-squared	0.37	0.39	0.38	0.40	0.39	0.42	0.40	0.42	0.43	0.44	0.38	0.40
F-statistic	16.95	18.12	11.34	12.27	11.91	13.21	12.40	13.00	13.78	14.61	11.61	12.25
Obs.	60	60	59	59	59	59	59	59	59	59	59	59
Overidentifying restriction tests (chi-squared critical value at 5%: 3.84)												
Value of the chi-squared	1.346	0.946	0.882	0.435	0.985	0.589	0.951	0.580	1.016	0.637	0.463	0.200
p-value	0.2459	0.3306	0.3476	0.5095	0.3211	0.4428	0.3294	0.4464	0.3135	0.4249	0.4963	0.6550

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double at 10% level of confidence.

Table 9c: Overidentification and bureaucracy index as institutional indicator.

Dependent variable: log of GDP per capita												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Quality of bureaucracy	1.00* (0.19)	0.94* (0.17)	0.97* (0.17)	0.95* (0.16)	0.97* (0.17)	0.95* (0.16)	0.93* (0.17)	0.89* (0.15)	0.90* (0.17)	0.86* (0.15)	0.69* (0.12)	0.68* (0.11)
% land within 100km coast			0.76* (0.38)	0.76* (0.37)								
% land within 100km coast or river					0.81* (0.37)	0.81* (0.36)						
% population within 100km coast							0.81* (0.37)	0.82* (0.35)				
% population within 100km coast or river									0.82* (0.35)	0.84* (0.34)		
SSA dummy											-0.93* (0.19)	-0.94* (0.19)
Instruments												
Log mortality rates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
% land area in temperate zones	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
% population in temperate zones	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Statistics from the first stage regression												
R-squared	0.32	0.38	0.34	0.38	0.34	0.38	0.34	0.38	0.34	0.38	0.36	0.41
F-statistic	13.65	17.14	9.55	11.45	9.63	11.47	9.33	11.43	9.33	11.40	10.50	12.87
Obs.	60	60	59	59	59	59	59	59	59	59	59	59
Overidentifying restriction tests (chi-squared critical value at 5%: 3.84)												
Value of the chi-squared	1.974	2.807	1.010	1.119	1.136	1.129	1.098	1.663	1.303	1.709	0.026	0.063
p-value	0.1600	0.0939	0.3148	0.2902	0.2866	0.2880	0.2947	0.1973	0.2536	0.1911	0.8716	0.8011

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double at 10% level of confidence.

Table 10: Overidentification and risk of repudiation and malaria indicators as explanatory variables.

Dependent variable: log of GDP per capita												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Risk of repudiation of contracts by government	0.34* (0.13)	0.35* (0.14)	0.39* (0.12)	0.40* (0.13)	0.35* (0.16)	0.34** (0.18)	0.40* (0.13)	0.42* (0.14)	0.44* (0.12)	0.46* (0.13)	-1.20 (11.90)	0.23 (0.68)
% of 1995 pop. living in areas with malaria	-1.32* (0.42)	-1.31* (0.44)	-1.04* (0.37)	-1.02* (0.38)	-1.20 (0.76)	-1.28 (0.76)						
Falciparum malaria index, 1994							-1.23* (0.45)	-1.20* (0.47)	-0.99* (0.41)	-0.97* (0.42)	-20.37 (136.5)	-3.93 (6.82)
Landlocked dummy (not W&C Europe)			-0.48* (0.21)	-0.47* (0.21)					-0.43** (0.24)	-0.42** (0.24)		
SSA dummy					-0.08 (0.32)	-0.04 (0.31)					13.06 (90.99)	2.10 (4.47)
Instruments												
Log mortality rates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
% land in temperate zones	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
% pop. in temperate zones	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Malaria Ecology	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Statistics from the first stage regression												
Risk of repudiation of contracts by government												
R-squared	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.52	0.51
F-statistic	19.37	19.23	14.33	14.13	14.40	14.21	19.20	19.00	14.33	14.13	14.40	14.21
% of 1995 pop. living in areas with malaria												
R-squared	0.55	0.59	0.57	0.61	0.59	0.62	0.49	0.51	0.50	0.52	0.76	0.77
F-statistic	23.16	26.60	18.06	20.84	19.76	22.17	17.68	18.84	13.46	14.46	43.40	45.64
Obs.	60	60	59	59	59	59	59	59	59	59	59	59
Overidentifying restriction tests (chi-squared critical value at 5%: 3.84)												
Value of the chi-squared	0.000	0.026	0.009	0.042	0.052	0.213	0.058	0.349	0.000	0.155	0.012	0.665
p-value	0.9888	0.8718	0.9253	0.8378	0.8204	0.6449	0.8095	0.5550	0.9973	0.6938	0.9138	0.4150

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double at 10% level of confidence.

Table 11: Overidentification with risk of repudiation and malaria as explanatory variables and a different set of instruments.

Dependent variable: log of GDP per capita									
Instruments: Malaria Ecology, percentage of population living in temperate zones and newstate									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Risk of repud. of contracts by gov.	0.31* (0.13)	0.33* (0.12)	0.30* (0.12)	0.34* (0.12)	0.31* (0.12)	0.33* (0.12)	0.39* (0.14)	0.31* (0.11)	0.33* (0.11)
% of pop. living in malaria areas	-1.47* (0.46)	-1.31* (0.44)	-1.36* (0.44)	-1.27* (0.44)	-1.34* (0.44)	-1.24* (0.41)	-1.56* (0.46)	-1.41* (0.42)	-1.22* (0.40)
% pop. within 100km coast		0.28** (0.15)							0.32* (0.14)
% pop. within 100km coast/river			0.29** (0.18)						
% land within 100km coast				0.30* (0.15)					
% land 100km coast/river					0.29** (0.16)				
Landlocked (not W&C Europe)						-0.57* (0.16)			
Temperature (average)							0.89* (0.44)		
Log of Hydrocarbons								0.02** (0.01)	0.02* (0.01)
Statistics from the first stage regression									
Risk of repudiation of contracts by government									
R-squared	0.54	0.54	0.55	0.54	0.54	0.54	0.54	0.54	0.55
F-statistic	39.29	29.45	30.44	29.31	29.56	29.22	30.18	29.81	23.92
% of 1995 pop. living in areas with malaria									
R-squared	0.67	0.67	0.67	0.67	0.67	0.68	0.68	0.68	0.68
F-statistic	69.77	51.84	52.33	51.86	51.82	52.70	52.57	52.67	41.76
Obs.	106	106	106	106	106	106	106	105	105
Overidentifying restriction tests (chi-squared critical value at 5%: 3.84)									
Value of the test	0.553	0.684	0.780	0.801	0.919	1.482	0.347	0.139	0.174
p-value	0.4570	0.4081	0.3770	0.3708	0.3378	0.2234	0.5559	0.7095	0.6769

Notes: standard errors in brackets. A star means that the coefficient is significant at 5% level of confidence. A double star that the coefficient is significant at 10% level of confidence. Ethnolinguistic fractionalisation; ethnic tensions; SSA and other continental dummies; the log of distance to the closest major market; the percentage of land area and of population living in the tropics; life expectancy in 1965; the legitimacy dummy; the Sachs-Warner openness indicator; the humidity indicators; other temperature indexes; latitude; the log of the years of school attainment; and the indicators of coastal and inland density in 1965 and in 1995 are not significant. Data not shown for reasons of space.