

NBER WORKING PAPER SERIES

INSTITUTIONS DON'T RULE:
DIRECT EFFECTS OF GEOGRAPHY
ON PER CAPITA INCOME

Jeffrey D. Sachs

Working Paper 9490
<http://www.nber.org/papers/w9490>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2003

Director, The Earth Institute at Columbia University. All data are available from the author upon request. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2003 by Jeffrey D. Sachs. All rights reserved. Short sections of text not to exceed two paragraphs, may be quoted without explicit permission provided that full credit including notice, is given to the source.

Institutions Don't Rule: Direct Effects of Geography on Per Capita Income

Jeffrey D. Sachs

NBER Working Paper No. 9490

February 2003

JEL No. O11, P16, P51, R11

ABSTRACT

In a series of papers, my colleagues and I have demonstrated that levels of per capita income, economic growth, and other economic and demographic dimensions are strongly correlated with geographical and ecological variables such as climate zone, disease ecology, and distance from the coast. Three recent papers purport to show that the role of geography in explaining cross-country patterns of income per capita operates predominantly or exclusively through the choice of institutions, with little direct effect of geography on income after controlling for the quality of institutions. This note shows that malaria transmission, which is strongly affected by ecological conditions, directly affects the level of per capita income after controlling for the quality of institutions.

Jeffrey D. Sachs

Columbia University

314 Low Library

535 West 116th Street, MC4327

New York, NY 10027

and NBER

sachs@columbia.edu

Introduction

In a series of papers, my colleagues and I (Gallup, Sachs, and Mellinger, 1998, 2000; Gallup and Sachs, 2001; Sachs and Malaney, 2002) have demonstrated that levels of per capita income, economic growth, and other economic and demographic dimensions are strongly correlated with key geographical and ecological variables, such as climate zone, disease ecology, and distance from the coast. Several recent papers have explored this link between geography and development, testing whether the effects of geography on economic activity work mainly through various direct channels (e.g. through effects on productivity, population growth, trade and investment) or indirect channels (e.g. through effects on the choice of political and economic institutions). Of course, both kinds of channels may be present simultaneously, and Gallup, Sachs, and Mellinger (1998) offer a theoretical illustration as to why a disadvantageous physical environment may be conducive to the adoption of less productive institutional arrangements as well.

Three papers in particular (Acemoglu, Johnson, and Robinson, 2001; Easterly and Levine, 2002; and Rodrik, Subramanian, and Trebbis, 2002) purport to show that the role of geography in explaining cross-country patterns of income per capita operates predominantly or exclusively through the choice of institutions, with little direct effect of geography on income. This would indeed be a surprising conclusion, since many of the reasons why geography seems to have affected institutional choices in the past (e.g. the suitability of locations for European technologies, the disease environment and risks to survival of immigrants, the productivity of agriculture, the transport costs between far-flung regions and major markets) are indeed based on *direct* effects of geography on production systems, human health, and environmental sustainability, and many of those very same channels would still be likely to apply today. Thus,

the logic of the geography-institutions linkage is also the logic of a direct geography-productivity linkage.

Fortunately, these three papers propose a testable – and refutable – hypothesis that geography matters only through institutions, based on the following cross-country regression model:

$$(1) \ln(Y_i) = \beta_0 + \beta_1 QI_i + \beta_2' Z_i + \varepsilon_i$$

The log of income per capita of country i , $\ln(Y_i)$, is specified to be a function of the “quality of institutions,” according to an index QI_i , and one or more other variables Z_i , which may include geographical, historical or policy variables. Since income is affected by geography (Z), the argument suggests that Z works only through QI , with no direct effects once QI is entered into the regression equation. Specifically, the papers advance the null hypothesis that the coefficient vector β_2' is zero, so geographical (and other “non-institutional”) variables have no additional explanatory power beyond their possible indirect role in affecting the quality of institutions:

$$(2) H_0 : \beta_2' = 0$$

In the empirical test of H_0 in the three recent papers, the income variable is purchasing-power-parity adjusted GNP per capita in 1995 and the institutions variable is a proxy measure of institutional quality.

This current note shows that the null hypothesis in (2) is rejected when the Z_i variable is a measure of malaria transmission. In other words, I provide evidence that malaria transmission directly impacts the level of per capita income after controlling for institutional quality. The likely reasons for this linkage are described in Gallup and Sachs, 2001, and Sachs and Malaney, 2002. Before proceeding to the statistical test, however, it is important to stress that the model in (1) is worryingly oversimplified in any case (and thus is certainly not the model of choice that

I would specify or prefer to test). No reliable conclusions about the primacy of institutions over other variables should. The first obvious specification problem is one of statics versus dynamics. Economic theory suggests that the determination of per capita income should be specified as a *dynamic process* (e.g. Barro and Sala-i-Martin, 1997), in which the growth of income during a time interval $[0,T]$ is a function of the income level at the start of the period and some kind of average of the values of the “forcing variables” during the time interval $[0,T]$:

$$(3) \quad (1/T)\ln(Y_{iT}/Y_{i0}) = \beta_0 + \beta_1 QI_{i[0,T]} + \beta_2 Z_{i[0,T]} + \beta_3 \ln(Y_{i0}) + \varepsilon_i$$

It is much more likely that the quality of institutions in a given time period will affect the *growth rate* of the economy during that period (controlling for initial income), as opposed to the contemporaneous *level* of national income. It is also very doubtful that a process as complex as economic development can possibly be explained by two or three variables alone, much less the particular “geography” variables stressed by AJR and RST. Distance from the equator, the centerpiece of testing in RST, is an exceedingly poor choice for a serious test of geographical variables.¹ It is at best a proxy, and a poor one at that, for climate or possibly for distance from major markets, and should not be used as the basis of the bulk of the tests in the RST paper when much better alternatives are available. In any event, most geographical analyses stress several factors (climate, geographical isolation, disease environment), so that testing these variables one at a time is subject to extreme risk of left-out-variable error.

The purpose of this note, however, is to show that even within narrow confines of equation (1) the null hypothesis is easily refuted. I focus on one particular variable, malaria risk. It is easy to show that highly malarious regions of the world (i.e. regions where there is a high

¹ I have made this point elsewhere in Sachs (2000). It is interesting in this regard that when this variable was introduced by Hall and Jones (1999), it was taken to be a proxy for European settlement, not a proxy for geography per se. It is, incidentally, a fairly miserable measure for European settlement as well, another case in which much more direct measures are readily available.

risk of malaria transmission) have lower per capita income than non-malarious regions, controlling for the quality of government institutions. Moreover, since malaria transmission is strongly related to ecological conditions (specifically the type of mosquito vectors and the climate conditions) there is an excellent instrumental variable that can be used for malaria prevalence to help us to test the proposition that malaria “causes” poverty rather than vice versa.

Malaria Risk

As the main measure of malaria risk, I use an indicator based on the 1994 WHO world map of malaria risk. By overlaying the world map of malaria risk with a map of world population, my colleagues and I have calculated the proportion of each country’s population that live with risk of malaria transmission, labeled MAL94P. An alternative measure, introduced in Gallup and Sachs (1998), and used in AJR and RST, multiplies the MALP94 index 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 (*P. vivax*, *P. malariae*, and *P. ovale*). This index is called MALFAL. Sub-Saharan Africa has a very high proportion of malaria cases that are *falciparum* malaria, whereas in the Americas, Europe, and much of Asia, a higher proportion of malaria *vivax* cases is found.

Neither measure is a completely satisfactory index of the intensity of disease transmission or the burden of the disease in terms of morbidity and mortality. The basic problem is that the number of cases of malaria reported annually to WHO is a tiny fraction of the total number of malaria cases in a given year. That is why we rely on a measure of population at risk rather than actual infections. Most malaria cases in Africa are self treated if they are treated at all. Even deaths due to malaria are often unreported or are not classified by cause of death. There are also

deep clinical issues in the attribution of cause of death in the case of malaria. Two kinds of errors can occur. Some deaths may occur due to other causes but be attributed to malaria since the individual may simultaneously be infected by malaria. Alternatively, some deaths attributed to other causes may indeed have malaria as a co-factor, but not the sole or even principle cause. Also, multiplying MALP94 by the share of falciparum introduces one error while correcting another. It is important to distinguish countries burdened by falciparum from those suffering mainly or solely from non-fatal variants. On the other hand, in a case where falciparum is, say, 50% of all clinical cases, it may still be true that 100% of the public is at risk of falciparum infection. It would be best to make an index of percent of population at risk of falciparum infection, but we don't have data of that sort available.

Malaria is intrinsically a disease of warm environments because a key part of the life cycle of the parasite (sporogony) depends on a high ambient temperature. This is, in essence, why malaria is a disease of the tropics and the sub-tropics. Falciparum malaria requires even higher temperatures than vivax malaria. Malaria also depends on adequate conditions for mosquito breeding, mainly pools of clean water, usually due to rainfall ending up in puddles, cisterns, discarded tires, and the like. As a result, malaria has a distinct seasonality in the sub-humid tropics, where wet and dry seasons alternate, and mosquito breeding and hence malaria follows the rainy season. Additionally, the intensity of malaria transmission depends on the specific mosquito vectors that are present. All malaria is transmitted by mosquitos of the genus anopheles. Some anopheles species, especially those in sub-Saharan Africa, show a high preference for taking their blood meals from humans (anthropophagy) as opposed to animals such as cattle. These human-biting vectors lead to much more intensive transmission of the disease.

The temperature, mosquito abundance, and vector specificity can be combined into a single measure of Malaria Ecology (ME), an ecologically-based variable that is predictive of malaria risk. (Kiszewski et. al., forthcoming). The basic formula for ME includes temperature, species abundance, and vector type. The underlying index is measured on a highly disaggregated sub-national level, and then is averaged for the entire country. Because ME is built upon climatological and vector conditions on a country-by-country basis, and is therefore exogenous to public health interventions and economic conditions, ME provides an ideal instrumental variable for malaria risk.²

Estimation Strategy

The key is to estimate equation (1), in order to test whether malaria risk helps to explain cross-country patterns of national income after controlling for the quality of institutions. In order to do this, both the institutional quality and the malaria prevalence variables should be instrumented. I use three instruments.

KGPTEMP. This variable measures the share of a country's population in temperate ecozones, based on the Koeppen-Geiger ecozone classification system. (Mellinger, Gallup, and Sachs, 2000).

LOGMORT. This variable, created by AJR based on a range of source materials, is an estimate of mortality rates of British soldiers and other populations in the early 19th century, and has been used as an instrumental variable for institutional quality in all three papers.

ME. This variable, described above, provides an instrument for malaria risk that controls for the fact that causation may run not only from malaria to income but also from income to malaria.

Since the AJR sample of ex-colonies covers only 64 countries, it is convenient to use KGPTEMP as an additional instrument for institutional quality in order to estimate (1) for a larger sample of

² This variable was generated together with colleagues of mine at the Harvard Center for International Development.

countries in the AJR and RST data sets.³ Institutional quality is strongly positively correlated with the percentage of the population living in temperate-zone ecologies, so that KGPTTEMP is a useful instrument.⁴

The left-hand-side variable in each case is GNP per capita, PPP adjusted, in 1995. Each paper uses a slightly different version of this variable, so I use the same definition as in the original paper. Each paper also uses a distinct measure of institutional quality, and again, I use the same variable as in the respective papers.

The regression results are shown in Table 1. Each regression uses two-stage-least-squares to estimate the log income in 1995 as a function of institutional quality and malaria risk. For the large sample of countries, we use two instruments (KGPTTEMP and ME), while for the smaller set of countries in the AJR and EL papers, LOGMORT is added as an instrument. For each of the three papers there are four regressions, based on the two malaria variables, and the smaller or larger set of instruments.

The results are strong and robust. In every single regression, both the quality of institutions and the malaria risk variables are statistically significant at the 0.05 percent level, and in most cases at the 0.01 percent level. *The null hypothesis is decisively rejected in all three data sets, both for the larger number of countries (using KGPTTEMP and ME as instruments) and for the smaller number of countries (using all three instrumental variables).*

In Table 2, I add an additional right-hand-side geographical variable, the share of the national population living within 100km of the coast. For a variety of reasons, including overland transport costs, we should expect that countries with coastal populations should enjoy

³ In the EL paper, the quality of institutions variable is defined only for 62 countries in any event.

⁴ The positive relationship between institutional quality and KGPTTEMP may be due to the fact that European settlers preferred to settle in familiar temperate-zone settings, as suggested by AJR, though many alternative hypotheses are also possible.

higher per capita incomes than countries with substantially hinterland or even landlocked populations. We see that this is true in the large sample of countries (using AJR and RST) but is not evident in the smaller sample of countries in the EL data set.

Conclusions

The chief strength of the AJR, EL, and RST papers is that a clear null hypothesis is advanced. As Karl Popper taught many years ago, the failure to reject a null hypothesis on one set of data does not imply the correctness of the hypothesis more generally. This paper shows, indeed, that the null hypothesis may be readily rejected using an alternative geographical variable: malaria risk. A more complete model of development, however, will require that we move beyond the simple specifications tested in the three papers reviewed in this note. There is good theoretical and empirical reason to believe that the development process reflects a complex interaction of institutions, policies, and geography.

References

- Acemoglu, Daron, Simon Johnson and James A. Robinson. "The Colonial Origins of Comparative Development: an Empirical Investigation." *American Economic Review*, vol. 91, no. 5, December 2001.
- Barro, Robert J. and Xavier Sala-i-Martin. "Technological Diffusion, Convergence, and Growth." *Journal of Economic Growth*, vol. 2, no.1, March 1997, pgs. 1-26.
- Easterly, William and Ross Levine. "Tropics, Germs and Crops: How Endowments Influence Economic Development." NBER Working Paper 9106, August 2002.
- Gallup, John Luke, and Jeffrey Sachs D. with Andrew D. Mellinger. "Geography and Economic Development." in *Annual World Bank Conference on Development Economics 1998* (April), The World Bank: Washington, DC.
- Gallup, John Luke and Jeffrey D. Sachs. "The Economic Burden of Malaria," The Supplement to *The American Journal of Tropical Medicine & Hygiene*. Vol. 64, no. 1, 2, January/February 2001.
- Hall, Robert E. and Charles I. Jones. "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.
- Kiszewski, Anthony, Andrew Mellinger, Pia Malaney, Andrew Spielman, Sonia Ehrlich, Jeffrey Sachs. "A Global Index of the Stability of Malaria Transmission Based on the Intrinsic Properties of Anopheline Mosquito Vectors." Forthcoming.
- Mellinger, Andrew D., Jeffrey D. Sachs and John L. Gallup. "Climate, Coastal Proximity, and Development," in *Oxford Handbook of Economic Geography*. Clark, Gordon L., Maryann P. Feldman, and Meric S. Gertler, eds. Oxford University Press, 2000.
- Rodrik, Dani, Arvind Subramanian and Francesco Trebbi. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." NBER Working Paper 9305, October 2002.
- Sachs, Jeffrey D. and Pia Malaney. "The Economic and Social Burden of Malaria." *Nature Insight*, Vol. 415, no. 6872, Feb. 7, 2002.
- Sachs, Jeffrey D.. "Tropical Underdevelopment." prepared for Economic History Association Annual Meeting, September 2000 and CID Working Paper No. 57, December 2000 available at <http://www.cid.harvard.edu/cidwp/057.htm>

Table 1

Paper	(1) AJR	(2) AJR	(3) AJR	(4) AJR	(5) EL	(6) EL	(7) EL	(8) EL	(9) RST	(10) RST	(11) RST	(12) RST
Dependent variable:	logppp95	logppp95	logppp95	logppp95	lgdpppc95	lgdpppc95	lgdpppc95	lgdpppc95	lgdpp95	lgdpp95	lgdpp95	lgdpp95
<u>Independent variable</u>												
Average protection against expropriation risk, 1985-1995	0.29 (2.87)	0.45 (3.03)	0.53 (3.06)	0.56 (4.12)								
kk (Institutions Index)		1.38 (3.40)	1.32 (3.60)	1.59 (4.24)	1.51 (4.45)							
Rule						0.60 (3.45)	0.53 (2.74)	0.78 (5.60)	0.68 (4.01)			
MAL94P	-1.43 (-3.99)	-1.07 (-2.83)			-1.32 (-2.55)	-1.37 (-2.76)			-1.33 (-4.60)	-1.43 (-4.37)		
MALFAL			-1.04 (-2.74)	-0.83 (-2.47)			-1.15 (-2.35)	-1.20 (-2.52)			-1.13 (-4.53)	-1.25 (-4.26)
R ²	0.73	0.69	0.62	0.60	0.67	0.68	0.61	0.63	0.78	0.71	0.77	0.68
N	101	59	73	59	62	62	62	62	133	69	133	69
Instruments for variables above	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT	ME, LOGMORT

t-statistics are indicated in parentheses.

All regression equations are estimates with two-stage least squares and include a constant term (not reported).

First-stage regressions and (where relevant) overidentification tests support the use of the instruments in each case.

MALFAL and LOGMORT cover slightly different countries across papers, for consistency the corresponding authors' version of the variable is used.

The sample size in each regression varies slightly compared to the original regressions in AJR, EL and RST, respectively, due to differing coverage of the malaria and KGPTTEMP variables. These minor variations in sample size do not appear to affect the substantive results.

Table 2		(1)	(2)	(3)
Paper		AJR	EL	RST
Dependent variable:		logppp95	lgdppc95	lcgdp95
<u>Independent variable</u>				
Average protection against expropriation risk, 1985-1995		0.32 (3.32)		
kk (Institutions Index)			1.39 (3.36)	
RULE				0.60 (3.60)
mal94p		-1.19 (-3.37)	-1.31 (-2.23)	-1.25 (-4.59)
pop100km		0.47 (2.73)	0.016 (0.04)	0.34 (2.51)
R ²		0.76	0.67	0.79
N		101	62	133

t-statistics are indicated in parentheses.

All regressions include KGPTMP and ME as instruments for the dependent variables except pop100km

First-stage regressions and support the use of the instruments in each case.

POP100KM cover slightly different countries across papers, for consistency the

corresponding authors' version of the variable is used