

INSTITUTIONS RULE: THE PRIMACY OF INSTITUTIONS OVER
GEOGRAPHY AND INTEGRATION IN ECONOMIC DEVELOPMENT

Dani Rodrik
Harvard University

Arvind Subramanian
IMF

Francesco Trebbi
Harvard University

Revised
February 2004

ABSTRACT

We estimate the respective contributions of institutions, geography, and trade in determining income levels around the world, using recently developed instrumental variables for institutions and trade. Our results indicate that the quality of institutions “trumps” everything else. Once institutions are controlled for, conventional measures of geography have at best weak direct effects on incomes, although they have a strong indirect effect by influencing the quality of institutions. Similarly, once institutions are controlled for, trade is almost always insignificant, and often enters the income equation with the “wrong” (i.e., negative) sign. We relate our results to recent literature, and where differences exist, trace their origins to choices on samples, specification, and instrumentation.

The views expressed in this paper are the authors’ own and not of the institutions with which they are affiliated. We thank three referees, Chad Jones, James Robinson, Will Masters, and participants at the Harvard-MIT development seminar, joint IMF-World Bank Seminar, and the Harvard econometrics workshop for their comments, Daron Acemoglu for helpful conversations, and Simon Johnson for providing us with his data. Dani Rodrik gratefully acknowledges support from the Carnegie Corporation of New York.

Commerce and manufactures can seldom flourish long in any state which does not enjoy a regular administration of justice, in which the people do not feel themselves secure in the possession of their property, in which the faith of contracts is not supported by law, and in which the authority of the state is not supposed to be regularly employed in enforcing the payment of debts from all those who are able to pay. Commerce and manufactures, in short, can seldom flourish in any state in which there is not a certain degree of confidence in the justice of government.

-- Adam Smith, Wealth of Nations

I. Introduction

Average income levels in the world's richest and poorest nations differ by a factor of more than 100. Sierra Leone, the poorest economy for which we have national income statistics, has a per-capita GDP of \$490, compared to Luxembourg's \$50,061.¹ What accounts for these differences, and what (if anything) can we do to reduce them? It is hard to think of any question in economics that is of greater intellectual significance, or of greater relevance to the vast majority of the world's population.

In the voluminous literature on this subject, three strands of thoughts stand out. First, there is a long and distinguished line of theorizing that places geography at the center of the story. Geography is a key determinant of climate, endowment of natural resources, disease burden, transport costs, and diffusion of knowledge and technology from more advanced areas. It exerts therefore a strong influence on agricultural productivity and the quality of human resources. Recent writings by Jared Diamond and Jeffrey Sachs are among the more notable works in this tradition (see Diamond 1997; Gallup, Sachs, and Mellinger 1998, and Sachs 2001).

A second camp emphasizes the role of international trade as a driver of productivity change. We call this the integration view, as it gives market integration, and impediments thereof, a starring role in fostering economic convergence between rich and poor regions of the world. Notable recent research in this camp includes Frankel and Romer (FR, 1999) and the pre-geography work of Sachs (Sachs and Warner 1995).² It may be useful to distinguish between "moderate" and "maximal" versions of this view. Much of the economics profession would accept the hypothesis that trade can be an underlying source of growth once certain institutional prerequisites have been fulfilled. But a more extreme perspective, and one that

¹ These are figures for 2000, and they are expressed in current "international" dollars, adjusted for PPP differences. The source is the World Development Indicators CD-Rom of the World Bank.

² One can question whether it is appropriate to treat trade as one of the ultimate determinants of economic prosperity, but here we are simply following a long literature that has attached central causal importance to it.

has received wide currency in public debates, is that trade/integration is the major determinant of whether poor countries grow or not. It is the latter perspective that characterizes such widely cited papers as Sachs and Warner (1995) and Dollar and Kraay (2004).

Finally, a third group of explanations centers on institutions, and in particular the role of property rights and the rule of law. In this view, what matters are the rules of the game in a society and their conduciveness to desirable economic behavior. This view is associated most strongly with Douglass North (1990). It has received careful econometric treatment recently in Hall and Jones (1999), who focus on what they call “social infrastructure,” and in Acemoglu, Johnson, and Robinson (AJR, 2001), who focus on the expropriation risk that current and potential investors face.

Growth theory has traditionally focused on physical and human capital accumulation, and, in its endogenous growth variant, on technological change. But accumulation and technological change are at best proximate causes of economic growth. No sooner have we ascertained the impact of these two on growth—and with some luck their respective roles also—that we want to ask: But why did some societies manage to accumulate and innovate more rapidly than others? The three-fold classification offered above—geography, integration, and institutions—allows us to organize our thoughts on the “deeper” determinants of economic growth. These three are the factors that determine which societies will innovate and accumulate, and therefore develop, and which will not.

Since long-term economic development is a complex phenomenon, the idea that any one (or even all) of the above deep determinants can provide an adequate accounting of centuries of economic history is, on the face of it, preposterous. Historians and many social scientists prefer nuanced, layered explanations where these factors interact with human choices and many other not-so-simple twists and turns of fate. But economists like parsimony. We want to know how well these simple stories do, not only on their own or collectively, but more importantly, vis-à-vis each other. How much of the astounding variation in cross-national incomes around the world can geography, integration, and institutions explain? Do these factors operate additively, or do they interact? Are they all equally important? Does one of the explanations “trump” the other two?

The questions may be simple, but devising a reasonable empirical strategy for answering them is far from straightforward. This is not because we do not have good empirical proxies for each of these deep determinants. There are many reasonable measures of “geography,” such as distance from the equator (our preferred measure), percentage land mass located in the tropics, or average temperature. The intensity of an economy’s integration with the rest of the world can be measured by flows of trade or the height of trade barriers. The quality of institutions can be measured with a range of perceptions-based indicators of property rights and the rule of law. The difficulty lies instead in sorting out the complex web of causality that entangles these factors.

The extent to which an economy is integrated with the rest of the world and the quality of its institutions are both endogenous, shaped potentially not just by each other and by geography, but also by income levels. Problems of endogeneity and reverse causality plague any empirical researcher trying to make sense of the relationships among these causal factors. We illustrate this with the help of Figure 1, adapted from Rodrik (2003). The plethora of arrows in the figure, going in both directions at once in many cases, exemplifies the difficulty.

The task of demonstrating causality is perhaps easiest for the geographical determinists. Geography is as exogenous a determinant as an economist can ever hope to get, and the main burden here is to identify the main channel(s) through which geography influences economic performance. Geography may have a direct effect on incomes, through its effect on agricultural productivity and morbidity. This is shown with arrow (1) in Figure 1. It can also have an indirect effect through its impact on distance from markets and the extent of integration (arrow [2]) or its impact on the quality of domestic institutions (arrow [3]). With regard to the latter, economic historians have emphasized the disadvantageous consequences for institutional development of certain patterns of factor endowments, which engender extreme inequalities and enable the entrenchment of a small group of elites (e.g., Engerman and Sokoloff, 1994). A similar explanation, linking ample endowment of natural resources with stunted institutional development, also goes under the name of “resource curse” (Sala-i-Martin and Subramanian, 2003).

Trade fundamentalists and institutionalists have a considerably more difficult job to do, since they have to demonstrate causality for their preferred determinant, as well as identify the effective channel(s) through which it works. For the former, the task consists of showing that arrows (4) and (5)—capturing the direct impact of integration on income and the indirect impact through institutions, respectively—are the relevant ones, while arrows (6) and (7)—reverse feedbacks from incomes and institutions, respectively—are relatively insignificant. Reverse causality cannot be ruled out easily, since expanded trade and integration can be mainly the result of increased productivity in the economy and/or improved domestic institutions, rather than a cause thereof.

Institutionalists, meanwhile, have to worry about different kinds of reverse causality. They need to show that improvements in property rights, the rule of law and other aspects of the institutional environment are an independent determinant of incomes (arrow [8]), and are not simply the consequence of higher incomes (arrow [9]) or of greater integration (arrow [5]).

In econometric terms, what we need to sort all this out are good instruments for integration and institutions—sources of exogenous variation for the extent of integration and institutional quality, respectively, that are uncorrelated with other plausible (and excluded) determinants of income levels. Two recent papers help us make progress by providing plausible instruments. FR (1999) suggests that we can instrument for actual trade/GDP ratios by using trade/GDP shares constructed on the basis of a gravity equation for bilateral trade flows. The FR approach consists of first regressing bilateral trade flows (as a share of a country’s GDP) on measures of country mass, distance between the trade partners, and a few

other geographical variables, and then constructing a predicted aggregate trade share for each country on the basis of the coefficients estimated. This constructed trade share is then used as an instrument for actual trade shares in estimating the impact of trade on levels of income.

Acemoglu, Johnson, and Robinson (AJR, 2001) use mortality rates of colonial settlers as an instrument for institutional quality. They argue that settler mortality had an important effect on the type of institutions that were built in lands that were colonized by the main European powers. Where the colonizers encountered relatively few health hazards to European settlement, they erected solid institutions that protected property rights and established the rule of law. In other areas, their interests were limited to extracting as much resources as quickly as possible, and they showed little interest in building high-quality institutions. Under the added assumption that institutions change only gradually over time, AJR argue that settler mortality rates are therefore a good instrument for institutional quality. FR (1999) and AJR (2001) use their respective instruments to demonstrate strong causal effects from trade (in the case of FR) and institutions (in the case of AJR) to incomes. But neither paper embeds their estimation in the broader framework laid out above. More specifically, AJR control for geographical determinants, but do not check for the effects of integration. FR do not control for institutions.

Our approach in this paper consists of using the FR and AJR instruments simultaneously to estimate the structure shown in Figure 1. The idea is that these two instruments, having passed what might be called the AER (American Economic Review)-test, are our best hope at the moment of unraveling the tangle of cause-and-effect relationships involved. So we systematically estimate a series of regressions in which incomes are related to measures of geography, integration, and institutions, with the latter two instrumented using the FR and AJR instruments, respectively. These regressions allow us to answer the question: what is the independent contribution of these three sets of deep determinants to the cross-national variation in income levels? The first stage of these regressions provides us in turn with information about the causal links among the determinants.

This exercise yields some sharp and striking results. Most importantly, we find that the quality of institutions trumps everything else. Once institutions are controlled for, integration has no direct effect on incomes, while geography has at best weak direct effects. Trade often enters the income regression with the “wrong” (i.e., negative) sign, as do many of the geographical indicators. By contrast, our measure of property rights and the rule of law always enters with the correct sign, and is statistically significant, often with t-statistics that are very large.

On the links among determinants, we find that institutional quality has a positive and significant effect on integration. Our results also tend to confirm the findings of Easterly and

Levine (EL, 2003), namely that geography exerts a significant effect on the quality of institutions, and via this channel on incomes.³

Our preferred specification “accounts” for about half of the variance in incomes across the sample, with institutional quality (instrumented by settler mortality) doing most of the work. Our estimates indicate that an increase in institutional quality of one standard deviation, corresponding roughly to the difference between measured institutional quality in Bolivia and South Korea, produces a 2 log-points rise in per-capita incomes, or a 6.4-fold difference—which, not coincidentally, is also roughly the income difference between the two countries. In our preferred specification, trade and distance from the equator both exert a negative, but insignificant effect on incomes (see Table 2B, panel A, column 6).

Much of our paper is devoted to checking the robustness of our central results. In particular, we estimate our model for three different samples: (a) the original 64-country sample used by AJR; (b) a 79-country sample which is the largest sample we can use while still retaining the AJR instrument; and (c) a 137-country sample that maximizes the number of countries at the cost of replacing the AJR instrument with two more widely available instruments (fractions of the population speaking English and Western European languages as the first language, from Hall and Jones, 1999.) We also use a large number of alternative indicators of geography and integration. In all cases, institutional quality emerges as the clear winner of the “horse race” among the three. Finally, we compare and contrast our results to those in some recent papers that have undertaken exercises of a similar sort. Where there are differences in results, we identify and discuss the source of the differences and explain why we believe our approach is superior on conceptual or empirical grounds.⁴

³ The EL approach is in some ways very similar to that in this paper. EL estimate regressions of the levels of income on various measures of endowments, institutions, and “policies.” They find that institutions exert an important effect on development, while endowments do not, other than through their effect on institutions. Policies also do not exert any independent effect on development. The main differences between our paper and EL are the following. First, we use a larger sample of countries (79 and 137) to run the “horse” race between the three possible determinants. The EL sample is restricted to 72 countries. Second, EL do not test in any detail whether integration has an effect on development. For them, integration or open trade policy is part of a wider set of government policies that can affect development. Testing for the effect of policies in level regressions is, however, problematic as discussed in greater detail below. Policies pursued over a short time span, say 30-40 years, are like a flow variable, whereas development, the result of a much longer cumulative historical process, is more akin to a stock variable. Thus, level regressions that use policies as regressors conflate stocks and flows.

⁴ We note that many of the papers already cited as well as others have carried out similar robustness tests. For example, AJR (2001) document that geographic variables such as temperature, humidity, malaria risk exert no independent direct effects on income once institutions are controlled for. A follow-up paper by the same authors (Acemoglu, Johnson
(continued)

One final word about policy. As we shall emphasize at the end of the paper, identifying the deeper determinants of prosperity does not guarantee that we are left with clearcut policy implications. For example, finding that the “rule of law” is causally implicated in development does not mean that we actually know how to increase it under the specific conditions of individual countries. Nor would finding that “geography matters” necessarily imply geographic determinism—it may simply help reveal the roadblocks around which policy makers need to navigate. The research agenda to which this paper contributes is one that clarifies the priority of pursuing different objectives—improving the quality of domestic institutions, achieving integration into the world economy, or overcoming geographical adversity—but says very little about how each one of these is best achieved.

The plan of the paper is as follows. Section II presents the benchmark results and robustness tests. Section III provides a more in-depth interpretation of our results and lays out a research agenda.

II. Core Results and Robustness

A. Data and Descriptive Statistics

Table 1 provides descriptive statistics for the key variables of interest. The first column covers the sample of 79 countries for which data on settler mortality have been compiled by AJR.⁵ Given the demonstrated attractiveness of this variable as an instrument that can help illuminate causality, this will constitute our preferred sample. The second column contains summary statistics for a larger sample of 137 countries for which we have data on alternative instruments for institutions (fractions of the population speaking English and other European languages). Data for the FR instrument on trade, on which we will rely heavily, are also available for this larger sample.

GDP per capita on a PPP basis for 1995 will be our measure of economic performance. For both samples, there is substantial variation in GDP per capita: for the 79-country sample, mean GDP in 1995 is \$3072, the standard deviation of log GDP is 1.05, with the poorest country’s (Congo, DRC) GDP being \$321 and that of the richest (Singapore) \$ 28,039. For

and Robinson 2003) shows that macroeconomic policies have limited effects after institutions are controlled. Easterly and Levine (2003) produce similar robustness results on the geography front. Our contribution is to put these and other tests in a broader framework, including trade, and to provide an interpretation of the results which we think is more appropriate.

⁵ AJR actually compiled data on settler mortality for 81 countries, but data on our other variables are unavailable for Afghanistan (for per capita PPP GDP for 1995) and the Central African Republic (for rule of law).

the larger sample, mean income is \$4492, the standard deviation is 1.14, with the richest country (Luxembourg) enjoying an income level of \$34,698.

The institutional quality measure that we use is due to Kaufmann, Kraay, and Zoido-Lobaton (2002). This is a composite indicator of a number of elements that capture the protection afforded to property rights as well as the strength of the rule of law.⁶ This is a standardized measure that varies between -2.5 (weakest institutions) and 2.5 (strongest institutions). In our sample of 79 countries, the mean score is -0.25, with Zaire (score of -2.09) having the weakest institutions and Singapore (score of 1.85) the strongest.

Integration, measured using the ratio of trade to GDP, also varies substantially in our sample. The average ratio is 51.4 percent, with the least “open” country (India) posting a ratio of 13 percent and the most “open” (Singapore) a ratio of 324 percent. Our preferred measure of geography is a country’s distance from the equator (measured in degrees). The typical country is about 15.4 degrees away from the equator.

B. OLS and IV Results in the core specifications

Our paper represents an attempt to estimate the following equation:

$$\log y_i = \mu + \alpha INS_i + \beta INT_i + \gamma GEO_i + \varepsilon_i \quad (1)$$

where y_i is income per capita in country i , INS_i , INT_i , and GEO_i are respectively measures for institutions, integration, and geography, and ε_i is the random error term. Throughout the paper, we will be interested in the size, sign, and significance of the three coefficients α , β , and γ . We will use standardized measures of INS_i , INT_i , and GEO_i in our core regressions, so that the estimated coefficients can be directly compared.

Before we discuss the benchmark results, it is useful to look at the simple, bivariate relationships between income and each of the “deep determinants.” Figure 2 shows these scatter plots, with the three panels on the left hand side corresponding to the sample of 79 countries and the three panels on the right to the larger sample of 137 countries. All the plots show a clear and unambiguously positive relationship between income and its possible determinants. Thus, any or all of them have the potential to explain levels of income. This positive relationship is confirmed by the simple OLS regression of equation (1) reported in column (6) of Table 2A. The signs of institution, openness, and geography are as expected and statistically significant or close to being so. Countries with stronger institutions, more

⁶ AJR use an index of protection against expropriation compiled by Political Risk Services. The advantage of the rule of law measure used in this paper is that it is available for a larger sample of countries, and in principle captures more elements that go toward determining institutional quality. In any case, measures of institutional quality are highly correlated: in our 79-country sample, the two measures have a simple correlation of 0.78.

open economies, and more distant from the equator are likely to have higher levels of income.

To get a sense of the magnitude of the potential impacts, we can compare two countries, say Nigeria and Mauritius, both in Africa. If the OLS relationship is indeed causal, the coefficients in column (6) of Table 2A would suggest that Mauritius's per capita GDP should be 10.3 times that of Nigeria, of which 77 percent would be due to better institutions, 9 percent due to greater openness, and 14 percent due to better location. In practice, Mauritius's income (\$11,400) is 14.8 times that of Nigeria (\$770).

Of course, for a number of reasons described extensively in the literature—reverse causality, omitted variables bias, and measurement error—the above relationship cannot be interpreted as causal or accurate. To address these problems, we employ a two-stage least squares estimation procedure. The identification strategy is to use the AJR settler mortality measure as an instrument for institutions and the FR measure of constructed trade shares as an instrument for integration. In the first-stage regressions, INS_i and INT_i are regressed on all the exogenous variables. Thus:

$$INS_i = \lambda + \delta SM_i + \phi CONST_i + \psi GEO_i + \varepsilon_{INS_i} \quad (2)$$

$$INT_i = \theta + \sigma CONST_i + \tau SM_i + \omega GEO_i + \varepsilon_{INT_i} \quad (3)$$

where SM_i refers to settler mortality and $CONST_i$ to the FR instrument for trade/GDP. The exclusion restrictions are that SM_i and $CONST_i$ do not appear in equation 1.

Equations (1)-(3) are our core specification. This specification represents, we believe, the most natural framework for estimating the respective impacts of our three deep determinants. It is general, yet simple, and treats each of the three deep determinants symmetrically, giving them all an equal chance. Our proxies for institutions, integration, and geography are the ones that the advocates of each approach have used. Our instruments for institutions and integration are sensible, and have already been demonstrated to “work” in the sense of producing strong second-stage results (albeit in estimations not embedded in our broader framework).

Panel A of Table 2B reports the two-stage least squares estimates of the three coefficients of interest. The estimation is done for three samples of countries: (i) for the sample of 64 countries analyzed by AJR; (ii) for an extended sample of 79 countries for which AJR had compiled data on settler mortality; and (iii) for a larger sample of 137 countries that includes those that were not colonized. In AJR, the quality of institutions was measured by an index of protection against expropriation. We use a rule of law index because it is available for a larger sample. The IV estimates of the coefficient on institutions in the first three columns of Panel A are very similar to those in AJR, confirming that these two indexes are capturing broadly similar aspects of institutions, and allowing us to use the larger sample for which data on settler mortality are available.

Columns (4)-(6) report our estimates for the extended AJR sample (which as we shall explain below will be our preferred sample in this paper). Columns (5) and (6) confirm the importance of institutions in explaining the cross-country variation in development. Once the institutional variable is added, geography and openness do not have any additional power in explaining development. Institutions trump geography and openness. In our preferred specification (column (6)), not only are institutions significant, their impact is large, and the estimated coefficients on geography and openness have the “wrong” sign! The coefficient on institutions in the IV estimation is nearly three times as large as in the corresponding OLS estimation (2 versus 0.7), suggesting that the attenuation bias from measurement error in the institution variables swamps the reverse causality bias that would tend to make the OLS estimates greater than the IV estimates.

The results are similar for the larger sample of countries (Panel A, columns (6) to (9)). In this sample, we follow Hall and Jones (1999) in using the following two variables as instruments for institutional quality (in lieu of settler mortality): ENGFRAC, fraction of the population speaking English, and EURFRAC, fraction of the population speaking other European languages. Once again, institutions trump geography and openness, although the size of the estimated coefficient is smaller than that for the smaller sample. Figure 3 plots the conditional relationship between income and each of the three determinants for the 79-country (left panels) and 137-country (right panels) samples. In contrast to Figure 2, which showed a positive partial relationship between income and all its determinants, Figure 3 shows that only institutions have a significant and positive effect on income once the endogenous determinants are instrumented.⁷

The first-stage regressions (reported in Panel B) are also interesting. In our preferred specification, settler mortality has a significant effect on integration: the coefficient is correctly signed and significant at the 1 percent level. This result holds for the range of specifications that we estimate as part of the robustness checks reported below. The geography variable has a significant impact in determining the quality of institutions as does integration, although its coefficient is significant only at the 5 percent level. The table also reports a number of diagnostic statistics on weak instruments. These provide little evidence that our results suffer from the presence of weak instruments. The F-statistic for both first-stage regressions is well above the threshold of 10 suggested by Staiger and Stock (1997);⁸

⁷ The finding that trade enters with a negative (albeit insignificant) coefficient may be considered puzzling. In further results (available from the authors), we found that this is due largely to the adverse effects of trade in primary products. When total trade is broken into manufactures and non-manufactures components, it is only the latter that enters with a negative coefficient.

⁸ Although this threshold applies strictly to the case where there is a single endogenous regressor, it is nevertheless reassuring that our specifications yield F-statistics well above it.

the partial R-squares are reasonable; and the correlation between the fitted values of the endogenous variables appears to be sufficiently low.⁹

While all three samples provide qualitatively similar results, our preferred sample will be the 79-country sample: obviously this sample Pareto-dominates the 64-country sample. We also prefer this sample to the 137-country sample because settler mortality appears to be a superior instrument to those used in the 137-country sample (ENGFRAC and EURFRAC). Panel B shows that the instruments for the IV regressions in the 137-country sample fail to pass the over-identification tests despite the well-known problems of these tests having low power. Indeed, this turns out to be true not just for the core specifications in Table 2B, but for many of the robustness tests that we discuss below. Thus, while it is reassuring that the main result regarding the primacy of institutions also holds in the larger sample, we will focus mainly on the 79-country sample in the rest of the paper (referring to results for the larger sample in passing).¹⁰ We shall examine the robustness of our main results in the next section.

Table 2C illustrates the inter-relationships between integration and institutions in the 79-country sample. We regress trade and institutional quality separately on geography and on each other (instrumenting the endogenous variables in the manner discussed previously). While it is possible to envisage non-linear relationships among these determinants—trade may have sometimes positive and sometimes negative effects on the quality of institutions, for example—we keep the specifications simple and linear as in all our core specifications. The IV regressions show that each of these exerts a positive impact on the other, with the larger quantitative and statistically significant impact being that of institutional quality on trade. A unit increase in institutional quality increases the trade share by 0.45 units, while a unit increase in trade increases institutional quality by 0.22 units.¹¹

Taking these indirect effects into account, we can calculate the total impacts on incomes of these two determinants by combining the estimated parameters. Our estimates of α and β (the direct effects) in our preferred sample and specification are 1.98 and -0.31 , respectively (column 6). We can solve the system of equations implied by the additional results in

⁹ In Appendix A, we explain that our core specification also passes a more formal test (suggested by Stock and Yogo, 2002) for weak instruments in the presence of two endogenous regressors.

¹⁰ We emphasize that we haven't found an instance in which the use of one sample or another makes a qualitative difference to our results.

¹¹ Breaking trade into manufactures and non-manufactures components as before, we find that it is only manufactures trade that has significant positive effect on institutional quality (results are available from the authors).

columns (1) and (2) of Table 2C to calculate the total effects on log incomes of “shocks” to the error terms in the institution and trade equations.¹²

The results are as follows. If we consider the point estimates in column (6) of Table 2B and in columns (1) and (2) in Table 2C as our best estimate of the various effects, a unit (positive) shock to the institutional quality equation ultimately produces an increase in log incomes of 1.85; a unit (positive) shock to the trade equation ultimately produces an increase in log incomes of 0.09. This is a twenty two-fold difference. Alternatively, we could consider only those impacts that are statistically significant. Under this assumption, a unit shock to the institutional quality equation is the estimate from column (6), namely 1.98. The corresponding unit shock to the trade equation has no impact on income at all. Institutions overwhelmingly trump integration.

The much greater impact of institutions is the consequence of four factors: (i) the estimated direct effect of institutions on incomes is positive and large; (ii) the estimated direct effect of trade on incomes is negative (but statistically insignificant); and (iii) the estimated indirect effect of trade on institutions is positive, but small and statistically insignificant; and (iv) the estimated indirect effect of institutions on trade is large and statistically significant but this has either a negative or no impact on incomes because of (ii).

Repeating this exercise, and taking into account only the statistically significant coefficients, we find that the total impact of a unit shock to geography on income is about 1.49, only a quarter less than that of institutions. The large impact of geography stems from the sizable indirect impact that it has in determining institutional quality (coefficient of 0.75 in Panel B, column 1 of Table 2C).¹³

We next analyze the channels through which the deep determinants influence incomes. The proximate determinants of economic growth are accumulation (physical and human) and productivity change. How do the deep determinants influence these channels? To answer this question, we regressed income per worker and its three proximate determinants, physical capital per worker, human capital per worker, and total factor productivity (strictly speaking a labor-augmenting technological progress parameter) on the deep determinants. Data for the left hand side variables for these regressions (i.e. income, physical, and human capital per worker, and factor productivity) are taken from Hall and Jones (1999). These results are

¹² Note that these calculations omit the feedback effect from income to trade and institutions, since we are unable to estimate these. Our numbers can hence be viewed as impact effects, taking both direct and indirect channels into account, but ignoring the feedback from income.

¹³ In light of our quantitative estimates, our main difference with Sachs (2003) seems to relate to whether the sizable effects of geography are direct (the Sachs position) or indirect, operating via institutions (our position).

reported in Table 3 for both the 79-country sample (columns 1-4) and the 137-country sample (columns 5-8).¹⁴ Three features stand out.

First, the regression for income per worker is very similar to the regressions for per capita income reported in Table 2, with institutions exerting a positive and significant effect on income, while integration and geography remain insignificant. Second, and interestingly, the same pattern holds broadly for the accumulation and productivity regressions; that is, institutions are an important determinant of both accumulation and productivity, while integration and geography are not influential in determining either accumulation or productivity.¹⁵ Finally, it is interesting to note that institutions have a quantitatively larger impact on physical accumulation than on human capital accumulation or productivity; for example, in the 79-country sample the coefficient on physical capital accumulation is about six times greater than on human capital accumulation and about 3.2 times greater than on productivity. One possible interpretation is that these results emphasize the particularly important role that institutions play in preventing expropriability of property which serves as a powerful incentive to invest and accumulate physical capital.

C. Robustness checks

Tables 4, 5, and 6 present our robustness checks. In Table 4 we test whether our results are driven by certain influential observations or by the 4 neo-European countries in our sample (Australia, Canada, New Zealand, and Australia), which are arguably different from the rest of the countries included. We also check to see whether the inclusion of regional dummies affects the results.

In columns (1)* and (1)** of Table 4 we use the Belsey-Kuh-Welsch (1980) test to check whether individual observations exert unusual leverage on the coefficient estimates, discarding those which do so. In the specification without regional dummies ((1)*), two observations—Ethiopia and Singapore—are influential. Once these are dropped, the coefficient estimate for institutions not only remains statistically unaffected, but increases in magnitude. In the equation with regional dummies, the test requires the observation for Ethiopia to be omitted, and the revised specification (column (1)**) yields results very similar to the baseline specification, with the coefficient estimate on institutions remaining strong and significant. The inclusion of regional dummies for Latin America, Sub-Saharan Africa, and Asia tends to lower somewhat the estimated coefficient on institutions, but its significance level remains unaffected. Note also that none of the regional dummies enters significantly, which is reassuring regarding the soundness of our parsimonious specification.

¹⁴ Actual sample sizes are smaller than for our core specifications because of the unavailability of data for some countries in the Hall and Jones (1998) data set.

¹⁵ In the larger sample, integration has a negative and significant effect on income and accumulation but this result is not robust to the inclusion of additional variables such as land and area.

The tests for influential observations suggest that there is no statistical basis for discarding neo-European countries. Nevertheless to confirm that these countries are not driving the results, we re-estimated the baseline specification without these observations. As the column labeled (1)*** confirms, the coefficient estimates are unaffected; indeed, once again the size of the coefficient on institutions rises substantially, suggesting the greater importance of institutions for the non-neo-European colonized countries. The remaining columns (columns (2)* and (2)**) confirm that our results are robust also for the larger sample of countries.

We then check whether our results are robust to the inclusion of dummies for legal origin (column (3)), for the identity of colonizer (column (4)), and religion (column (5)). La Porta et. al. (1999) argue that the type of legal system historically adopted in a country or imported through colonization has an important bearing on the development of institutions and hence on income levels. Similar claims are made on behalf of the other variables. In all cases, while these variables themselves tend to be individually and in some cases jointly significant, their inclusion does not affect the core results about the importance of institutions and the lack of any direct impact of geography and integration on incomes. Indeed, controlling for these other variables, the coefficient of the institutions variable increases: for example, in the 79-country sample, this coefficient increases from 2 in the baseline to 2.43 when the legal origin dummies are included.¹⁶

In Table 5 we check whether our particular choice of measure for geography (distance from the equator) influences our results. We successively substitute in our baseline specification a number of alternative measures of geography used in the literature. These include percent of a country's land area in the tropics (TROPICS), access to the sea (ACCESS), number of frost days per month in winter (FROSTDAYS), the area covered by frost (FROSTAREA), whether a country is an oil exporter (OIL), and mean temperature (MEAN TEMPERATURE). (Recall that we had already introduced regional dummies as part of our basic robustness check above.) The variables FROSTDAYS and FROSTAREA are taken from Masters and McMillan (2001), who argue that the key disadvantage faced by tropical countries is the absence of winter frost. (Frost kills pests, pathogens and parasites, thereby improving human health and agricultural productivity.) We find that none of these variables, with the exception of the oil dummy, is statistically significant in determining incomes. Equally importantly, they do not change qualitatively our estimates of the institution variable,

¹⁶ We do not report the results for the larger sample but they are very similar. For the 79-country sample, interesting results are obtained for some of the individual legal origin and other variables. For example, as in AJR (2001), the French legal origin dummy has a positive total effect on incomes; the total impact of having been colonized by the UK is negative and statistically significant even though former UK-colonies have better quality of institutions on average. As for religion, suffice it to say that Weber is not vindicated!

which remains significant, nor of the integration variable, which remains insignificant and “wrongly” signed.¹⁷

In response to the findings reported in an earlier version of this paper, Sachs (2003) has produced new empirical estimates which attribute a more significant causal role to geography. Arguing that distance from equator is a poor proxy for geographical advantage, Sachs uses two explanatory variables related to malaria incidence. The first of these is an estimate of the proportion of a country’s population that lives with risk of malaria transmission (MAL94P), while the second multiplies MAL94P by an estimate of the proportion of malaria cases that involve the fatal species, *Plasmodium falciparum* (MALFAL). Since malaria incidence is an endogenous variable, Sachs instruments for both of these using an index of “malaria ecology” (ME) taken from Kiszewski et al. (2003). In columns (8) and (9) we add these variables to our core specification, instrumenting them in the same manner as in Sachs (2003). Our results are similar to Sachs’, namely that malaria appears to have a strong, statistically significant, and negative effect on income. Note that the statistical significance of institutional quality is unaffected by the addition of the malaria variables, something that Sachs (2003) notes as well.

We are inclined to attach somewhat less importance to these results than Sachs does. First, it is difficult to believe that malaria, which is a debilitating rather than fatal disease, can by itself exert such a strong effect on income levels. If it is a proxy for something else, it would be good to know what that something else is and measure it more directly. Second, we are a bit concerned about the endogeneity of the instrument (ME). Sachs (2003, 7) asserts that ME is exogenous because it is “built upon climatological and vector conditions on a country-by-country basis,” but he does not go into much further detail. The original source for the index (Kiszewski et al. 2003), written for a public health audience, has no discussion of exogeneity at all. Third, the malaria variables are very highly correlated with location in Sub-Saharan Africa.¹⁸ The practical import of this is that it is difficult to tell the effect of malaria variables apart from those of regional dummies. This is shown in columns (10) and (11) where we add regional dummies to the specifications reported earlier. Both malaria variables now drop very far below statistical significance, while institutional quality remains significant (albeit at the 90% level in one case).

Finally, we experimented with a series of specifications (not reported) that involved interacting the different geography variables with each other as well as introducing different functional forms (for example, exponential) for them. These did not provide evidence in favor of additional significant direct effects of geography on income. Overall, we conclude

¹⁷ In most of these regressions (columns (1)-(7)), the geography variable is a significant determinant of institutions in the first stage regressions.

¹⁸ Regressing MALFAL and MAL94P, respectively, on a Sub-Saharan African dummy yields t-stats of above 10 and R²s above 0.40.

that there seems to be some, albeit modest, support for the direct impact of geography on income, although one that is not as robust as that of institutional quality.¹⁹ The first stage regressions, however, point clearly in favor of an important indirect role of geography via institutions.

In Table 6, we undertake a series of robustness checks on the side of trade. First, we check whether our results are sensitive to our omission of market size variables. Frankel and Romer (1999) argue that smaller countries tend to trade more, and that one should therefore control for country size when looking for the effect of trade on incomes. The column labeled (1) in Table 6 includes two measures of country size—area and population. These variables do not have additional explanatory power in the income equation, which is different from the results in Frankel and Romer (1999). The size and significance of the coefficient on institutions are unaffected. The coefficient on openness becomes positive, but is highly insignificant. Column (3) replicates this exercise for the larger sample. The coefficient on institutions does not change qualitatively (but the standard error is sharply reduced as is the coefficient estimates), while the coefficient on openness is still negatively signed.

Alcalá and Cicconé (AC, 2002) have recently advocated the use of what they call “real openness”, which is measured as the ratio of trade to PPP GDP. They argue that this is a better measure of integration than the simple ratio of trade to nominal GDP (that FR and we favor) in the presence of trade-driven productivity change. We have a number of conceptual and empirical problems with the AC approach, which we discuss in Appendix A at the end of the paper and to which the interested reader can turn. Here we simply point out that our results are robust to the use of this alternative measure of openness. Column (5) presents the findings when we substitute the AC measure of openness for ours. This integration measure is also “wrongly” signed and insignificant, while the coefficient on institutions increases in size and remains significant, albeit at the 5 percent level.²⁰

Columns (2), (4) and (6) replicate the three robustness checks described above but with an instrument for openness that is slightly different from that in Frankel and Romer (1999). To obtain their instruments, FR estimated a gravity equation with the dependent variable defined as trade to PPP GDP. Strictly speaking therefore, theirs was an appropriate instrument for AC’s “real openness.” We re-estimated the gravity equation on the original FR sample of 63

¹⁹ See also Acemoglu, Johnson, and Robinson (2002) who document that the regions of the world that were relatively rich around 1500 underwent a “reversal of fortune” subsequently. They argue that this militates against a geographic determinist view of income difference.

²⁰ An alternative, albeit imperfect, way of taking the AC concern into account is to use a measure of trade openness that excludes services (the main source of non-tradables) from the measure of GDP in the denominator. When we use this measure of openness (trade as a share of “goods GDP”) we get similar results to those reported above (these results are available from authors).

countries, with trade to GDP as the dependent variable. We then used the coefficients from this gravity equation to construct the instrument for openness for all the 137 countries in our larger sample. The results in columns (2), (4), and (6) are very similar to those using the original FR instruments. The choice of instruments thus does not affect our main results.

Finally, in column (7) we substitute a “policy” measure for the trade variable. For reasons explained later, we believe that it is not appropriate to use policy variables in level regressions. We nevertheless sought to test the robustness of our results to one of the most-widely used measures in the trade and growth literature due to Sachs and Warner (1995), which has been endorsed recently by Krueger and Berg (2002).²¹ The results show that the institutional variable remains significant at the 5 percent level and the Sachs-Warner measure is itself wrongly signed like the other openness measures.

III. What Does It All Mean?

The present paper represents in our view the most systematic attempt to date to estimate the relationship between integration, institutions, and geography, on the one hand, and income, on the other. In this section, we evaluate and interpret our results further. This also gives us an opportunity to make some additional comments on the related literature. We group the comments under three headings. First, we argue that an instrumentation strategy should not be confused with building and testing theories. Second, we relate our discussion on institutions to the discussion on “policies.” Third, we discuss the operational implications of the results.

A. An instrument does not a theory make

Insofar as our results emphasize the supremacy of institutions, they are very close to those in AJR. Note that we have gone beyond AJR by using larger sample sizes, and by including measures of integration in our estimation. We now want to clarify a point regarding the interpretation of results. In particular, we want to stress the distinction between using an instrument to identify an exogenous source of variation in the independent variable of interest and laying out a full theory of cause and effect. In our view, this distinction is not made adequately clear in AJR and is arguably blurred by Easterly and Levine (2003).

One reading of the AJR paper, and the one strongly suggested by their title—“The Colonial Origins of Comparative Development”—is that they regard experience under the early period of colonization as a fundamental determinant of current income levels. While the AJR paper is certainly suggestive on this score, in our view this interpretation of the paper’s central message would not be entirely correct. One problem is that AJR do not carry out a direct test of the impact of colonial policies and institutions. Furthermore, if colonial experience were

²¹ The shortcomings of the Sachs-Warner index as a measure of trade policy are discussed at length in Rodriguez and Rodrik (2001).

the key determinant of income levels, how would we account for the variation in incomes among countries that had never been colonized by the Europeans?

To illustrate the second point, Figure 5 presents histograms of per-capita incomes for 163 countries for which we have data on per-capita GDP in 1995. The sample is split into two groups, a group of 103 countries that were colonized by one of the major Western European powers sometime before the twentieth century, and a group of 60 countries that were not colonized. The latter group includes some very high-income countries such as Finland and Luxembourg as well very poor countries such as Ethiopia,²² Yemen, and Mongolia. (Afghanistan is another low-income non-colonized country, but we do not have income data for it.) As the figure reveals, the dispersion of incomes within the colonized sample is not hugely different than that in the non-colonized sample. The standard deviations of log income per capita are 1.01 and 0.89 for the colonized and non-colonized samples, respectively. The income gaps that separate Ethiopia from Turkey, or China from Luxembourg are huge, and can obviously not be explained by any of these countries' colonial experience.

While colonial history does not quite provide a satisfactory account of income differences around the world, it can still provide a valid instrument. And that, in our view, is where the AJR paper is successful. An instrument is something that simply has some desirable statistical properties. It need not be a large part of the causal story.²³ To illustrate the distinction between a theory and an instrument, here is an analogy that draws on a well-known paper by Angrist and Krueger (1991).

Angrist and Krueger (1991) use quarter of birth as an instrument for the level of educational attainment, to disentangle the effects of schooling on personal earnings from those of unobserved attributes (such as “ability”). The story is that compulsory schooling requirements, requiring schooling until age 16 or 17, interacting with school-entry requirements, imply variation in the level of schooling that is correlated with quarter of birth but not with other personal attributes. The authors show for example that students born in the first quarter of the year have a systematically lower level of average schooling in the population. This is a plausible strategy for identification, but it obviously does not imply a

²² Ethiopia was included in the AJR sample of colonies, even though this country has never been colonized. (It was occupied for a period of several years by Italy during 1936-1941, but this neither counts as colonization, nor could have had much to do with the settler mortality rates from the 19th century.) Excluding Ethiopia from the AJR sample makes no difference to the basic AJR results—and in fact it improves these results, as eyeballing AJR's Figure 1 and 2 would indicate.

²³ In Acemoglu, Johnson, and Robinson (2002), the authors provide a fuller account of comparative development, emphasizing the interaction between the quality of institutions, on the one hand, and the prevailing opportunities to industrialize, on the other.

quarter-of-birth theory of earnings. Similarly, the AJR strategy does not amount to a direct test of a colonial-origins theory of development.²⁴

Easterly and Levine (2003) also assign a causal role to the settler mortality instrument and interpret it as a geographical determinant of institutions such as “crops and germs,” rather than viewing it as a device to capture the exogenous source of variation in institutions. Indeed, although they stress the role of institutions, they appear to come close to a geography theory of development. Our view is that we should not elevate settler mortality beyond its status as an instrument, and avoid favoring either a colonial view of development (as some readings of AJR would have it) or a geography-based theory of development (as some readings of EL would have it).

B. The primacy of institutional quality does not imply policy ineffectiveness

Easterly and Levine (2003) assert that (macroeconomic) policies do not have an effect on incomes, once institutions are controlled for. Our view on the effectiveness of policy is similar to that expressed in AJR (2001, 1395): there are “substantial economic gains from improving institutions, for example as in the case of Japan during the Meiji Restoration or South Korea during the 1960s” or, one may add, China since the late 1970s. The distinction between institutions and policies is murky, as these examples illustrate. The reforms that Japan, South Korea, and China undertook were policy innovations that eventually resulted in a fundamental change in the institutional underpinning of their economies.

We find it helpful to think of policy as a flow variable, in contrast to institutions, which is a stock variable. We can view institutions as the cumulative outcome of past policy actions. Letting p_i denote policy on dimension i ($i =$ fiscal, trade, monetary, etc.), I institutional quality, and δ the rate at which institutional quality decays absent countervailing action, the evolution of institutional quality over time can be written as $\dot{I} = \sum \alpha_i p_i - \delta I$, where α_i denotes the impact of policy i on institutional quality.

This suggests that it is inappropriate to regress income levels on institutional quality and policies, as Easterly and Levine (2003) do. The problem is not just that incomes move slowly while policies can take sudden turns. In principle this could be addressed by taking long-term averages of policies. (Easterly and Levine average their policy measures over a number of decades.) It is that measures of institutional quality already contain all the

²⁴ AJR themselves are somewhat ambiguous about this. They motivate settler mortality as an instrument, but then their account gravitates towards a colonial origins theory of institutional development. And their title strongly suggests that they consider the contribution of their paper to have been a theory as opposed to an identification strategy. In personal communication, one of the authors has explained that the colonial experience allows them to exploit the *exogenous* source of variation in institutions and not all the variation. The fit of the first-stage regressions of about 25 percent leaves room for most of the variation to be explained by factors other than colonization.

relevant information about the impact of policies. If the appropriate specification for income is $\ln y = \beta I + u$, the effect of policies should be sought in a regression of the form $d \ln y / dt = \beta \dot{I} + v = \alpha_0 + \beta \sum \alpha_i p_i + v$. In other words, one should look for the effect of policies in a regression of growth of income on policies.

Moreover, a geography theory of institutions can understate the impact that policies can play in changing them over time. As an empirical matter, institutions have changed remarkably in the last three decades. For example, one indicator of institutional quality—the index measuring the constraint on the executive in the Gurr Polity IV dataset, which is available on a consistent basis for several decades—shows a marked improvement between the 1970s and 1990s. For 71 countries in our core sample, this index had a mean value of 3.21 in the 1970s, 3.52 in the 1980s, and 4.37 in the 1990s. A purely geographical theory of institutions would have difficulty in accounting for these changes. Indeed, if the first stage regressions reported in Panel B of Table 2B are run over the last three decades, the coefficient on settler mortality, declines from -1.14 in the 1970s to -1.04 in the 1980s and 0.86 in the 1990s, illustrating the mutability of institutions, and the declining importance of history (on the AJR interpretation of settler mortality) or geography (on the EL interpretation of settler mortality) in explaining the cross-national variation in institutions.²⁵

C. *The hard work is still ahead*

How much guidance do our results provide to policymakers who want to improve the performance of their economies? Not much at all. Sure, it is helpful to know that geography is not destiny, or that focusing on increasing the economy's links with world markets is unlikely to yield convergence. But the operational guidance that our central result on the primacy of institutional quality yields is extremely meager.

Our indicators of institutional quality are investors' and other observers' ratings of the institutional environment. They quantify these observers' views as to the likelihood that investors will retain the fruits of their investments, the chances that the state will expropriate them, or that the legal system will protect their property rights. While it is important to know that these ratings matter—and matter a great deal in fact—it remains unclear how the underlying evaluations and perceptions can be altered. In terms of the formulation developed above, what we have estimated is β , while what policy makers need to know are the α_i (policy impacts) for the policies at their disposal. In fact, since our identification strategies rely on exogenous sources of variation in these evaluations, they are doubly unhelpful from a policy perspective.

We illustrate the difficulty of extracting policy-relevant information from our findings using the example of property rights. Obviously, the presence of clear property rights for investors is a key, if not the key, element in the institutional environment that shapes economic

²⁵ The R-squares also decline over time for these first-stage regressions.

performance. Our findings indicate that when investors believe their property rights are protected, the economy ends up richer. But nothing is implied about the actual form that property rights should take. We cannot even necessarily deduce that enacting a private property-rights regime would produce superior results compared to alternative forms of property rights.

If this seems stretching things too far, consider the experiences of China and Russia. China still retains a socialist legal system, while Russia has a regime of private property rights in place. Despite the absence of formal private property rights, Chinese entrepreneurs have felt sufficiently secure to make large investments, making that country the world's fastest growing economy over the last two decades. In Russia, by contrast, investors have felt insecure, and private investment has remained low. Our institutional quality indicators bear this out, with Russia scoring considerably lower than China despite a formal legal regime that is much more in line with European norms than China's. Credibly signaling that property rights will be protected is apparently more important than enacting them into law as a formal private property rights regime.

So our findings do not map into a determinate set of policy desiderata. Indeed, there is growing evidence that desirable institutional arrangements have a large element of context specificity, arising from differences in historical trajectories, geography, political economy, or other initial conditions. As argued in Mukand and Rodrik (2002), this could help explain why successful developing countries—China, South Korea, and Taiwan among others—have almost always combined unorthodox elements with orthodox policies. It could also account for why important institutional differences persist among the advanced countries of North America, Western Europe, and Japan—in the role of the public sector, the nature of the legal systems, corporate governance, financial markets, labor markets, and social insurance mechanisms, among others.

It is important to underscore that this does not mean economic principles work differently in different places. We need to make a distinction between economic principles and their institutional embodiment. Most first-order economic principles come institution-free. Economic ideas such as incentives, competition, hard-budget constraints, sound money, fiscal sustainability, property rights do not map directly into institutional forms. Property rights can be implemented through common law, civil law, or, for that matter, Chinese-type socialism. Competition can be maintained through a combination of free entry and laissez-faire, or through a well-functioning regulatory authority. Macroeconomic stability can be achieved under a variety of fiscal institutions. Institutional solutions that perform well in one setting may be inappropriate in other setting without the supporting norms and complementary institutions. In the words of Douglass North:

“economies that adopt the formal rules of another economy will have very different performance characteristics than the first economy because of different informal norms and enforcement. The implication is that transferring the formal political and economic rules of successful Western economies to third-world and Eastern European

economies is not a sufficient condition for good economic performance.” (North 1994, 366)

In addition, since policy makers always operate in second-best environments, optimal reform trajectories—even in apparently straightforward cases such as price reform—cannot be designed without regard to prevailing conditions and without weighting the consequences for multiple distorted margins.

Consequently, there is much to be learned still about what improving institutional quality means on the ground. This, we would like to suggest, is a wide open area of research. Cross-national studies of the present type are just a beginning that point us in the right direction.

Figure 1: The “deep” determinants of income

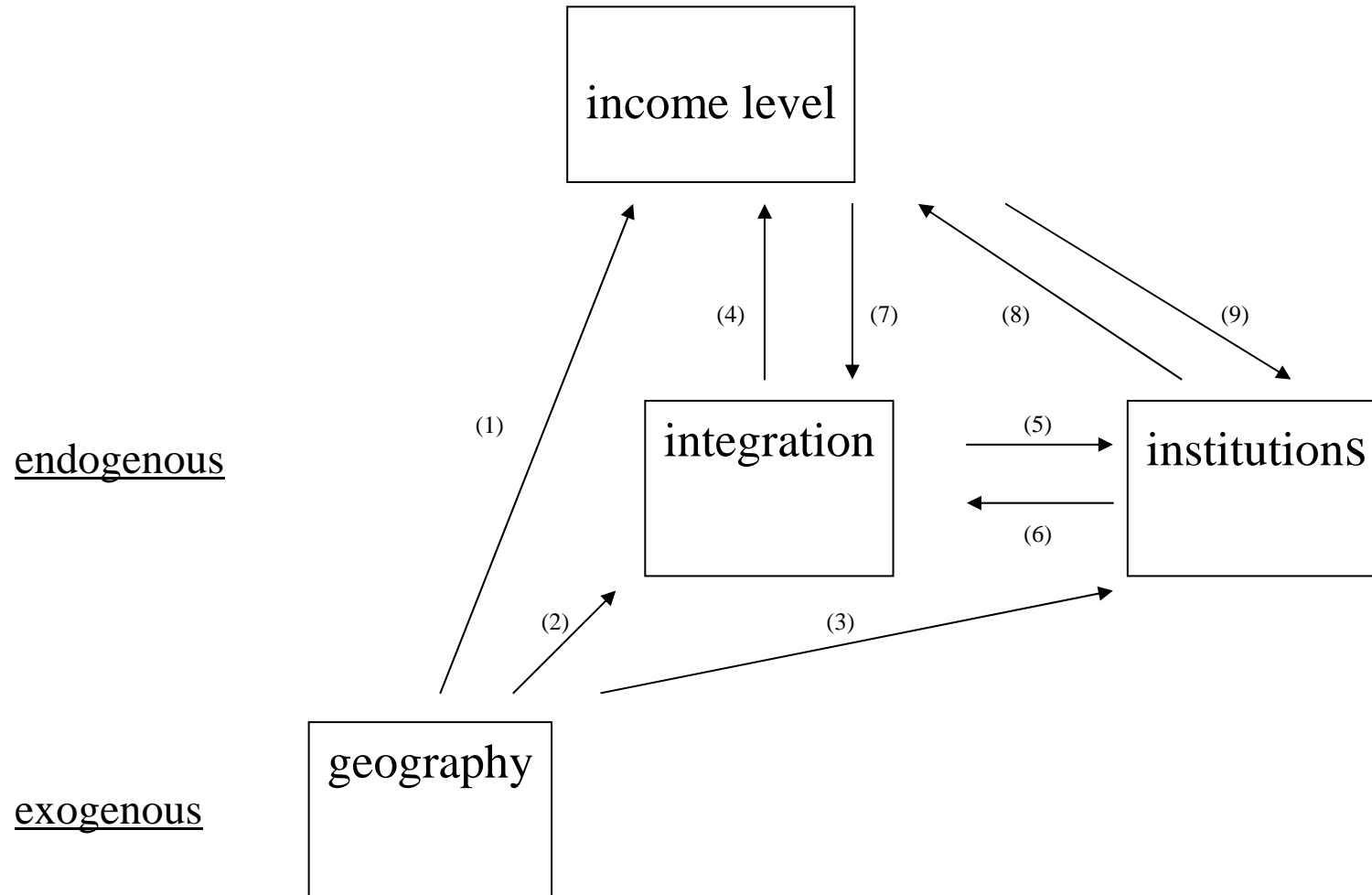


Figure 2: Simple Correlations between Income and its Determinants
(Sample of 79 countries for (a), (b), and (c); sample of 137 countries for (d), (e), and (f))

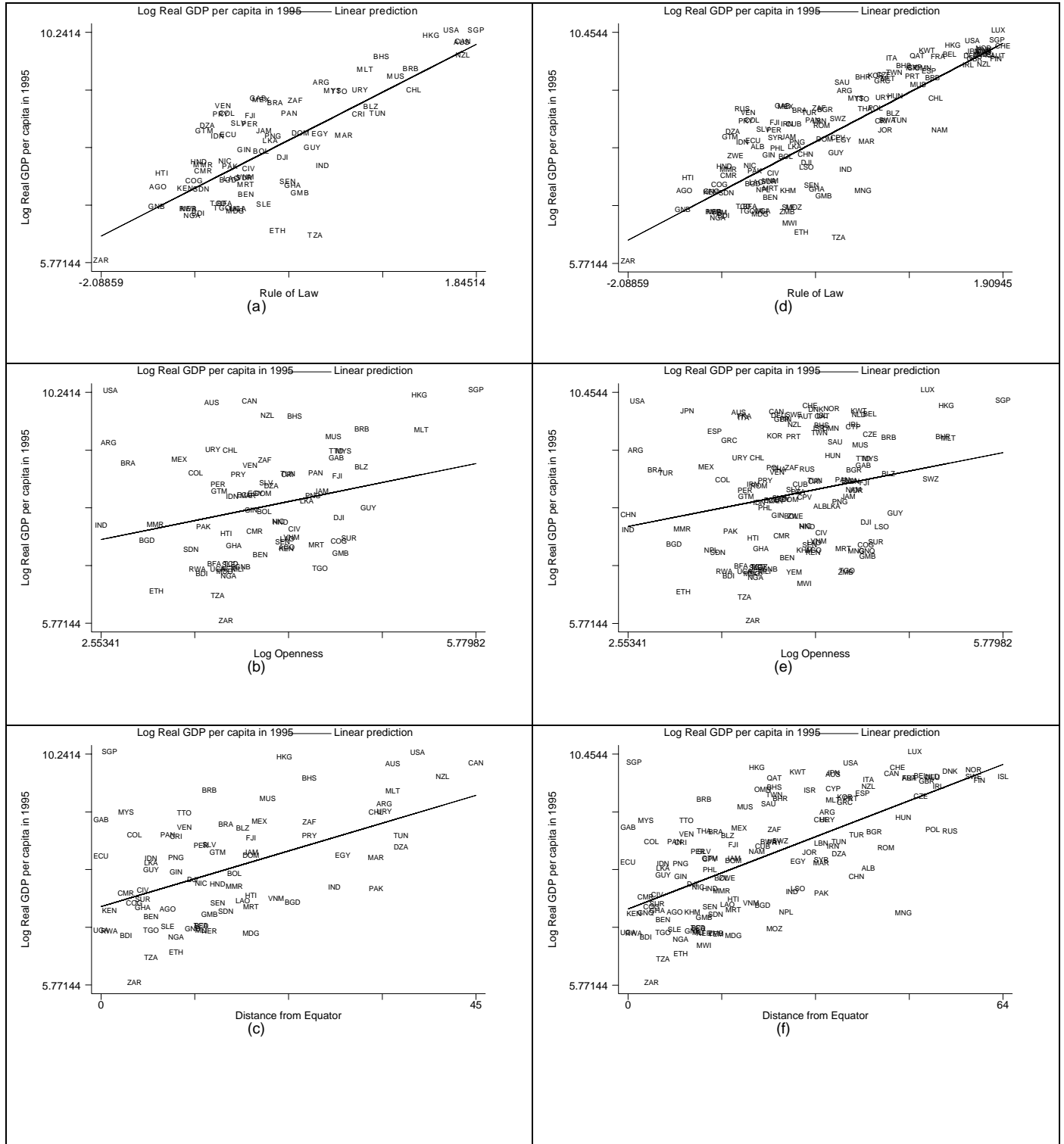


Figure 3: Conditional Correlations between Income and its Determinants
(Sample of 79 countries for (a), (b), and (c); sample of 137 countries for (d), (e), and (f))

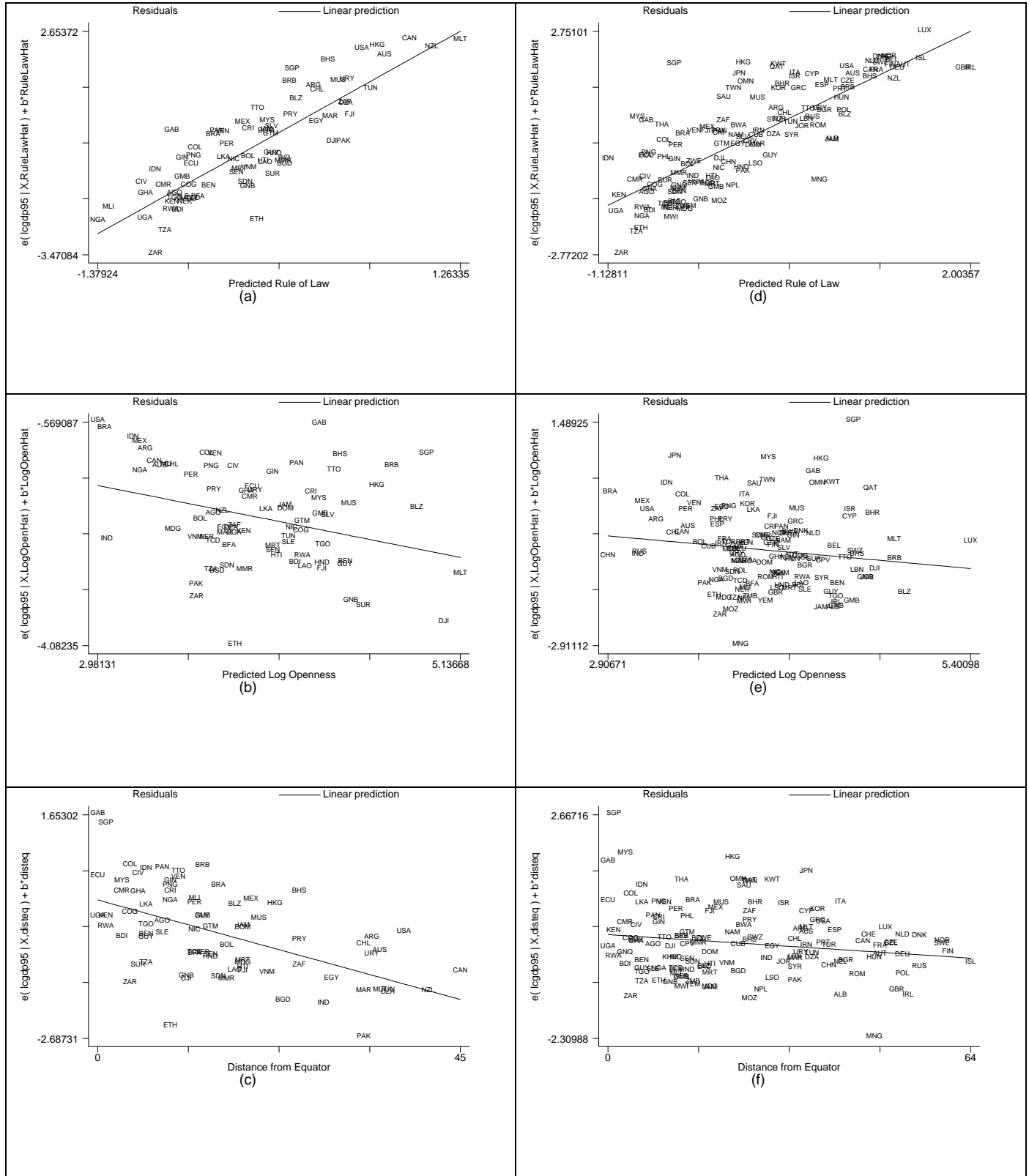


Figure 4: Distribution of incomes for colonized and non-colonized countries

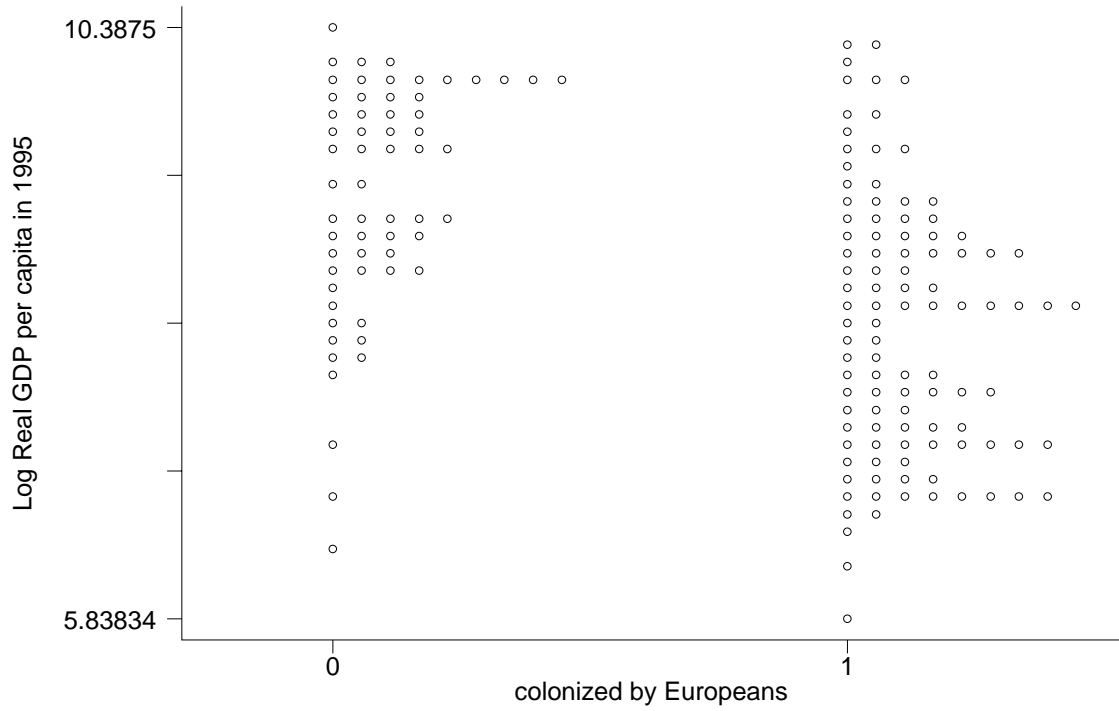


Table 1. Descriptive Statistics

	<i>Extended AJR Sample (79 countries)</i>	<i>Large Sample (137 countries)</i>
Log GDP per capita (PPP) in 1995 (LCGDP95)	8.03 <i>1.05</i>	8.41 <i>1.14</i>
Rule of law (RULE)	-0.25 <i>0.86</i>	0.08 <i>0.95</i>
Log openness (LCOPEN)	3.94 <i>0.61</i>	4.01 <i>0.57</i>
Distance from equator in degrees (DISTEQ)	15.37 <i>11.16</i>	23.98 <i>16.26</i>
Log European settler mortality (LOGEM4) (deaths per annum per 1000 population)	4.65 <i>1.22</i>
Log constructed openness (LOGFRANKROM)	2.76 <i>0.76</i>	2.91 <i>0.79</i>
Fraction of population speaking other European language (EURFRAC)	0.30 <i>0.41</i>	0.24 <i>0.39</i>
Fraction of population speaking English (ENGFRAC)	0.11 <i>0.29</i>	0.08 <i>0.24</i>

Notes: Standard deviations are reported below the means. Rule of law ranges between -2.5 and +2.5. Openness is measured as the ratio of trade to GDP. Constructed openness—the instrument for openness—is the predicted trade share and is from Frankel and Romer (1999). The Appendix describes in detail all the data and their sources.

Table 2A: Determinants of Development: Core Specifications, Ordinary Least Squares Estimates

	<i>AJR sample</i>			<i>Extended AJR sample</i>			<i>Large sample</i>		
<i>Dependent variable</i>	<i>Log GDP per capita</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Geography (DISTEQ)	0.74 (4.48) *	0.20 (1.34)	0.32 (1.85) **	0.80 (5.22) *	0.22 (1.63)	0.33 (2.11) **	0.76 (10.62) *	0.20 (2.48) **	0.23 (2.63) *
Institutions (RULE)		0.78 (7.56) *	0.69 (6.07) *		0.81 (9.35) *	0.72 (6.98) *		0.81 (12.12) *	0.78 (10.49) *
Integration (LCOPEN)			0.16 (1.48)			0.15 (1.53)			0.08 (1.24)
Observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.57	0.59	0.26	0.61	0.62	0.42	0.71	0.71

Notes: The dependent variable is per capita GDP in 1995, PPP basis. There are three samples for which the core regressions are run: (i) the first three columns correspond to the sample of 64 countries in Acemoglu, Johnson, and Robinson (2001; AJR); (ii) columns (4) to (6) use a sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by AJR; and (iii) columns (7) to (9) use a larger sample of 137 countries. The regressors are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. T-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”.

Table 2B: Determinants of Development: Core Specifications, Instrumental Variables Estimates

	<i>AJR sample</i>			<i>Extended AJR sample</i>			<i>Large sample</i>					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)			
<i>Panel A: Second-stage: Dependent variable = Log GDP per capita</i>												
Geography (DISTEQ)	0.74 (4.48) *	-0.42 (-1.19)	-0.56 (-1.23)	0.80 (5.22) *	-0.45 (-1.26)	-0.72 (-1.38)	0.76 (10.62) *	-0.06 (-0.5)	-0.14 (-0.93)			
Institutions (RULE)		1.68 (4.29) *	1.78 (3.78) *		1.75 (4.42) *	1.98 (3.56) *		1.19 (8.02) *	1.30 (7.09) *			
Integration (LCOPEN)			-0.18 (-0.71)			-0.31 (-1.10)			-0.15 (-1.09)			
No. of observations	64	64	64	79	79	79	137	137	137			
R-square	0.25	0.54	0.56	0.26	0.51	0.52	0.417	0.51	0.56			
Test for over-identifying restrictions (p-value)							(0.0089)		(0.0354)			
<i>Panel B: First Stage for Endogenous Variables (Institutions (RULE) and Integration (LCOPEN))</i>												
Dependent variable		RULE	RULE	LCOPEN		RULE	RULE	LCOPEN				
Geography (DISTEQ)		0.41 (2.8) *	0.47 (3.21) *	-0.25 (-2.00) **		0.47 (3.34) *	0.54 (3.87) *	-0.18 (-1.37)	0.67 (10.81) *	0.66 (11.23) *	-0.05 (-0.84)	
Settler mortality (LOGEM4)		-0.39 (-3.87) *	-0.40 (-4.1) *	-0.30 (-3.51) *		-0.34 (-3.69) *	-0.34 (-3.82) *	-0.27 (-3.22) *				
Population speaking English (ENGFRAC)									0.19 (2.69) *	0.18 (2.69) *	0.17 (2.65) *	
Population speaking other European languages (EURFRAC)									0.14 (1.94) **	0.17 (2.55) **	-0.11 (-1.67) **	
Constructed openness (LOGFRANKROM)			0.20 (1.95) **	0.90 (10.32) *		0.19 (2.16) **	0.80 (9.67) *		0.23 (3.99) *	0.70 (12.33) *		
F-statistic	n.a.	22.9	17.2	41.7	n.a.	24	18.5	36.9	n.a.	50.09	45.79	41.39
R-square	n.a.	0.41	0.44	0.66	n.a.	0.37	0.40	0.58	n.a.	0.52	0.57	0.54
Partial R-square			0.16	0.58			0.12	0.51		0.18		0.52
corr(RULEFIT, LCOPENFIT)				0.14				0.21				0.27

Notes: The dependent variable in Panel A is per capita GDP in 1995, PPP basis. There are three samples for which the core regressions are run: (i) the first three columns correspond to the sample of 64 countries in Acemoglu, Johnson, and Robinson (2001; AJR); (ii) columns (4) to (6) use a sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by AJR; and (iii) columns (7) to (9) use a larger sample of 137 countries for which the instrument for institutions is similar to that in Hall and Jones (1999). The regressors in Panel A are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. The dependent variables in Panel B are measures of institutions (RULE) and/or integration (LCOPEN) depending on the specification. The regressors in Panel B are: (i) DISTEQ described above; (ii) settler mortality (LOGEM4) in the first six columns; (iii) the proportion of the population of a country that speaks English (ENGFRAC) and the proportion of the population that speaks any European language (EURFRAC) in the last three columns; (iv) instrument for openness (LOGFRANKROM) obtained from Frankel and Romer (1999). All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. T-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”.

Table 2C: Inter-Relations Between Integration and Institutions

Dependent variable	<i>Extended AJR sample</i>		<i>Large Sample</i>	
	RULE	LCOPEN	RULE	LCOPEN
	(1)	(2)	(3)	(4)
<i>Panel A. OLS estimates</i>				
Geography (DISTEQ)	0.79 (5.79) *	-0.43 (-3.48) *	0.66 (10.81) *	-0.2 (-2.8) *
Institutions (RULE)		0.34 (4.17) *		0.27 (3.8) *
Integration (LCOPEN)	0.32 (3.35) *		0.27 (3.55) *	
R-square	0.42	0.23	0.52	0.12
<i>Panel B. Second stage IV estimates</i>				
Geography (DISTEQ)	0.75 (5.19) *	-0.51 (-3.08) *	0.66 (10.88) *	-0.02 (-0.16)
Institutions (RULE)		0.45 (2.33) **		0.02 (0.09)
Integration (LCOPEN)	0.22 (1.02)		0.27 (1.34)	
No. of observations	79	79	137	137
R-square	0.31	0.09	0.49	0.02
<i>Panel C. First stage IV estimates</i>				
Dependent variable	LCOPEN	RULE	LCOPEN	RULE
Geography (DISTEQ)	0.02 (0.13)	0.47 (3.34) *	-0.04 (-0.7)	0.67 (10.81) *
Settler mortality (LOGEM4)		-0.34 (-3.69) *		
Constructed openness (LOGFRANKROM)	0.79 (9.08) *		0.71 (12.35) *	
Population speaking English (ENGFRAC)				0.19 (2.69) *
Population speaking other European languages (EURFRAC)				0.14 (1.94) ***
F-statistic	44.7	24.0	76.3	50.1
R-square	0.54	0.39	0.53	0.53

Notes: The dependent variable in Panels A and B are per capita GDP in 1995, PPP basis. There are two samples for which the core regressions are run: (i) the first two columns correspond to the sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by AJR; and (ii) columns (3) and (4) use a larger sample of 137 countries for which the instrument for institutions is similar to that in Hall and Jones (1999). The regressors in Panels A and B are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. The dependent variables in Panel C are measures of institutions (RULE) and/or integration (LCOPEN) depending on the specification. The regressors in Panel C are: (i) DISTEQ described above; (ii) settler mortality (LOGEM4) in the first two columns; (iii) the proportion of the population of a country that speaks English (ENGFRAC) and the proportion of the population that speaks any European language (EURFRAC) in the last two columns; (iv) instrument for openness (LOGFRANKROM) obtained from Frankel and Romer (1999). All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. T-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”.

Table 3. Determinants of Development: Channels of Influence

Dependent variable	<i>Extended AJR sample</i>				<i>Larger sample</i>			
	Income per worker	Capital per worker	Human capital per worker	Total factor productivity	Income per worker	Capital per worker	Human capital per worker	Total factor productivity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geography (DISTEQ)	-0.97 (-1.52)	-1.72 (-1.63)	-0.26 (-1.54)	-0.33 (-1.02)	-0.25 (-1.18)	-0.38 (-1.14)	-0.05 (-1.00)	-0.13 (-0.85)
Institutions (RULE)	2.21 (3.30) *	3.39 (3.03) *	0.56 (3.14) *	1.06 (3.08) *	1.32 (5.30) *	1.90 (4.72) *	0.34 (5.64) *	0.69 (3.74) *
Integration (LCOPEN)	-0.42 (-1.36)	-0.70 (-1.30)	-0.15 (-1.86) ***	-0.13 (-0.84)	-0.30 (-1.79) **	-0.46 (-2.10) **	-0.11 (-3.00) *	-0.11 (-0.84)
R-square	0.60	0.52	0.51	0.44	0.58	0.54	0.58	0.36
No. of observations	73	73	73	73	119	119	119	119

Notes: The four dependent variables—income per worker, capital per worker, human capital per worker, and the level of total factor productivity—are expressed in natural logarithms and are from Hall and Jones (1999). IV estimates for the AJR sample use settler mortality (LOGEM4) as the instrument for institutions and EURFRAC and ENGFRAC as the instrument for the larger sample. All regressors, except RULE, are in logarithms and are scaled. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. T-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”.

Table 4. Determinants of Development: Robustness to "Influential" Observations, Neoeuropes, Legal Systems, Origin of Colonizer, and Religion

	Baseline 1	(1)*	(1)**	(1)***	(1)****	Baseline 2	(2)*	(2)**	(2)***	(3)	(4)	(5)
<i>Two-stage least squares: Dependent variable is log GDP per capita in 1995</i>												
Geography (DISTEQ)	-0.72 (-1.38)	-1.37 (-1.71) ***	-0.71 (-1.42)	-0.92 (-1.18)	-0.62 (-0.82)	-0.14 (-0.93)	-0.14 (-0.94)	0.02 (0.20)	-0.34 (-1.48)	-1.00 (-1.53)	-0.78 (-1.05)	-0.82 (-1.31)
Institutions (RULE)	1.98 (3.56) *	2.66 (3.06) *	1.86 (3.26) *	2.77 (2.45) **	1.99 (1.64)	1.30 (7.09) *	1.30 (7.14) *	0.90 (8.54) *	1.64 (5.15) *	2.43 (3.08) *	2.32 (2.46) **	2.09 (3.04) *
Integration (LCOPEN)	-0.31 (-1.10)	-0.45 (-1.12)	-0.33 (-1.26)	-0.74 (-1.31)	-0.44 (-0.80)	-0.15 (-1.09)	-0.15 (-1.02)	0.02 (0.18)	-0.31 (-1.92) **	-0.42 (-1.48)	-0.28 (-0.87)	-0.32 (-1.08)
REGIONAL DUMMIES												
Latin America (LAAM)			0.42 (1.18)		0.15 (0.28)			0.25 (1.65) ***				
Sub-Saharan Africa (SAFRICA)			-0.17 (-0.43)		-0.41 (-1.05)			-0.62 (-3.70) *				
East Asia (ASIAE)			0.22 (0.50)		0.05 (0.10)			0.13 (0.65)				
Legal origin										[0.135]		
Identity of colonizer											[0.080] ***	
Religion												[.021] **
R-square	0.52	0.56	0.65	0.44	0.63	0.56	0.59	0.68	0.55	0.54	0.55	0.59
No. of observations	79	77	78	75	75	137	136	134	133	79	79	79
Omitted observations	None	Singapore Ethiopia	Ethiopia	Australia Canada New Zealand USA	Australia Canada New Zealand USA	None	Singapore	Cuba Czech Rep. Germany	Australia Canada New Zealand USA	None	None	None

Notes: The dependent variable is per capita GDP in 1995, PPP basis. Baseline 1 corresponds to the specification in column (6) of Table 2B. Baseline 2 corresponds to the specification in column (9) of Table 2B. In columns labeled with 1 and 2 asterisks, influential observations are defined according to the Belsey, Kuh, and Welsch (1980) DFITS statistic, which requires omitting those observations for which DFITS exceeds $2(k/n)^{1/2}$, where k is the number of regressors and n is the sample size. In columns labeled with three or four asterisks, observations for Australia, Canada, New Zealand, and Canada (Neoeuropes) are omitted. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. T-statistics are reported under coefficient estimates. For legal origin, identity of colonizer, and religion, p-values for joint significance of the underlying variables (LEGFR and LEGSO for legal origin, COLUK and COLFR for colonizer's identity, and CATH, PROT, and MUSL for religion) are reported. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by "***", "**", and "*". All regressors are scaled as described in the notes to Tables 2A-2C.

Table 5. Determinants of Development: Robustness to Alternative Measures of Geography

<i>Two-stage least squares: Dependent variable is log GDP per capita in 1995</i>												
	Baseline	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Institutions (RULE)	1.98 (3.56) *	1.47 (2.96) *	1.99 (3.57) *	1.48 (6.09) *	1.45 (7.03) *	1.97 (3.39) *	1.89 (3.03) *	1.81 (4.23) *	0.65 (2.34) *	0.94 (3.90) *	0.91 (1.67) ****	1.04 (2.47) *
Integration (LCOPEN)	-0.31 (-1.10)	-0.19 (-0.91)	-0.35 (-1.00)	-0.10 (-0.52)	0.00 (0.01)	-0.42 (-1.10)	0.00 (0.01)	-0.25 (-0.65)	-0.02 (-0.01)	-0.11 (-0.71)	-0.08 (-0.46)	-0.14 (-1.00)
Geography (DISTEQ)	-0.72 (-1.38)	-0.41 (-0.97)							-0.25 (-0.91)	-0.34 (-1.35)	-0.26 (-1.07)	-0.28 (-1.11)
REGIONAL DUMMIES												
Latin America (LAAM)		0.43 (1.57)									0.25 (0.61)	0.29 (0.74)
Sub-Saharan Africa (SAFRICA)		-0.31 (-0.97)									-0.20 (-0.60)	-0.06 (0.10)
East Asia (ASIAE)		0.29 (0.83)									0.26 (0.97)	0.34 (1.22)
Area under tropics (TROPICS)			0.63 (1.45)									
Access to sea (ACCESS)				0.06 (0.19)								
Major oil exporter (OIL)					0.23 (2.10) **							
Days under frost (FROSTDAYS)						-1.07 (-1.47)						
Area under frost (FROSTAREA)							-0.62 (-1.17)					
Temperature (MEANTEMP)								0.52 (1.27)				
Malaria proportion in population (MALFAL94)									-0.69 (-5.26) *		-0.36 (-0.67)	
Malaria proportion in population*proportion of fatal species (MALFAL)										-0.47 (-5.28) *		-0.30 (-0.62)
R-square	0.52	0.63	0.54	0.52	0.54	0.53	0.51	0.64	0.52	0.52	0.63	0.63
No. of observations	79	79	76	76	67	76	66	69	75	75	75	75

Notes: The dependent variable is per capita GDP in 1995, PPP basis. Baseline corresponds to the specification in column (6) of Table 2. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. t-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”.

Table 6. Determinants of Development: Robustness to Alternative Measures of and Instruments for Integration

<i>Two-stage least squares: Dependent variable is log GDP per capita in 1995</i>								
	Baseline	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Geography (DISTEQ))	-0.72 (-1.38)	-0.56 (-0.83)	-0.63 (-0.88)	0.13 (0.38)	0.12 (0.35)	-1.16 (-1.25)	-1.24 (-1.14)	-0.86 (-1.17)
Institutions (RULE)	1.98 (3.56) *	1.83 (2.64) **	1.90 (2.58) **	0.97 (2.39) **	0.99 (2.46) **	2.70 (2.14) **	2.84 (1.82) **	2.55 (2.11) **
Integration (LCOPEN)	-0.31 (-1.38)	0.12 (0.10)	-0.01 (-0.01)	-0.87 (-0.90)	-0.85 (-0.94)			
Land area (AREA)		0.27 (0.77)	0.24 (0.67)	-0.40 (-0.97)	-0.39 (-0.99)			
Population (POP)		0.11 (0.16)	0.39 (0.05)	-0.43 (-0.63)	-0.42 (-0.65)			
"Real openness" (LNOOPEN)						-0.77 (-0.83)	-0.94 (-0.70)	
"Policy openness" (SW)								-2.04 (-1.07)
R-square	0.52	0.61	0.61	0.60	0.60	0.55	0.55	0.61
No. of observations	79	79	79	136	136	71	71	69

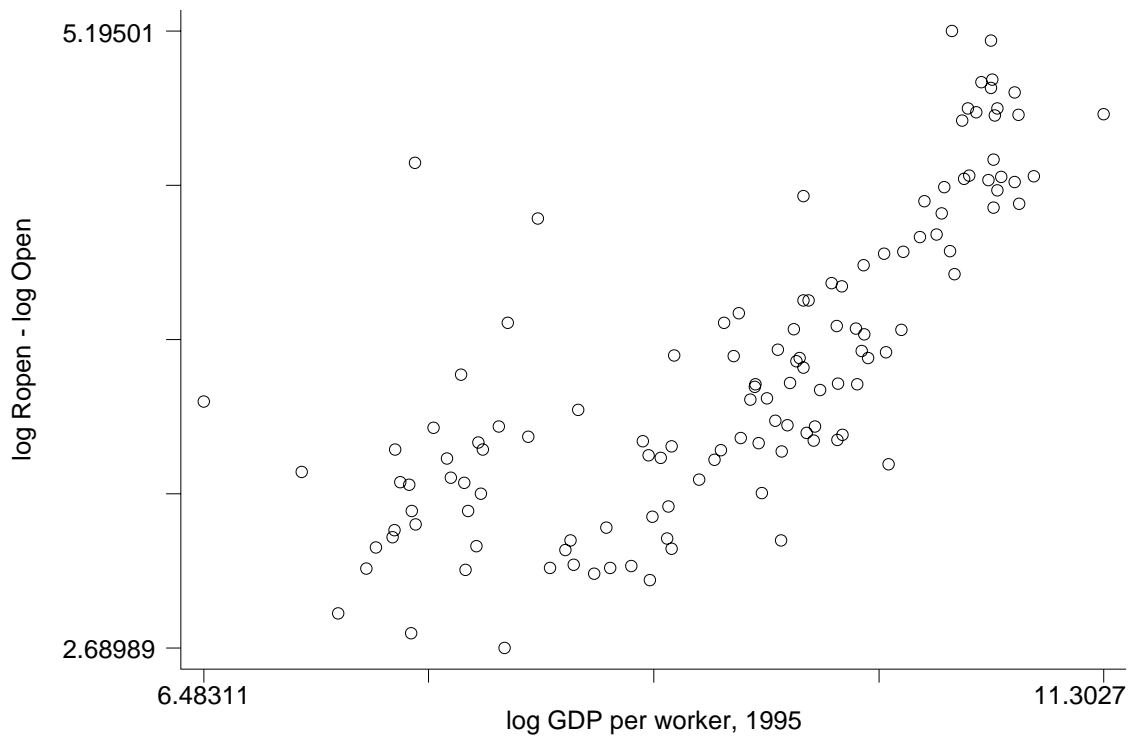
Notes: The dependent variable is per capita GDP in 1995, PPP basis. All regressors, except DISTEQ, RULE, and SW, are expressed in logs. Baseline corresponds to the specification in column (6) of Table 2B. In columns (1), (3) and (5) the instrument for openness (LOGFRANKROM) is from Frankel and Romer (1999). In columns (2), (4) and (6), the instrument for openness (LOGFRANKROMR) is derived by re-estimating the gravity equation in Frankel and Romer (1999) with the left-hand side variable defined as nominal bilateral trade to nominal GDP. In FR, the left hand side variable was defined as nominal trade divided by PPP GDP. Standard errors are corrected, using the procedure described in Frankel and Romer (1999), to take into account the fact that the openness instrument is estimated. t-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by “*”, “**”, and “***”. All regressors are scaled as described in the notes to Tables 2A-2C.

Appendix A: Problems with the use of the AC measure of “real openness”

Alcalá and Ciccone (AC, 2004) propose a different measure of integration, which they claim is superior to that used by FR and others. They argue that the conventional measure of openness—nominal trade divided by nominal GDP—can yield an estimate of trade on incomes that is biased downwards. The logic is as follows. Suppose that an increase in trade raises productivity, but that it does so predominantly in the tradables sector. Unless non-tradables are inferior in demand, this will raise the relative price of non-tradables. This will in turn tend to depress the ratio of trade to nominal GDP. The result is that the initial increase in the openness ratio will be attenuated. AC therefore prefer to use what they call “real openness,” nominal trade divided by PPP GDP.

A first pass. AC find a relationship between “real openness” and income within their empirical framework that they claim is more robust than when the conventional measure of openness is used. This seems to be the case even when institutional quality is entered, which shows up significantly in their regressions as well. Since we were unable to obtain their data set, we could not replicate their results exactly. However, as we have already discussed, the use of “real openness” within our empirical specification does not alter the central results of our paper, namely the importance of institutions and the insignificance of openness (see columns 5 and 6 of Table 6). Here we discuss a number of misgivings we have with the AC approach.

As a first step, it is useful to explain the mechanics of why “real openness” (Ropen) does much “better” than openness (Open) in AC-type regressions. Ropen and Open are linked by the identity $\log \text{Ropen} = \log \text{Open} + \log P$, where P is a country’s price level. We know from the Balassa-Samuelson argument that P has a close relationship to a country’s income/productivity level. This can be seen explicitly in the chart below, which shows the relationship between $(\log \text{Ropen} - \log \text{Open})$ and GDP per worker. (The simple correlation coefficient between the two is 0.76.)



Hence what AC are doing is to augment the standard measure of openness (Open) with a component that is highly correlated with income/productivity. This procedure is virtually guaranteed to generate a high correlation between productivity and openness. (The correlation between log Open and log P is extremely weak, 0.05, while the correlation between log Ropen and log P is 0.66.) Even after instrumentation, what ends up doing the heavy lifting in their paper appears to be the strong correlation between price levels and labor productivity.²⁶

Theoretical issues. The AC argument strikes us as being misleading on purely conceptual grounds. The use of “real openness” can yield in fact an opposite, and potentially more severe, bias. What AC do not recognize is that the actual null hypothesis that is tested is that trade does not cause

²⁶ Indeed the AC argument that the true relationship between trade and productivity can be ascertained only by holding the price level constant suggests estimating a more general framework of the kind: $\log y = \alpha + \beta_1 \log Open + \beta_2 \log P + v$. When we do so, using an instrument set close to that in AC, we find that the coefficient on openness is negative and insignificant, and that on the price level positive and highly significant. The comparable equation estimated with real openness yields a coefficient that is positive and significant. Whatever effect Ropen has on productivity, it seems to be operating via P, not via Open. So this more general framework yields little evidence that there is a significant causal effect from openness to productivity, holding the price level constant. Indeed, if we are to interpret these results literally, they suggest that causality runs from the price level to productivity.

productivity. Under that null, AC's real openness measure generates a positive correlation between income and openness that is entirely spurious. In effect, the AC adjustment has the consequence that any and all increases in the productivity of tradables, regardless of source, can produce a rise in their measure of openness. Any increase in tradables productivity, whether driven by trade or not, will raise nontradables prices at home and the price level of an economy relative to others. "Adjusting" for this by using PPP GDP as the denominator drives up measured openness. The conventional measure of openness does not suffer from that shortcoming, and hence is preferable. We explain and illustrate this point using a simple model.

Imagine a symmetric world populated with a large number of small endowment economies. Each economy i has a fixed endowment of nontraded and traded goods, denoted by N_i and T_i respectively. Let each country produce a different traded good (the Armington assumption), but consume all the varieties of traded goods produced around the world. If there is a very large number of countries, each country's consumption of its own endowment of the traded good will be negligible: (almost) all of its traded good will be exported in exchange for imports of the traded goods produced elsewhere. Let PN_i stand for the price of nontraded goods in country i and let the prices of all traded goods be fixed at unity. Since the sum of exports and imports are given by $2T_i$, conventionally measured openness in a country i can then be expressed as $ON_i = 2T_i/(PN_i*N_i + T_i)$.

Assume that preferences in each country take the Cobb-Douglas form, such that nontraded goods and traded goods (in aggregate) have fixed budget shares. Under this assumption, $2T_i/(PN_i*N_i + T_i)$ will be constant and independent of a country's endowments of T and N . (This is because $dPN_i/PN_i = dT_i/T_i - dN_i/N_i$.) Cross-country differences in conventionally measured openness, ON_i , will arise solely from differences in Cobb-Douglas budget shares. Assume also that differences in the endowment of the traded good are the only source of cross-country differences in income. That is, all countries have identical N_i but varying T_i . Countries with larger T_i are richer.

Under the above assumptions, there is no causal relationship that goes from trade to incomes. Cross-country differences in income are due entirely to differences in endowments. And if we run a regression of income on openness, we will get nothing. Trade shares either do not vary across countries, or they vary "randomly" with the Cobb-Douglas parameter. They have no systematic relationship to levels of income. So the econometrics will provide a good guide to the underlying reality.

Now suppose that we follow AC, and construct their real openness measure, OR_i . This adjustment consists of expressing the value of i 's nontraded production at some benchmark country's prices, PB , instead of domestic prices, PN_i . The AC measure of real openness is therefore $OR_i = 2T_i/(PB*N_i + T_i)$. Note that OR_i is increasing in T_i . When we correlate OR_i with incomes across countries, we will get a positive relationship. This is a spurious relationship, since the only source of productivity differences in this model is differences in endowments.

Hence, in this benchmark model, the conventional measure of openness does exactly what we would like a measure of openness to do under the null hypothesis that trade does not cause productivity. The AC variant, meanwhile, imparts a positive bias to the estimated trade-income relationship. A key feature of the model above is that the elasticity of substitution in demand between T and N is

unity. This ensures that the rise in PN is just enough to keep (conventional) openness invariant to changes in the endowment (or productivity) of tradables. When the elasticity of substitution differs from one, conventional openness does not always deliver such a helpful result, but the bias is not unidirectional. So with an elasticity of substitution greater than one, a regression of income on conventional openness will yield (misleadingly) a positive coefficient, while with an elasticity less than one, the regression will yield (misleadingly) a negative coefficient. However, the AC real openness measure is invariant to the elasticity of substitution and hence is always positively biased.

Empirics. Another point relates to the choice between real openness and openness on econometric grounds. Recall that the authors' original argument on behalf of $Ropen$ is based on the idea that there is reverse causality from productivity to $Open$, via the price level. If the Frankel-Romer constructed trade share is a valid instrument, in the sense of being uncorrelated with productivity through any channel other than trade, any type of reverse causality—positive or negative—is already taken care of. The reverse causality that AC worry about should be handled by the instrument as well! For the authors' argument to be valid, instrumentation should fail when $Open$ is used, but work when $Ropen$ is used (even though the same instruments are used in both cases). The authors do not provide any justification for this, and it is unclear to us that any justification could be produced.

Moreover, it is possible that the AC strategy does exactly the reverse and that it weakens the instrument. As we mentioned above, we were unable to obtain AC's data and could not replicate their results exactly. But in our attempted replications of their baseline specification, we repeatedly found that the first-stage F statistics were lower, sometimes substantially so, when real openness was used in lieu of openness. In fact, the F-statistic was typically below 10 when real openness was used (and always above 10 when openness was used).²⁷ A more formal test for weak instruments with two endogenous regressors (which in our case are institutions and openness) is the G statistic proposed by Stock and Yogo (2002). Our baseline specification with $Open$ yielded a value for this statistic of 4.85 exceeding the critical value of 4.58, while the specification with $Ropen$ yielded a value of 3.89.²⁸

²⁷ A little exploration reveals why the instruments work much better with openness than with real openness. The first stage regressions associated with estimating the equation described in the previous footnote, which is based on the decomposition of real openness into openness and price, show that the first-stage for the price level equation has an F-statistic of 1.92. Apparently, the instruments do much worse with real openness because of the very weak correlation between the instrument set and the price level. Another issue is why AC use such an odd instrument list, entering the levels of population and land area, as well as their logs, whereas the second-stage equation has only the logs. It is hard to defend the idea that the level of land area, say, can be safely excluded from the second stage when its log belongs in it.

²⁸ The tabulated critical values for the weak instrument test for an exactly identified problem of two endogenous variables such as ours are, respectively, 7.03, 4.58, and 3.95 for desired maximal sizes of r in a 5 percent linear restrictions Wald test of 0.10, 0.15, and 0.20.

In sum, we do not find the case for “real openness” particularly compelling. We worry that the “more robust” results that AC claim for it derive from the interaction of strong reverse causality with imperfections of the instrument.

Appendix B: Data and sources

AFRICA = Dummy variable taking value 1 if a country belongs to Africa, 0 otherwise.

ASIA = Dummy variable taking value 1 if a country belongs to Asia, 0 otherwise.

ACCESS = Dummy variable taking value 1 for countries without access to the sea, 0 otherwise.

AREA = Land area (thousands sq. mt.) Source: Frankel and Romer (1999).

ASIAE = Dummy variable taking value 1 if a country belongs to South-East Asia, 0 otherwise.

CATH = Dummy variable taking value 1 if the country’s population is predominantly catholic.

COLFR = Dummy variable taking value 1 if the colonizer was France.

COLUK = Dummy variable taking value 1 if the colonizer was England.

DISTEQ = Distance from Equator of capital city measured as $\text{abs}(\text{Latitude})/90$.

ENGFRAC = Fraction of the population speaking English. Source: Hall and Jones (1999).

EURFRAC = Fraction of the population speaking one of the major languages of Western Europe: English, French, German, Portuguese, or Spanish. Source: Hall and Jones (1999).

FROSTAREA = Proportion of land with >5 frost-days per month in winter. Source: Masters and McMillan (2001).

FROSTDAYS = Average number of frost-days per month in winter. Source: CID Harvard University (2002) from Masters and McMillan (2001).

LAAM = Dummy variable taking value 1 if a country belongs to Latin America or the Caribbean, 0 otherwise.

LCGDP95 = Natural logarithm of per capita GDP in Purchasing-Power-Parity US dollars (PPP GDP). Source: Penn World Tables, Mark 6, in 1995. For the following countries the 1996 CGDP value in the PWT was utilized: Bahrain; Kuwait; Mongolia; Oman; Qatar; Saudi Arabia; Swaziland. The 1995 value for CDGP95 in PWT, Mark 6, was obtained by dividing the 1996 observation by the corresponding percentage change in real GDP computed using International Financial Statistics (IFS, IMF 2002) nominal GDP and GDP deflator series.

LCOPEN = Natural logarithm of nominal openness. Nominal openness is given by the ratio of nominal imports plus exports relative to GDP (in exchange rate US dollars). Source: Penn World Tables (PWT), Mark 6. Average over all 1950-98 available data. For the following countries the 1996 observation available in the PWT has been utilized: Bahrain; Bahamas; Bermuda; Bhutan; Djibouti; Eritrea; Georgia; Kuwait; Laos; Mongolia; Oman; Qatar; Saudi Arabia; Sudan; Swaziland; Tajikistan; Turkmenistan. For Myanmar and Suriname LCOPEN (Average over all 1950-95 available data) is constructed from IFS (IMF) using Exports and Imports in Goods and Services in national currencies and nominal GDP in national currencies for the available years.

LFR = Dummy variable taking a value of 1 if a country has a legal system deriving from that in France.

LNOPEN = Natural logarithm of “real” openness. Real openness is given by the ratio of nominal imports plus exports to GDP in Purchasing-Power-Parity US dollars (PPP GDP). Source: Penn World Tables, Mark 5.6 and World Bank (2002).

LOGA = Labor-augmenting technological progress parameter in 1998. Source: Hall and Jones (1999)

LOGEM4 = Natural logarithm of estimated European settlers’ mortality rate. Source: Acemoglu, Johnson, and Robinson (2001)

LOGFRANKROM = Natural logarithm of predicted trade shares computed following Frankel and Romer (1999) from a bilateral trade equation with “pure geography” variables. Source: Frankel and Romer (1999).

LOGFRANKROMR = Natural logarithm of predicted trade shares computed as for LOGFRANKROM except that the dependent variable in the bilateral trade (gravity) equation is nominal trade divided by nominal GDP (both in US dollars). Source: Authors’ estimates.

LOGHL = Natural logarithm of human capital per worker in 1988. Source: Hall and Jones (1999).

LOGKL = Natural logarithm of physical capital per worker in 1988. Source: Hall and Jones (1999).

LOGYL = Natural logarithm of GDP in Purchasing-Power-Parity US dollars (PPP GDP) per worker in 1988. Source: Hall and Jones (1999).

LSO = Dummy variable taking a value of 1 if a country has a socialist legal system.

MALFAL94 = Malaria index, year 1994. Source: Gallup and Sachs (1998).

MEANTEMP = Average temperature (Celsius). Source: CID Harvard University (2002).

MUSL = Dummy variable taking value 1 if the country’s population is predominantly Muslim.

OIL = Dummy variable taking value 1 for a country being major oil exporter, 0 otherwise.

POP = Population.

PROT = Dummy variable taking value 1 if the country's population is predominantly protestant.

RULE = Rule of Law index. Refers to 2001 and approximates for 1990's institutions Source: Kaufmann, Kraay, and Zoido-Lobaton (2002).

SAFRICA = Dummy variable taking value 1 if a country belongs to Sub-Saharan Africa, 0 otherwise.

SW = Dummy variable taking value 0 if the country had $BMP = 1$, $MON = 1$, $SOC = 1$, $TAR > 0.4$, or $NTB > 0.4$; 1 otherwise. Source: Sachs and Warner (1995)

TROPICS = Percentage of tropical land area. Source: Gallup and Sachs (1998).

XCONST1970 = Constraint on the executive in the 1970s. Source: Polity IV dataset.

REFERENCES

- Acemoglu, D., S. Johnson, and J. A. Robinson. (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91,1369-1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson. (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics* 117, 1231-1294.
- Acemoglu, D., S. Johnson, J. Robinson and Y. Thaicharoen. (2003). "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," *Journal of Monetary Economics* 50, 49-123.
- Alcalá, F., and A. Ciccone. (2004). "Trade and Productivity," *Quarterly Journal of Economics*, forthcoming.
- Angrist, J. D., and A. B. Krueger. (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106, 979-1014.
- Belsey, D.A., E. Kuh, and R.E. Welsch. (1980). *Regression Diagnostics*. New York: John Wiley and Sons.
- Diamond, J. (1997). *Guns, Germs, and Steel*, New York: W.W. Norton & Co.
- Dollar, D., and A. Kraay. (2004). "Trade, Growth, and Poverty," *The Economic Journal* 114, 22-49.
- Easterly, W., and R. Levine. (2003). "Tropics, Germs, and Crops: How Endowments Influence Economic Development," *Journal of Monetary Economics* 50, 3-40.
- Engerman, S. L., and K. L. Sokoloff. (2004). "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States," *National Bureau of Economic Research Working Paper No. H0066*.
- Frankel, J., and D. Romer. (1999). "Does Trade Cause Growth?," *American Economic Review* 89, 379-399.
- Gallup, J. L. and J. D. Sachs. (1998). "The Economic Burden of Malaria," *Center for International Development, Harvard University, mimeo*.
- Gallup, J. L., J. D. Sachs, and A. D. Mellinger. (1998). "Geography and Economic Development," *National Bureau of Economic Research Working Paper No. w6849*.
- Hall, R., and C. I. Jones. (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114, 83-116.

Kaufmann, D., A. Kraay, and P. Zoido-Lobaton. (2002). "Governance Matters II—Updated Indicators for 2000/01," World Bank Policy Research Department Working Paper No. 2772, Washington D.C.

Kiszewski, A., et al. . (2003). "A Global Index Representing the Stability of Malaria Transmission," Center for International Development, Harvard University, mimeo.

Krueger, A., and A. Berg. (2002) "Trade, Growth, and Poverty," Paper presented at the 2002 World Bank Annual Conference on Development Economics, Washington D.C.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R.W. Vishny. (1998). "The Quality of Government," *Journal of Law, Economics, and Organization* 15, 1113-55.

Masters, W. A., and M. S. McMillan. (2001). "Climate and Scale in Economic Growth," *Journal of Economic Growth* 6, 167-186.

Mukand, S. and D. Rodrik. (2002). "In Search of the Holy Grail: Policy Convergence, Experimentation, and Economic Performance," Harvard University, mimeo.

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

North, D. C. (1994). "Economic Performance Through Time," *The American Economic Review* 84, 359-368.

Rodríguez, F., and D. Rodrik. (2001). "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In Ben Bernanke and Kenneth S. Rogoff (eds), *Macroeconomics Annual 2000*, Cambridge, MA: MIT Press.

Rodrik, D. (2003). "Institutions, Integration, and Geography: In Search of the Deep Determinants of Economic Growth." In Dani Rodrik, (ed), *In Search of Prosperity: Analytic Country Studies on Growth*, Princeton, NJ: Princeton University Press.

Sachs, J. D. (2001). "Tropical Underdevelopment," National Bureau of Economic Research Working Paper No. w8119.

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

Sachs, J., and A. Warner. (1995). "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity* 1, 1-118.

Sala-i-Martin, X., and A. Subramanian. (2003). "Addressing the Curse of Natural Resources: An Illustration from Nigeria," National Bureau of Economic Research Working Paper No. 9804.

Staiger, D., and J.H. Stock. (1997). "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65, 557-86.

Stock, J.H., and M. Yogo. (2002). "Testing for Weak Instruments in Linear IV Regressions." In Donald W. K. Andrews and James H. Stock (eds), *Festschrift in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, forthcoming.