1. Geography, Institutions and the Poverty of Nations

There are tremendous differences in incomes and standards of living between the rich and the poor countries of the world today. For example, average per capita income in sub-Saharan Africa is less than $1/20^{th}$ of per capita income in the United States – and this is after adjusting for differences in purchasing power, which helps African incomes. For those of us lucky enough to be living in North America or Western Europe, it is difficult even to imagine how people can survive at such income levels.

Explanations for why the economic fortunes of countries have diverged so much abound. Poor countries, such as those in sub-Saharan Africa, Central America or South Asia, often lack functioning markets, their populations are poorly educated, and their machinery and technology are outdated or nonexistent. These are, however, only proximate causes of poverty, in turn beggning the question of why these places don't have better markets, better human capital, more investments, better machinery, and better technology. There must be some fundamental causes of poverty, leading to these outcomes, and via these channels, to dire poverty.

The two main popular candidates for the fundamental causes of cross-country differences in prosperity are geography and institutions. The geography hypothesis, which has a large following both in the popular imagination and in academia, maintains that the geography, climate, and ecology of a society’s location shape both its technology and the incentives of its inhabitants.

There are at least three main versions of the geography hypothesis, each emphasizing a different mechanism for how geography affects prosperity. First, climate may be an important determinant of work effort, incentives, or even productivity. This
idea dates back at least to the famous French philosopher, Montesquieu ([1748], 1989), who wrote in his classic book *The Spirit of the Laws* in a forceful, if somewhat exaggerated way: "The heat of the climate can be so excessive that the body there will be absolutely without strength. So, prostration will pass even to the spirit; no curiosity, no noble enterprise, no generous sentiment; inclinations will all be passive there; laziness there will be happiness," and "People are…. more vigorous in cold climates…. The inhabitants of warm countries are, like old men, timorous; the people in cold countries are, like young men, brave". One of the founders of modern economics Alfred Marshall is another prominent figure who emphasized the importance of climate, arguing: "vigor depends partly on race qualities: but these, so far as they can be explained at all, seem to be chiefly due to climate" (1890, p. 195).

Second, geography may determine the technology available to a society, especially in agriculture. This view is developed by an early Nobel Prize winner in economics, Gunnar Myrdal, who wrote "serious study of the problems of underdevelopment... should take into account the climate and its impacts on soil, vegetation, animals, humans and physical assets – in short, on living conditions in economic development" (1968, volume 3, p. 2121). More recently, Jared Diamond in his bestseller, *Guns, Germs and Steel*, espouses this view, "...proximate factors behind Europe's conquest of the Americas were the differences in all aspects of technology. These differences stemmed ultimately from Eurasia's much longer history of densely populated...[societies dependent on food production]...." which was in turn determined by geographical differences between Europe and the Americas (1997, p. 358). The economist Jeffrey Sachs has also argued in favor of the importance of geography in agricultural productivity, stating that "By the start of the era of modern economic growth, if not much earlier, temperate-zone technologies were more productive than tropical-zone technologies..." (2001, p. 2).

The third variant of the geography hypothesis, popular especially over the past decade, links poverty in many areas of the world to their "disease burden," emphasizing that: "The burden of infectious disease is similarly higher in the tropics than in the temperate zones" (Sachs, 2000, p. 32). Bloom and Sachs (1998) claim that the prevalence of malaria, a disease which kills millions of people, particularly children, every year in sub-Saharan Africa, reduces the annual growth rate of sub-Saharan African economies by more than 1.3 percent a year (this is an enormous effect, implying that with malaria eradicated in 1950, income *per capita* in sub-Saharan Africa would have been double of what it is today).

In this chapter, we will argue that differences in "institutions" are far more important in understanding the divergent economic and social conditions of nations than differences in geography.

While the geography hypothesis emphasizes "forces of nature" as a primary factor in the poverty of nations, the *institutions hypothesis* is about "man-made" influences. According to this view, some societies are organized in a way that upholds the rule of law, encourages investment in machinery, in human capital, and in better technologies, facilitates broad-based participation in economic and political life by the citizens, and supports market transactions. Loosely speaking, we can refer to these societies as possessing (or as having developed) "good institutions". Three crucial elements of these good institutions are: first, enforcement of property rights for a broad cross-section of
society, so that a variety of individuals have incentives to invest and take part in
economic life; second, constraints on the actions of elites, politicians and other powerful
groups so that these people cannot expropriate the incomes and investments of others in
the society or create a highly uneven playing field; and finally, some degree of equal
opportunity for broad segments of the society, so that they can make investments,
especially in human capital, and participate in productive economic activities. These
good institutions – or *institutions of private property*, a term emphasizing the importance
of the enforcement of rule of law and property rights – contrast with conditions in many
societies of the world throughout history and today, where the rule of law is selectively
applied and property rights are nonexistent for the vast majority of the population. In
these societies the political and economic power of elites is without bounds, and only a
small fraction of citizens have access to education, investment, and production
opportunities.

Like the geography hypothesis, the institutions hypothesis has an impeccable
pedigree. It goes back at least to John Locke, Adam Smith and John Stuart Mill, and
features prominently in many current academic contributions and popular debates (e.g.,
Jones 1981). John Locke, for example, stressed the importance of property rights,
writing: "...there must of necessity be a means to appropriate them some way or other,
before they can be of any use, or at all beneficial to any particular man" ([1690], 1980, p.
10). He further argued that the main purpose of government was "the preservation of the
property of ... members of the society" (p. 47). The economist Douglass North was
awarded a Nobel Prize in part for articulating the role of institutions in understanding
economic development.

At some level, it is perhaps surprising that some societies have dysfunctional
institutions, despite the large economic and social costs that these bring. Our perspective
in this essay is that there are no compelling reasons to think that societies will naturally
gravitate towards good institutions. In fact, appreciating why this is so will be key to
understanding why institutions vary across countries. Institutions not only affect
economic prospects of nations, but are also central to the *distribution* of income among
various individuals and groups in society – in other words, institutions not only affect the
size of the social pie, but also how it is distributed. This perspective implies that a
potential change from dysfunctional and bad institutions towards better ones, which will
increase the size of the social pie, may be nonetheless *blocked* when such a change
significantly reduces the size of the slice that powerful groups receive from the pie and
when they cannot be *credibly* compensated for this loss after the change in institutions. ¹
By the same token, powerful groups will often opt for institutions that do not provide any
rights to the majority of the population so that they can extract resources or labor from
them, or monopolize the most lucrative businesses. Motivated by this reasoning, we will
refer to bad and dysfunctional institutions as *extractive institutions*, emphasizing the fact
that they are there, or were introduced in the first place, as a means of supporting the
extraction of resources by one group at the expense of the rest of the society.

In the rest of this essay, we develop the case for the importance of institutions. To
build this case, we will go back to the history of European colonization, which provides
us with a natural laboratory where, while geography remained constant, European

¹ See North (1981), Bates (1981) and Olson (1982) for a general discussion, Acemoglu and
Robinson (2000, 2002) for why elites may block beneficial institutional change because they fear
losing their politically privileged position, and Acemoglu (2003) for problems associated with the
credibility of striking deals between powerful groups and the rest of the society so as to
compensate them after institutional changes take place.
colonists radically transformed institutions in many of these societies. That institutions matter, naturally, does not imply that geography is not important. The two explanations could be complementary rather than competing. Geographic and ecological factors, for example, have undoubtedly played a major role in determining where early civilizations located and where humans migrated during their early history. Nevertheless, the evidence we discuss in this essay also suggests that the role of geography is relatively limited in understanding the sources of prosperity and poverty today.

2. Geography vs. Institutions: What We See Today

If you want to believe that geography matters, look at a world map. Locate the poorest places in the world, with per capita income levels less than 1/20th of the United States. You will find almost all of them close to the equator, in very hot regions with periodic torrential rains. If you believe, like Montesquieu, Marshall and many others, that climate matters for economic activity, then this is all supportive of that view.

Next look at some recent writings on agricultural productivity. You will see many ecologists and economists claim that the tropical areas do not have enough frost to clean the soil and are suffering from soil depletion because of heavy rains. Here seems to be evidence that, as Myrdal and others have claimed, tropical agriculture is less productive than its temperate counterpart.

Next turn to sources on tropical diseases, for example, the recent report by the World Health Organization (2001). Not surprisingly, given the word "tropical" disease, areas infested with these diseases are at the tropics and much poorer than the United States and Europe, where such diseases are entirely absent. Here seems to be evidence that the burden of disease condemns these places to poverty.

Does this evidence establish that geography is a first-order influence on prosperity? No. It is true there is a correlation between geography and prosperity, i.e., a simple statistical association. But statistical association does not prove causation. Most important, there are often omitted factors driving the associations we observe in the data.

Consider an example from the history of malaria, the quintessential tropical disease, to illustrate this point. In the nineteenth century doctors did not understand what caused malaria. To make progress towards protecting European troops stationed in the tropics, they developed an "empirical theory" of malaria by observing that people who lived or traveled close to swamps caught malaria. In other words, they turned the association between the incidence of malaria and swamps into a causal relationship, that the incidence of malaria was caused by swamps, and elaborated on this theory, by arguing that malaria was transmitted by mists, bad airs and miasmas emitted by swamps and bogs. Of course they were wrong, and a few decades later, other scientists proved that this statistical association was caused by an omitted factor, mosquitoes. Malaria is caused by parasites transmitted by mosquito bites, primarily by the mosquitoes of the genus Anopheles, which breed well in swamps, explaining the statistical association between swamps and malaria infection.

In the same way, it is quite possible that an omitted factor, some institutional feature, is the root cause of the poverty of many tropical countries, and the statistical association between geography and poverty is a mere correlation and no more.
In fact, if you want to find a similar statistical association between institutions and prosperity, there is plenty of evidence for that as well. For example, Figure 1 measures institutions in terms of the protection for entrepreneurs’ property rights – protection against expropriation risk. This is the result of assessments by Political Risk Services between 1985 and 1995, an organization which collects and compiles this information and sells it to businessmen contemplating investment in these places. A high score means a high degree of protection against expropriation. Figure 1 shows the relationship between this measure of institutions and income per capita today (more accurately, the logarithm of income per capita in 1995, adjusted for purchasing power parity differences across countries). Figure 2 uses a different measure of institutions due to Ted Gurr (1997), constraints placed on the executive in the early post-war years, more closely corresponding to our notion of constraining elites and powerful groups. A high score now means effective constraints against arbitrary actions by politicians and the executive. Both figures exhibit a strong correlation between institutions and income per capita.

But, as was the case with geography, this statistical association does not prove causation. It could once again be omitted factors, or even reverse causality, the fact that richer countries can afford better institutions, better protection against arbitrary behavior and better constitutions, which account for the associations depicted in Figures 1 and 2.

How can we make progress in distinguishing between the roles of geography and institutions as fundamental causes of prosperity and poverty? There is relatively little we can learn by looking at correlations, but a lot we can gather by going back in history and making use of the “experiments” that it offers us.

3. Natural Experiments of History

In the natural sciences, causal theories are tested by conducting controlled experiments. For example, to investigate whether Tylenol helps with headaches, we would randomly allocate a large number of otherwise similar subjects with headaches into one of two groups, either the treatment group, which will receive Tylenol, or the control group, which will receive a placebo, an apparently identical but actually inactive pill. We will then see whether there is an improvement in the headaches of the treatment group relative to the control group. If the answer is yes, subject to caveats related to statistical power, we can conclude that it is Tylenol that has the causal effect on headaches. This has to be so, since in our experiment all other conditions were kept the same between the two groups.

Controlled experiments are much harder to conduct in the social sciences. We cannot change a country’s institutions and watch what happens to the incomes and welfare of its citizens (and that's fortunate!). However, even if we cannot use controlled experiments to test what determines prosperity, history offers many natural experiments, where we can convincingly argue that one factor changes while other potential determinants of the outcomes of interest remain constant. For our focus of tracing the effects of institutions on economic prosperity, the most direct sort of natural experiment would be one where a homogeneous country is divided into two, each part with very different institutions. Fortunately for us, and unfortunately for many of its inhabitants, the history of Korea since World War II offers such a natural experiment (Yeon 1988).

Until the end of World War II, Korea was under Japanese occupation. Korean independence came shortly after the Japanese Emperor Hirohito announced the Japanese
surrender on August 15, 1945. After this date, Soviet forces entered Manchuria and North Korea and took over the control of these provinces from the Japanese. The major fear of the United States during this time period was the takeover of the entire Korea either by the Soviet Union or by communist forces under the control of the former guerrilla fighter, Kim Il Sung. U.S. authorities therefore supported the influential nationalist leader Syngman Rhee, who was in favor of separation rather than a united communist Korea.

Elections in the South were held in May 1948, amidst a widespread boycott by Koreans opposed to separation. The newly elected representatives proceeded to draft a new constitution and established the Republic of Korea to the south of the 38th parallel. The North became the Democratic People's Republic of Korea, under the control of Kim Il Sung. These two independent countries organized themselves in very different ways and adopted completely different sets of institutions. The North followed the model of Soviet socialism and the Chinese Revolution in abolishing private property of land and capital. Economic decisions were not mediated by the market, but by the communist state. The South instead maintained a system of private property and the government, and especially after the rise to power of Park Chung Hee in 1961, it attempted to use markets and incentives in order to develop the economy.

Before this "experiment" in institutional change, North and South Korea shared the same history and cultural roots. In fact, Korea exhibited an unparalleled degree of ethnic, cultural, geographic and economic homogeneity. There are few geographic distinctions between the North and South, and both share the same disease environment. For example, the CIA Factbook describes the climate of North Korea as “temperate with rainfall concentrated in summer” and that of South Korea as “temperate, with rainfall heavier in summer than winter”. In terms of terrain North Korea is characterized as consisting of “mostly hills and mountains separated by deep, narrow valleys; coastal plains wide in west, discontinuous in east,” while South Korea is “mostly hills and mountains; wide coastal plains in west and south”. In terms of natural resources North Korea is better endowed with significant reserves of coal, lead, tungsten, zinc, graphite, magnesite, iron ore, copper, gold, pyrites, salt, fluorspar, hydropower. South Korea’s natural resources are “coal, tungsten, graphite, molybdenum, lead, hydropower potential.” Both countries share the same geographic possibilities in terms of access to markets and the cost of transportation.

Other man-made initial economic conditions were also similar, and if anything, advantaged the North. For example, there was significant industrialization during the colonial period with the expansion of both Japanese and indigenous firms. Yet this development was concentrated more in the North than the South. For instance, the large Japanese zaibatsu of Noguchi, which accounted for one third of Japanese investment in Korea, was centered in the North. It built large hydroelectric plants, including the Suiho dam on the Yalu river, second in the world only to the Boulder dam on the Colorado river. It also created Nippon Chisso, the second largest chemical complex in the world that was taken over by the North Korean state. Finally, in Ch’ongjin North Korea also had the largest port on the Sea of Japan. All in all, despite some potential advantages for North, Maddison (2001) estimates that at the time of separation, North and South Korea had approximately the same income per capita.

We can therefore think of the splitting on the Koreas 50 years ago as a natural experiment that we can use to identify the causal influence of (a particular dimension of)
institutions on prosperity. Korea was split with the two halves organized in radically different ways, and with geography and many other potential determinants of economic prosperity held fixed. Thus any differences in economic performance can be attributed to differences in institutions.

Consistent with the institutions hypothesis, since separation, the two Koreas have experienced dramatically diverging paths of economic development. By the late 1960's South Korea was transformed into one of the Asian “miracle” economies, experiencing one of the most rapid surges of economic prosperity in history while North Korea stagnated. By 2000 the level of income in South Korea was $16,100 while in North Korea it was only $1,000. Thus though in 1950 there was little difference between the prosperity of the two Koreas, by 2000 the South had become a member of the Organization of Economic Cooperation and Development, the rich nations club, while the North had an income level about the same as a typical sub-Saharan African country. There is only one plausible explanation for the radically different economic experiences on the two Koreas after 1950: their very different institutions led to divergent economic outcomes.

It is also interesting to note that there has yet been little tendency for the bad institutions of North Korea to gravitate towards their better counterparts south of the border. It is possible that Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country. However, by the 1990s or even by the 1980s, it was clear that the communist economic policies in the North were not working. The continued efforts of the leadership to cling to these policies and to power can only be explained as those leaders wishing to look after their own interests at the expense of the interests of the population at large. Bad institutions are therefore kept in place, clearly not for the benefit of the society as a whole, but for the benefit of the ruling elite, and this is a pattern we encounter in most cases of institutional failures.

However convincing in its own right, the evidence from this natural experiment is not sufficient for our purposes of establishing the importance of institutions as the primary factor shaping cross-country differences in economic prosperity. First, this is only one case, and in the better-controlled experiments in the natural sciences, a relatively large sample is essential. Second, here we have an example of an extreme case, the difference between a market-oriented economy and a communist one. Few social scientists today would deny that a lengthy period of totalitarian centrally planned rule has significant economic costs. And yet, many might argue that differences in institutions among capitalist economies or among democracies are not the major factor leading to differences in their economic trajectories. To establish the major role of institutions in the prosperity and poverty of nations we need to look at a larger scale "natural experiment" in institutional divergence.

4. The Reversal of Fortune

The colonization of much of the globe by Europeans starting in the 15th century provides such a natural experiment. The colonization experience transformed the institutions in many lands conquered or controlled by Europeans, but, by and large, had no effect on their geographies. Therefore, if geography is the key factor determining the economic potential of an area or a country, the places that were rich before the arrival of the Europeans should continue to be rich after the colonization experience as well, in fact
also today. In other words, since the key determinant of prosperity remains the same, we should see a high degree of persistence in economic outcomes. If, on the other hand, it is institutions that are central, then those places where good institutions were introduced or developed should get richer compared to those where Europeans introduced or maintained extractive institutions.

Historical evidence suggests that Europeans indeed pursued very different colonization strategies, with very different associated institutions, in various colonies. At one extreme, Europeans set up extreme extractive institutions, exemplified by the Belgian colonization of the Congo, slave plantations in the Caribbean or forced labor systems in the mines of Central America. These institutions introduced neither protection for the property rights of regular citizens nor constraints on the power of elites. This is not surprising, since these institutions were designed to facilitate Europeans’ extraction of resources from the colonies.

At the other extreme, many Europeans went and settled in a number of colonies, creating settler societies, replicating, and often improving, the European form of institutions protecting private property. Primary examples of this mode of colonization include Australia, New Zealand, Canada, and the United States. The settlers in these societies also managed to place significant constraints on elites and politicians, even if they had to fight to achieve this objective. Both in North America and Australia, the plans of the British crown to develop a more hierarchical structure were thwarted by the protests, demonstrations and migrations of the lower strata of European settlers (indentured servants in North America and descendants of convicts in Australia).

So what happened to economic development after colonization? Did places that were rich before colonization remain rich, as suggested by the geography hypothesis? Or was there a systematic change in economic fortunes associated with the changes in institutions?

The historical evidence shows no evidence of the persistence suggested by the geography hypothesis. On the contrary, there is a remarkable Reversal of Fortune in economic prosperity. Societies like the Mughals in India, and the Aztecs and the Incas in America that were among the richest civilizations in 1500 are among the poorer societies of today. In contrast, countries occupying the territories of the less-developed civilizations in North America, New Zealand and Australia are now much richer than those in the lands of the Mughals, Aztecs and Incas.

The Reversal of Fortune is not confined to this comparison. Using certain proxies for prosperity before modern times, we can show that it is a much more widespread phenomenon. Two useful proxies for income per capita, especially in pre-industrial societies, are urbanization rates and population density. Only societies with a certain level of productivity in agriculture and a relatively developed system of transport and commerce could sustain large urban centers and a dense population. Figure 3 shows the relationship between income per capita and urbanization (fraction of the population living in urban centers with greater than 5,000 inhabitants) today, and demonstrates that even in the current period, there is a significant relationship between urbanization and prosperity. Naturally, high rates of urbanization do not mean that the majority of the population lived in prosperity. In fact, before the 20th century urban centers were often highly unhealthy and unequal. Nevertheless, urbanization is a good proxy for average income per capita in
society, which closely corresponds to the measure we are using to look at prosperity today.

Figures 4 and 5 show the relationship between income per capita today and urbanization rates and (log) population density in 1500. We pick 1500 since it is before European colonization had an effect on any of these societies. A strong negative relationship, indicating a reversal in the rankings in terms of economic prosperity between 1500 and today, is clear in both figures. In fact, the figures show that in 1500 the temperate areas were generally less prosperous than the tropical areas.

This reversal is prima facie evidence against the most standard versions of the geography hypothesis discussed above (we discuss some other, more "sophisticated" geography hypotheses below): it cannot be that the climate, ecology or disease environments of the tropical areas condemn them to poverty today, since these areas with the same climate, ecology and disease environments were richer than the temperate areas 500 years ago.

5. Institutions and the Reversal

Is the Reversal of Fortune consistent with the institutions hypothesis? The answer is yes. In fact, once we look at the variation in colonization strategies, we see that the Reversal of Fortune is exactly what the institutions hypothesis predicts.

European colonialism made Europeans the politically powerful group with the capability to influence institutions more than any indigenous group was able to at the time. As suggested by our discussion above, we expect Europeans to have done so not according to the interest of the society as a whole, but in order to maximize their benefits. And this is exactly what the historical evidence suggests happened.

In places where Europeans did not settle and thus did not care much about aggregate output or welfare, in places where there was a large population to be coerced and employed for cheap in mines or in agriculture, or simply taxed, in places where there was a lot to be extracted, Europeans pursued the strategy of setting up extractive institutions. In those colonies, there were no constraints on the power of the elites, that is, the Europeans themselves and their allies, and no civil or property rights for the majority of the population; in fact, many of them were forced laborers or slaves. Contrasting with this pattern, in other colonies Europeans settled in large numbers and developed the laws and institutions of the society to ensure that they themselves were protected, both in their political and economic lives. In these settler colonies, the institutions were therefore much more conducive to investment and economic growth.

This discussion also suggests that Europeans were more likely to invest in the development of institutions of private property in areas that were sparsely settled and previously relatively poor. And this is what the data show. The relatively densely settled and highly urbanized colonies ended up with extractive institutions, while sparsely-settled and non-urbanized areas received an influx of European migrants and developed institutions protecting property rights and constraining elites. European colonialism therefore led to an institutional reversal, in the sense that the richer places ended up with worse institutions.

To be fair, it is possible the Europeans did not actively introduce extractive institutions in many of these places. The structure of the Mughal, Aztec and Inca empires
were already very hierarchical, non-democratic and with power concentrated in the hands of rulers. Perhaps the Europeans simply took over these institutions. Whether this is so is secondary for our focus. What matters is that in densely-settled and relatively-developed places it was in the interests of Europeans to have extractive institutions, while in the sparsely-settled areas it was in their interests to develop institutions of private property, thus leading to the institutional reversal.

The institutional reversal combined with the institutions hypothesis predicts the Reversal of Fortune: relatively rich places got worse institutions, and if these institutions are really important, we should see them become relatively poor over time. This is exactly what we find with the Reversal of Fortune.²

We find further support for the view that the Reversal of Fortune is related to the institutional reversal and the effect of this institutional reversal on long-run growth, in the fact that there appears to be no comparable reversal among countries not colonized by Europeans between 1500 and today, and nothing of the sort in the colonized or non-colonized samples between 1000 and 1500. Something special, most probably related to changes in institutions, took place in these lands after colonization.

The timing and the nature of the reversal are also consistent with the institutions hypothesis. Figures 6 and 7 show that the previously-poor colonies surpassed the former highly-urbanized colonies starting in the late 18th and early 19th centuries, and this went hand in hand with industrialization in many of the booming areas. Figure 6 shows average urbanization in colonies with relatively low and high urbanization in 1500. The initially high-urbanization countries have higher levels of urbanization and prosperity until around 1800. At that time the initially low-urbanization countries start to grow much more rapidly and a prolonged period of divergence begins. Figure 7 shows industrial production per capita in a number of countries. Although not easy to see in the figure, there was more industry (per capita total) in India in 1750 than in the U.S. By 1860, the U.S. and other colonies with relatively good institutions, such as Australia and New Zealand, began to move ahead rapidly, and by 1953, a huge gap had already opened up.

Recall that the institutions hypothesis links incentives to invest in physical and human capital and in technology to institutions, and argues that economic prosperity results from these investments. Therefore, institutions should become more important when there are major new investment opportunities. The opportunity to industrialize is the major investment opportunity of the era. In fact, it would not be an exaggeration to say that countries that are rich today, both among the former European colonies and other countries, are those that industrialized successfully during the 19th century.

Moreover, industrialization is precisely the type of process that requires investment from various different segments of the society in new technology and commerce, market transactions supported by law and order, and a workforce investing in skills and human capital; in other words, a process that requires the protections offered by institutions of private property. In contrast, before industrialization, a country could be rich from agriculture, such as the relatively prosperous sugar colonies of Barbados, Cuba, Jamaica and Saint-Domingue, which had highly extractive institutions, concentrating power in the hands of plantation owners. These institutions were probably costly for economic performance even in these largely agricultural societies, but less costly than

² In Acemoglu, Johnson and Robinson (2002), we show that the Reversal of Fortune can be statistically accounted for by the differences in institutions during or after colonial times, further supporting the conclusion in this paragraph.
they would have been if economic activity relied on more decentralization and investment from a larger segment of society. The fact that former colonies with better institutions should industrialize and pull ahead of the rest during the 19th century is thus what the institutions hypothesis predicts.

Therefore, the institutions hypothesis is also consistent with the timing and nature of the reversal, taking place mainly in the 19th century and because societies with good institutions took advantage of the opportunity to industrialize, while those with extractive institutions failed to do so.

It is also useful to note that this evidence is not consistent with another hypothesis related to colonialism, that the reversal reflects the heavy plunder of the colonies by Europeans. This hypothesis would be an extension of the Marxist analyses of colonialism and of the development of the modern world economy. But if plunder were the cause of the reversal, we would expect the reversal to happen shortly after colonization, which witnessed the most intense plunder. Instead, it takes place mostly in the 19th century, significantly later, at least for the Americas. This indicates that the reversal is not the direct consequence of colonization per se, but results from the institutions that were put in place by the colonial powers with the aim of extracting resources.

6. Can the Geography Hypothesis Be Salvaged? Sophisticated Geography Hypotheses

We have so far documented the remarkable reversal in the rankings in terms of economic prosperity between 1500 and today among countries colonized by Europeans, and argued how this is supportive of the institutions hypothesis, and inconsistent with the persistence predicted by the geography hypothesis.

This evidence does not conclusively establish that geography had no major role in the evolution of incomes in these countries, however. It is possible to develop more sophisticated geography hypotheses predicting time-varying effects of climate, ecology or disease environments. Perhaps certain geographic characteristics that were not useful, or were even harmful, for successful economic performance in 1500 turned out to be beneficial later on.

A possible example, which can be called "the temperate drift hypothesis," argues that areas in the tropics had an early advantage, but later agricultural technologies, such as the heavy plow, crop rotation systems, domesticated animals, and high-yield crops, have favored countries in the temperate areas.

However, the evidence is not consistent with the temperate drift hypothesis. First, the reversal in relative incomes seems to be related to population density and prosperity before Europeans arrived, not to any inherent geographic characteristics of the area. Furthermore, according to the temperate drift hypothesis, the reversal should have occurred when European agricultural technology spread to the colonies. Yet, while the introduction of European agricultural techniques, at least in North America, took place earlier, as documented above, the reversal occurred mostly during the 19th century and is closely related to industrialization.

---

3 See, for example, Frank (1978) or Wallerstein (1974-1980).

4 See, for example, Bloch (1966) or White (1962).
Another related hypothesis is that certain geographic characteristics, such as the presence of coal reserves or easy access to the sea, facilitated industrialization. \(^5\) Although possible in theory, there is no evidence that geography either triggered or delayed industrialization via these channels. \(^5\) Neither the data nor history provide support for these sophisticated geography hypotheses.

Another related hypothesis is one of the main arguments of Jared Diamond's *Guns, Germs and Steel*, that the diseases, in particular measles and smallpox, brought by the Spanish conquistadors, made it easier for them to conquer the Inca and the Aztec empires. Diamond (1997, pp. 77-78) describes the general phenomenon eloquently (though his estimate of consequent population declines is at the very top of the range of available estimates): "Smallpox, measles, influenza, typhus, bubonic plague, and other infectious diseases endemic in Europe played a decisive role in European conquests, by decimating many peoples on other continents... Throughout the Americas, diseases introduced with Europeans spread from tribe to tribe in advance of the Europeans themselves, killing an estimated 95 percent of the pre-Colombian Native American population." \(^7\)

The absence of these diseases in the Americas may have helped these civilizations before 1500, but then their lack of immunity may have turned into a major disadvantage. This argument is therefore similar to a sophisticated geography hypothesis. It might even be argued that the deaths brought by European diseases were somewhat related to the relative economic decline of these areas.

It is definitely true that shortly after contact with the American Indians, deadly Eurasian diseases spread rapidly in these populations. But these astounding mortality rates among the native population were important precisely in the process of Europeans taking decisive control of these states and developing the institutions that were in their interests. Helped by these diseases, Europeans quickly shaped the social organization, institutions, and economic activities of these areas in order to profit from their colonial enterprise. The Spanish first confiscated large quantities of gold and silver, then took over existing Inca and Aztec tribute systems and set up mining based on forced labor. Other Europeans soon followed with similar extractive institutions, including slave plantations in Brazil and the Caribbean. Therefore, even if Diamond's argument is valid, it simply points out an effect of geography working via institutions, not a direct effect of geography on economic performance.

Moreover, the fates of the colonies and the New World are, on the whole, very similar to those of the Old World colonies, which were not adversely affected by European colonists' diseases. This also suggests that the diseases that Europeans brought were not crucial for the success and consequences of their colonization. Finally, once again the timing of the reversal does not provide any evidence that these early colonial experiences were crucial: the reversal happened centuries after these deaths, when places with better institutions pulled ahead because they invested and took advantage of the opportunity to industrialize.

---

\(^5\) See, for example, Pomeranz (2000) or Wrigley (1988).

\(^6\) Acemoglu, Johnson and Robinson (2002) show that these geographic characteristics are not associated with the reversal and do not play any role in explaining industrialization.

\(^7\) Mortality estimates vary widely, even among serious researchers, but a consensus view would be at least 50%; see Crosby (1972), McNeill (1976) and McEvedy and Jones (1978).
Overall, although it is possible to formulate sophisticated geography hypotheses that might be consistent with the reversal and thus salvage some role for geography as a major factor in the prosperity of nations today, none of these hypotheses seem to receive significant support from the evidence or explain the salient patterns in the data. The reversal appears to be closely related to institutions, and provides strong support to the institutions hypothesis.

7. Mortality of European Settlers: Another Source of Divergence in Institutions

We have so far seen that Europeans pursued different colonization strategies in different places, with very different associated institutions, and that a key determinant of whether they set up good institutions or not is whether they settled in large numbers. One factor explains much of the variation in settlement rates of Europeans: the disease environment they faced in the colonies.

We already mentioned in passing how the inhabitants of the New World had no immunity to the diseases brought by the Europeans. This is only part of the story. The Europeans, themselves, had no immunity to the diseases of the tropics, in particular, to malaria and yellow fever.

Yellow fever is largely eradicated today, but malaria is still endemic in many parts of sub-Saharan Africa, and as discussed above, causes the deaths of millions of children every year. Nevertheless, the majority of the adult inhabitants of areas endemic with malaria have either genetic or more often acquired immunity, ensuring that they do not die or are not incapacitated by even the most deadly strain of malaria, *falciparum* malaria. In contrast, malaria infection meant almost certain death for Europeans, especially in the 19th century before the causes and prevention of malaria were understood.8

As a result of the prevalence of yellow fever and malaria, potential European settlers and European troops faced very different mortality rates in the colonies. For example, before 1850, the annual mortality rates for a settlement size maintained at 1000 (through replacement) ranged from 8.55 in New Zealand, which was lower than in Europe at that time, to 49 in India, 130 in Jamaica, and around 500 in West Africa. These widely different mortality rates of potential settlers led to different settlement rates and to divergent institutional paths for various colonies.

Figure 8 shows a very strong association between (the log of) these mortality rates for European settlers and the measure of current institutions used in Figure 1, protection against expropriation risk between 1985 and 1995. Institutions today are much worse in places with higher settler mortality. Figure 9, in turn, shows a very strong association between these mortality rates and economic prosperity today, again as measured by income per capita: countries that had lower mortality rates for European settlers are now richer. What explain this pattern?

In Acemoglu, Johnson and Robinson (2001), we document that it reflects neither the current prevalence of malaria, nor current general health conditions, nor various geographic factors ranging from temperature to humidity, from natural resources to soil

---

quality. Instead, we develop the argument that the association shown in Figure 9 works through the effect of these mortality rates on European settlement and institutional development. In places where they faced high mortality rates, Europeans did not settle and typically introduced extractive institutions. Extractive institutions have a lot of staying power; for example, groups who benefit from using the power of the state to expropriate others will resist and attempt to block any move towards better institutions. As a result, in many cases extractive institutions persisted from colonial times to today and still adversely affect economic growth.

The idea that Figure 9 captures the effect of European settler mortality rates working via institutional development, not the direct effect of these diseases, is also supported by the mortality rates of indigenous peoples in these areas. While Europeans faced astounding death rates, the indigenous population had similar mortality rates to those of Europeans in their home countries. For example, the annual mortality rates of native troops serving in Bengal and Madras were respectively 11 and 13 in 1000, similar to, in fact lower than, the annual mortality rates of British troops serving in Britain, which were approximately 15 in 1000. In contrast, the death rates of British troops serving in these colonies were much higher because of their lack of immunity to local disease. For example, death rates in Bengal and Madras for British troops were between 70 and 170 in 1000.

That the relationship in Figure 9 does not reflect the direct effect of the disease environment is also consistent with the fact that using only information about the prevalence of yellow fever leads to similar results. Since yellow fever is largely eradicated today, this is unlikely to reflect the direct effect of yellow fever.

The advantage of exploiting this source of variation in institutions is that we can both get a rough sense of how important institutions are in explaining current differences in economic performance, and also test for possible direct effects of various geographic characteristics. The results indicate that differences in institutions across countries today account for the bulk of the differences in economic outcomes. Moreover, once we take the influence of institutions into account, none of these geographic characteristics have a significant effect on income per capita today.

Overall, the evidence both from the Reversal of Fortune and from the divergent patterns of institutional development driven by differences in European settler mortality rates points to the same conclusion: institutions have a large and quantitatively important effect on economic prosperity today. What's more, once we recognize the importance of institutions on economic performance, geography seems to play a relatively small role in the large cross-country differences in prosperity today.

8. Conclusion: Geography vs. Institutions

The evidence presented so far makes a fairly convincing case that it is institutional differences, not geographic factors, which are at the root of the very large differences in economic prosperity we observe today. Yes, countries near the tropics are poorer than those in temperate areas. But this does not reflect the effect of climate or ecology on economic outcomes, simply the fact that a key determinant of prosperity,

---

[5] Controlling for these geographic characteristics has little effect on the relationship of interest, partly because prevalence of malaria and yellow fever is not related to any simple geographic characteristics.


institutions, differs between these areas. Institutions differ, in turn, because institutions in many parts of the world today are shaped by the colonial history of these areas; Europeans were more likely to settle in the temperate areas and develop institutions encouraging investment and economic progress, and they were more likely to set up extractive institutions in tropical areas and in areas that were at the time more prosperous and densely settled, which were also typically the ones near the tropics.

Does this all mean that geography is unimportant? Yes and no. There is no evidence that geography plays a major direct role in the very large differences in income per capita and growth potential of countries today. But this does not mean that geography is unimportant. It is important in at least four major ways.

First, geography and diseases almost surely matter for economic outcomes. There can be no agriculture at the poles, and it is a truism that healthy individuals will be more productive and motivated in their work, in school and in their lives. The statement here is that the effects of geography and diseases are not a major factor in explaining the tremendous cross-country differences in prosperity – not that geography and diseases have no economic effects at all.\(^\text{12}\)

Second, that geography is not at the root of the tremendous differences in economic prosperity today does not mean that it was unimportant in history. It is quite possible that geographic differences shaped why some areas were richer than others more than 500 years ago. This must be the primary candidate for explaining why tropical areas among the colonies were more prosperous than in temperate ones in 1500. Going even further back, geographic characteristics must have been important in determining where settled agriculture developed and where humans migrated.

Third, geography could have an effect via institutions, especially during a particular historical juncture. After all, the disease environment is a geographic characteristic of many tropical areas. However, the major effect of disease environments was not direct, but indirect:\(^\text{13}\) during the particular episode of European colonization, they determined whether Europeans could settle and therefore, which types of institutions developed.\(^\text{14}\)

Finally, and most importantly, even if geography has no effect on income per capita, it does have significant effects on "social welfare", properly measured. Many parts of the world, especially many parts in the tropics, suffer from poorer health and higher mortality and morbidity than North America and Western Europe, partly because of their geographic characteristics (and partly because the corresponding diseases in North America and Europe have been eradicated as a result of the economic development of these societies!). It is important to understand the social and human costs of disease, and act upon them. Many scholars, journalists and commentators argue that we, the Western world, should invest in the health of less-developed populations and try to reduce mortality and morbidity in these areas because of the economic benefits that these

\(^{12}\) Existing evidence from micro data on the effect of health on individual economic outcomes indicates significant effects, which are quantitatively at least one order of magnitude smaller than cross-country differences in income per capita, consistent with this conclusion. See, for example, the survey in Acemoglu, Johnson and Robinson (2003).

\(^{13}\) Similarly, Stanley Engerman and Kenneth Sokoloff (1997) have emphasized how the geography of the Caribbean, which made it an ideal place for sugar production, was a key factor in the development of a plantation economy based on slavery, thus having adverse long-term economic consequences, but once again through institutions rather than directly.

\(^{14}\) Therefore, not only did these characteristics not have a direct effect, but also we should not expect them to have a universal effect on economic outcomes via their influence on institutions. Instead, they had an effect on institutional development in the context of a very specific historical episode, European colonialism. If it happened to be Indians colonizing Europe and the rest of the world, rather than the other way around, the prevalence of malaria would not have been associated with the development of extractive institutions.
investments will create. Our perspective is that we should undertake such investments on
humanitarian and social grounds. After all, we have as much reason to care about the
lives of people as their incomes. We do not need tenuous arguments that investing in
health will create economic miracles to justify such investments.

References
Commitment and Politics,” Journal of Comparative Economics, forthcoming.

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

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2002) “Reversal of Fortune:
Geography and Institutions in the Making of the Modern World Income Distribution,”

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2003) “Disease and
Development in Historical Perspective,” Journal of the European Economic Association,
forthcoming.

Acemoglu, Daron and James A. Robinson (2000) “Political Losers as a Barrier to

Perspective,” Unpublished.


Figure 1
Log income per capita in 1995 vs. perceived protection against expropriation risk, 1985-95

Figure 2
Figure 3

Log income per capita in 1995 vs. urbanization in 1995

Figure 4

Log income per capita in 1995 vs. urbanization in 1500
Figure 5
Log income per capita in 1995 vs. log population density in 1500

Figure 6
Timing of the Reversal
Urbanization in ex-colonies with low and high urbanization in 1500
(averages weighted within each group by population in 1500)
Figure 7
Reversal, Industrialization and Divergence
Industrial Production Per Capita, UK in 1900 = 100
(from Bairoch)

Figure 8
Perceived protection against expropriation risk, 1985-95 vs. log settler mortality
Figure 9
Log income per capita in 1995 vs. log settler mortality
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  
October 2002

ABSTRACT

We estimate the respective contributions of institutions, geography, and trade in determining income levels around the world, using recently developed instruments for institutions and trade. Our results indicate that the quality of institutions “trumps” everything else. Once institutions are controlled for, 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, although trade too has a positive effect on institutional quality. 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 Chad Jones, James Robinson, Will Masters, and participants at the Harvard-MIT development seminar and the Harvard econometrics workshop for their comments, Daron Acemoglu for helpful conversations, and Aart Kraay for providing us with his data. Dani Rodrik gratefully acknowledges support from the Carnegie Corporation of New York.

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).

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

---

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.

---

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
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 asccertained 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, forthcoming). 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 simple endowment of natural resources with stunted institutional development, also goes under the name of “resource curse.”

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 best suited to untangle the langle 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. 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. Importantly, integration also has a (positive) impact on institutional quality, suggesting that trade can have an indirect effect on incomes by improving institutional quality. Our results also tend to confirm the findings of Easterly and Levine (2002), namely that geography exerts a significant effect on the quality of institutions.

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—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—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. Our results tend to confirm the findings of Easterly and Levine (2002), namely that geography exerts a significant effect on the quality of institutions.

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 2, panel B, 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) an 80-country sample which is the largest sample we can use while still retaining the AJR instrument; and (c) a 140-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 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 80-country sample, mean GDP in 1995 is $3020, 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 the larger sample, mean income is $4452, 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. 2 This is a standardized

2 AJR actually compiled data on settler mortality for 81 countries, but data on per capita PPP GDP for 1995 are unavailable for Afghanistan.

3 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 (continued)
measure that varies between -2.5 (weakest institutions) and 2.5 (strongest institutions). In our sample of 80 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 51.5 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 15.2 degrees away from the equator.

B. OLS and IV Results in the core specifications

Our paper represents an attempt to estimate the following equation:

\[ \log y_i = \mu + \alpha INS_i + \beta INT_i + \gamma GEO_i + \epsilon_i \]  

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 \( \epsilon_i \) is the random error term. Throughout the paper, we will be interested in the size, sign, and significance of the three coefficients \( \alpha, \beta, \) and \( \gamma \). 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 80 countries and the three panels on the right to the larger sample of 140 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 Panel A of Table 2. The signs of institution, openness, and geography are as expected and statistically significant or close to being so. Countries with stronger institutions, more 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 Panel A in Table 2 would suggest that Mauritius’s per capita GDP should be 5.2 times that of Nigeria, of which 21 percent would be due to better institutions, 65 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:

\[ INS_i = \lambda + \delta SM_i + \phi CONST_i + \psi GEO_i + \epsilon_{INS_i} \]  

\[ INT_i = \theta + \sigma SM_i + \tau CONST_i + \omega GEO_i + \epsilon_{INT_i} \]  

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 B of Table 2 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 80 countries for which AJR had compiled data on settler mortality; and (iii) for a larger sample of 140 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 B are very similar to those in AJR, confirming that these two indexes are capturing the same concept. The estimated coefficients on 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...
institutions. The results are similar for the larger sample of countries (Panel B, columns (6) to (9)). In this sample, we follow Hall and Jones (1998) and Dollar and Kraay (2002) 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 80-country (left panels) and 140-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.

The first-stage regressions (reported in Panel C) 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.

While all three samples provide qualitatively similar results, our preferred sample will be the 80-country sample: obviously this sample Pareto-dominates the 64-country sample. We also prefer this sample to the 140-country sample because settler mortality appears to be a superior instrument to those used in the 140-country sample (ENGFRAC and EURFRAC). Panel B shows that the instruments for the IV regressions in the 140-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 2, but for many of the robustness tests that we discuss below. This raises questions about the results in Hall and Jones (1998) and in Dollar and Kraay (2002), which are based on the use of ENGFRAC and EURFRAC as instruments for institutions. 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 80-country sample in the rest of the paper (referring to results for the larger sample in passing). We shall examine the robustness of our main results in the next section.

Columns (10) and (11) show the inter-relationships between integration and institutions in the 80-country sample. We regress trade and institutional quality separately on geography and on each other (instrumenting the endogenous variables in the manner discussed previously). The IV regressions show that each of these exerts a positive impact on the other, with the larger quantitative impact being that of institutional quality on trade. A unit increase in institutional quality increases the trade share by 0.77 units, while a unit increase in trade increases institutional quality by 0.23 units. Hence these estimates suggest that integration can have an indirect effect on incomes via its effect on institutional quality.

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 $\alpha$ and $\beta$ (the direct effects) in our preferred sample and specification are 2.00 and –0.30, respectively (column 6). We can solve the system of equations implied by the additional results in columns (10) and (11) to calculate the total effects on log incomes of “shocks” to the error terms in the institutions and trade equations.

The results are as follows. If we consider the point estimates in equation (6) 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 2.15; a unit (positive) shock to the trade equation ultimately produces an increase in log incomes of 0.2. This is a ten-fold difference. Alternatively, we could consider the direct impact of trade on income to be nil, since the relevant estimate ($\beta$) is statistically indistinguishable from zero. Under this assumption, a unit shock to the institutional quality equation ultimately produces an increase in log incomes of 2, while a unit shock to the trade equation produces an increase in log incomes of 0.46. Institutions trump integration by a factor of 4.4.

The much greater impact of institutions is the consequence of three 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 effect of trade on institutions is positive, but small.

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 (1998). These results are reported in Table 3 for both the 80-country sample (columns 1-4) and the 140-country sample (columns 5-9).4 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

4 Sample sizes are reduced because of the unavailability of the independent variables for all countries.
productivity. 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 80-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 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.

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 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 80-country sample, this coefficient increases from 2 in the baseline to 2.38 when the legal origin dummies are included.

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 measures of geography used in the literature. These measures 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), prevalence of malaria (MALFAL94), and mean temperature (MEAN TEMPERATURE). 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, which remains significant, nor of the integration variable, which remains insignificant and “wrongly” signed.

In columns (9), (10), and (11), we test whether geography has an impact through a combination of effects captured by the different geography variables. In equation (9), we control jointly for distance from equator and the malaria variable. The p-value for the joint significance of the two geographical variables is well below one percent. The same happens when this specification is expanded to include the number of frost days per month in winter

---

5 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.

6 We do not report the results for the larger sample but they are very similar. For the 80-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, well, suffice it to say that Weber is not vindicated!

7 In six of the eight regressions (excluding the one with the oil dummy), the geography variable is a significant determinant of institutions in the first stage regressions.
However, when we slightly enlarge the pool to include the other two geography variables, tropical area and mean temperature, all the individual effects become insignificant as does the joint significance of all the geography variables (the corresponding p-value is 15 percent). As for individual effects, in columns (9) and (10), malaria seems to be important in explaining income differences and enters significantly at the 5 or 10 percent level. But its coefficient is about 4 times smaller than that for institutions. 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 significant direct effects of geography on income. Overall, we conclude that there seems to be some, albeit modest, support for the direct impact of geography on income. The first stage regressions, however, point clearly in favor of an important indirect role of geography via institutions.

In Table 6, we check whether our results are sensitive to our omission of market size variables, or our measures of and instruments for openness. 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 Ciccone (AC, 2002) argue that “real openness”, measured as the ratio of trade to PPP GDP is a better measure of integration than the simple ratio of trade to GDP that FR and we favor. In the next section, we examine in greater detail the merits of their argument, but here we test empirically whether this alternative measure affects our results. Column (5) presents the results. Once again, this integration measure is wrongly signed and insignificant, while the coefficient on institutions increases in size and remains significant, albeit at the 5 percent level.

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 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 140 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). 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. Recent Related Work

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. Recently a few other papers have carried out somewhat similar analyses and deserve discussion. The three papers we focus on are Easterly and Levine (EL, 2002), Alcalá and Ciccone (AC, 2002), and Dollar and Kraay (DK, 2002). Our reading of EL is that it is largely consistent with our results, although, as we shall discuss, the interpretations are somewhat different. The results reported in AC and DK are at variance with ours to a much greater extent, but in different ways. AC claim trade and institutions are both significant, while DK claim that the instrumented income-level regressions exhibit too much collinearity between the two determinants for their respective contributions to be ascertained. We will identify in this section the specific departures from the framework we laid out in this paper that account for the divergent results that these authors have found. In particular, we will show that the differences derive from choices on samples, specification, or instruments that

---

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

9 See also Frankel and Rose (2002), which builds on FR. This paper is critiqued by Rodrik (2000), who argues that the results are not robust to the exclusion of two highly influential variables (Hong Kong and Singapore) and the inclusion of institutional quality and geography as additional regressors.
we think are arbitrary. But the point of identifying the origin of the difference is to allow the reader to make up his/her own mind.

A. Easterly and Levine (EL, 2002) 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 (80 and 140) 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 confute stocks and flows. Testing for integration is less vulnerable to this critique because of the instrumentation strategy for measuring integration, which relies essentially on geography variables that are time-invariant.

Finally, we also differ from EL in the interpretation of the results. EL tend to a deterministic view of institutions, interpreting settler mortality, which is essentially an econometric instrument for capturing the exogenous source of variation in institutions, as a causal geographical determinant of institutions. As we show below, this would render institutions more immutable than they have actually proven to be.

B. Alcalá and Ciccone (AC, 2002) The key innovation in AC is the advocacy and justification of what they call “real openness” as a better measure for integration. They first note that the FR result on the significance of trade in determining income is not robust to the inclusion of distance from equator in the income equation. They then argue that part of the problem is that the conventional measure of openness that FR and others use—nominal trade divided by nominal GDP—can yield an estimate of trade on productivity 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.

AC find a relationship between “real openness” and income within their empirical framework that 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 columns 5 and 6 of Table 6 illustrate, 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.

Moreover, the AC argument strikes us as being misleading. To see why, first note that 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 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 in Appendix A using a simple model. We show, under fairly innocuous assumptions, that conventional openness will yield less biased results than real openness when productivity in the tradables sector is driven by non-trade factors.

The empirical counterpart of this point is that the AC measure of openness is much more strongly correlated with income levels than the conventional measure of openness. Note that real openness (Ropen) and openness (Open) are linked by the identity log Ropen = log 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 is confirmed by the scatter plot in Figure (4), which shows the difference between these two measures (i.e., log openness – log “real openness”) plotted against income. The correlation is extremely high (p = 0.75). Under the null hypothesis that trade does not cause productivity, this association is spurious. And even under the null that trade does cause productivity, the observed association would be biased upwards unless the only cause of productivity changes is trade (see Appendix A).
Our second 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). 14 On this ground alone, then, the AC strategy introduces an additional distortion to the estimation.

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.

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.

13 Staiger and Stock (1997) recommend a threshold value of 10 for the F-statistic to be confident that the instruments retain their validity. These results are available upon request.

14 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 footnote 13, 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.

C. Dollar and Kraay (DK, 2002)
Recently DK have analyzed the interaction between institutions, trade, and growth.16 They do so by estimating both level regressions and regressions where the dependent variable is the change in the growth rate of income. Their main argument for undertaking the second approach is the alleged multicollinearity between instruments for institutions and trade that militates against a proper disentangling of the two effects. That is, although institutions and trade are jointly significant in level regressions, it is difficult to identify the strength of the individual effects. Since our own results do not suggest multicollinearity to be a problem, we discuss the reasons for the difference.

Level regressions: Equation (1) in Table 7, which reproduces equation 12 in table 1 of DK, is the prime exhibit for their contention that the individual effects of trade and institutions cannot be disentangled. The estimated coefficients on institutions and trade are both insignificant which is allegedly a reflection of multicollinearity. (But notice that the coefficient on openness is negative!) It should be noted that DK follow AC by measuring integration as “real openness.” In the subsequent columns labeled with asterisks, we show how non-robust this finding is, and conversely how robust is the finding relating to the primacy of institutions. Either deleting population from the DK specification (which is insignificant in any case) as in column (1) * or replacing “real openness” with openness (columns (1)**, (1)***, and (1)****) restores the importance of institutions (and with a vengeance as the very high t-values indicate), while openness remains insignificant. Also from Table 6, we know that our preferred specification is unaffected by inclusion of geography variables and by the use of “real openness.”

Moreover, DK’s own results with their larger, 134-country sample are fully consistent with ours: institutions are significant, “real openness” is not. This is shown in equation (2) in Table 7, which reproduces column (6) in DK’s Table 1. DK argue that that the significance of the institutions variable in this larger sample is not robust to the exclusion of the 4 neo-European countries. (See equation (3) in Table 6, which corresponds to DK’s column (7) in Table 1). DK provide no justification for why it is appropriate to exclude the neo-European countries from this sample. We have already established for our preferred sample (Table 3) that: (i) our results are robust to the exclusion of influential observations; (ii) there is no statistical reason to exclude the neo-European countries; and (iii) nevertheless, excluding them leaves our results unchanged. The columns with asterisks in Table 7 confirm this. In the larger sample with neo-Europes (columns (2)*, (2)***, and (2)****) as well as in the sample without neo-Europes (columns (3)* and (3)**) institutions trump integration if the DK equations are estimated without population or if openness is used to replace “real openness”. Thus, the case that multicollinearity blurs the individual effects of trade and institutions is a hard one to make: it requires us to favor a problematic specification with an arbitrarily selected sample over all others, and to disregard much evidence to the contrary.

16 Unlike EL, FR and our paper, DK place less emphasis on geography.
Change regressions: The DK change regressions are difficult to understand. The measures for institutional quality are puzzling and arbitrary, and the identification strategy not clearly justified. In Panel B of Table 7, we reproduce regressions involving each of their institutional variables. In each case, we re-estimate the equation adding time-region dummies. In every instance, the coefficients on real openness cease to be significant. What the results essentially indicate is that the 1980s were a lousy decade for Africa and Latin America and a good decade for Asia; there is no other information in these regressions beyond that. 17

IV. What Does It All Mean?

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 want to highlight another possible difference, having to do with the interpretation of the results. In particular, we want to emphasize 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 (2002).

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 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, 18 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.

Where the AJR paper is successful is in the use of a plausible instrument to identify the causal relationship between institutional quality and income levels. An instrument is something that simply has some desirable statistical properties. It need not be a large part of the causal story. 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

17 More broadly, as Lant Pritchett has pointed out in his comments on the paper, the DK regressions are simply uninformative. That is, running these particular regressions with these instrument sets provides little information that would alter one’s priors one way or the other. The appropriateness of some of the measures of institutional quality used by DK—revolutions and coups and war deaths, for example—is not clear, and it is highly doubtful that these are adequate instruments for measuring institutional change over time. Even leaving aside the insignificance of trade once time-region dummies are included, the instruments for institutions are simply too weak in these decadal regressions for the results to be of much use.

18 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.
quartermillion-theory of earnings. Similarly, the AJR strategy does not amount to a direct
test of a colonial-origins theory of development.  

Easterly and Levine (2002) 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. Thus, both AJR and EL tend to elevate settler mortality beyond its
status as an instrument, with AJR favoring a colonial view of development and EL a
geography-based theory of development.

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

Easterly and Levine (2002) 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 (1999, 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.e., fiscal, trade, monetary, etc.), \( I_i \) institutional
quality, and \( \delta \) the rate at which institutional quality decays absent countervailing action, the
evolution of institutional quality over time can be written as

\[
I_i = \sum a_i \delta + \delta I_i
\]

where \( a_i \) 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 (2002) 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

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

number of decades.) It is that measures of institutional quality already contain all the
relevant information about the impact of policies. If the appropriate specification for income
is \( \ln y = \beta \ln I + \epsilon \), the effect of policies should be sought in a regression of the form

\[
\frac{d \ln y}{dt} = \beta \delta + \ln a_i + \sum \alpha_i \cdot p_i + \epsilon
\]

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 understated 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 C of Table 2 are run over the last three decades, the coefficient on settler mortality,
declines from 0.94 in the 1970s to 0.87 in the 1980s and 0.71 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.

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 \( \beta \), while what policy makers need to know are the \( \alpha_i \)
(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
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 80 countries for (a), (b), and (c); sample of 140 countries for (d), (e), and (f))

- 24 -
Figure 3: Conditional Correlations between Income and its Determinants.
(Sample of 80 countries for (a), (b), and (c); sample of 140 countries for (d), (e), and (f))

Note: The slopes of the linear prediction in (a), (b), and (c) above correspond to the (unstandardized) coefficients in column (6) of Table 2, while those in (d), (e), and (f) correspond to the (unstandardized) coefficients in column (9) of Table 2.

Figure 4: “Real Openness,” Openness, and Income
(Difference between logs of “real openness” and openness on the vertical axis and log per capita PPP GDP on the horizontal axis)
Figure 5: Distribution of incomes for colonized and non-colonized countries

Table 1. Descriptive Statistics

Panel A: Mean and Standard Deviations of Key Variables

<table>
<thead>
<tr>
<th></th>
<th>Extended AJR Sample (80 countries)</th>
<th>Large Sample (140 countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log GDP per capita (PPP) in 1995 (LGC95)</td>
<td>8.01 1.05</td>
<td>8.40 1.14</td>
</tr>
<tr>
<td>Rule of law (RULE)</td>
<td>-0.25 0.85</td>
<td>0.89 0.94</td>
</tr>
<tr>
<td>Log openness (LCOPEN)</td>
<td>3.94 1.05</td>
<td>4.01 0.57</td>
</tr>
<tr>
<td>Distance from equator in degrees (DISTEQ)</td>
<td>15.23 11.16</td>
<td>23.60 16.29</td>
</tr>
<tr>
<td>Log European settler mortality (LOGEM4)</td>
<td>4.66 1.22</td>
<td>.. 1.12</td>
</tr>
<tr>
<td>(deaths per annum per 1000 population)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log constructed openness (LOGFRANKROM)</td>
<td>2.76 0.76</td>
<td>2.92 0.80</td>
</tr>
<tr>
<td>Fraction of population speaking other European language (EURFRAC)</td>
<td>0.30 0.41</td>
<td>0.24 0.39</td>
</tr>
<tr>
<td>Fraction of population speaking English (ENGFRAC)</td>
<td>0.11 0.29</td>
<td>0.08 0.24</td>
</tr>
</tbody>
</table>

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). Appendix B describes in detail all the data and their sources.
Table 2: Determinants of Development: Core Specifications

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>AER sample</th>
<th>Extended AER sample</th>
<th>Large sample</th>
<th>Extended AER sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log GDP</td>
<td>log GDP</td>
<td>log GDP</td>
<td>log GDP</td>
</tr>
<tr>
<td></td>
<td>per capita</td>
<td>per capita</td>
<td>per capita</td>
<td>per capita</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>log GDP</td>
<td>log GDP</td>
<td>log GDP</td>
<td>log GDP</td>
</tr>
<tr>
<td></td>
<td>per capita</td>
<td>per capita</td>
<td>per capita</td>
<td>per capita</td>
</tr>
<tr>
<td>(5)</td>
<td></td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>(9)</td>
<td>(10)</td>
<td>(11)</td>
<td>(12)</td>
</tr>
</tbody>
</table>

Panel A: Ordinary least squares

| Geography (DISTEQ) | 0.74 0.20 0.32 | 0.81 0.25 0.36 | 0.76 0.21 0.24 | 0.92 0.72 |
|                   | (4.48) * | (1.34) | (1.85) | (10.59) * |
| Institutions (RULE) | 0.78 0.69 0.70 | 0.79 0.70 0.77 | 0.80 0.77 0.77 | 0.57 |
|                   | (7.06) * | (6.07) * | (6.06) * | (4.14) * |
| Integration (LCOPEN) | 0.16 0.15 | 0.16 0.15 | 0.16 0.15 | 0.34 |

Panel B: Two-stage least squares

| Geography (DISTEQ) | 0.74 -0.42 -0.96 | 0.81 -0.44 -0.70 | 0.76 -0.05 -0.14 | 0.78 -0.96 |
|                   | (4.40) * | (1.19) | (1.25) | (2.8) * |
| Institutions (RULE) | 1.67 1.78 1.88 | 1.76 2.00 3.34 | 1.99 3.26 3.28 | 0.77 |
|                   | (4.29) * | (1.80) | (2.00) | (1.66) |
| Integration (LCOPEN) | -0.18 -0.32 | -0.18 | 0.23 |
|                   | (1.05) | (1.05) | (2.09) |
| No. of observations | 64 64 64 | 80 80 80 | 140 140 140 | 80 80 80 |
| Resp. | 0.05 0.52 | 0.52 0.52 | 0.47 0.50 0.55 | 0.94 0.38 |

Panel C: First Stage for Endogenous Variables (Institutions (RULE) and Integration (LCOPEN))

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>RULE</th>
<th>RULE</th>
<th>LCOPEN</th>
<th>RULE</th>
<th>RULE</th>
<th>LCOPEN</th>
<th>RULE</th>
<th>RULE</th>
<th>LCOPEN</th>
<th>RULE</th>
<th>RULE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
<tr>
<td></td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
<td>LCOPEN</td>
<td>RULE</td>
<td>RULE</td>
</tr>
</tbody>
</table>

Panel D: Pairwise correlations (Sample of 80 countries)

Notes: p-values reported below the coefficients. Variables described in Appendix B.
Notes: The dependent variable in Panels A and B are 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 80 countries for which data on settler mortality (LOGEM4) have been compiled by AJR; and (iii) columns (7) to (9) use a larger sample of 140 countries for which the instrument for institutions is that in Dollar and Kraay (2002; DK) but is also similar to that in Hall and Jones (1999). The regressions 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 regressions 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 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 Appendix B 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

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Extended AJR sample</th>
<th>Larger sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income per worker</td>
<td>Capital per worker</td>
<td>Human capital</td>
</tr>
<tr>
<td></td>
<td></td>
<td>per worker</td>
</tr>
<tr>
<td>Geography (DISTEQ)</td>
<td>-0.94 (1.47)</td>
<td>-0.25 (1.5)</td>
</tr>
<tr>
<td>Institutions (RULE)</td>
<td>2.22 (3.29) *</td>
<td>0.57 (3.14) *</td>
</tr>
<tr>
<td>Integration (LCOPEN)</td>
<td>-0.41 (-1.31)</td>
<td>-0.15 (-0.79)</td>
</tr>
<tr>
<td>R-square</td>
<td>0.60</td>
<td>0.52</td>
</tr>
<tr>
<td>No. of observations</td>
<td>74</td>
<td>74</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Geography (DISTEQ)</th>
<th>Institutions (RULE)</th>
<th>Integration (LOOPEN)</th>
<th>Legal origin</th>
<th>Identity of colonizer</th>
<th>Religion</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.70 (-1.34)</td>
<td>2.00 (3.55)</td>
<td>-0.302 (-1.07)</td>
<td>0.44</td>
<td>None</td>
<td>[0.133]</td>
</tr>
<tr>
<td>-1.34 (-4.36)</td>
<td>(3.03) *</td>
<td>(1.86) (-1.23)</td>
<td>(1.25)</td>
<td>Singapore</td>
<td></td>
</tr>
<tr>
<td>-0.66 (-1.14)</td>
<td>1.82 (2.43)</td>
<td>-0.31 (-1.25)</td>
<td>(0.33)</td>
<td>Ethiopia</td>
<td></td>
</tr>
<tr>
<td>-0.90 (-1.44)</td>
<td>2.82 (1.67) *</td>
<td>-0.75 (-1.36)</td>
<td>(-0.51)</td>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td>-0.58 (-0.81)</td>
<td>1.97 (0.67) *</td>
<td>-0.42 (-0.81)</td>
<td>(1.11)</td>
<td>New Zealand</td>
<td></td>
</tr>
<tr>
<td>-0.14 (-0.91)</td>
<td>1.32 (0.67) **</td>
<td>-0.17 (-1.36)</td>
<td>(-0.79) *</td>
<td>Canada</td>
<td></td>
</tr>
<tr>
<td>-0.14 (-1.27)</td>
<td>0.90 (0.17)</td>
<td>-0.17 (-1.36)</td>
<td>(6.77) **</td>
<td>New Zealand</td>
<td></td>
</tr>
<tr>
<td>0.02 (-0.12) **</td>
<td>1.69 (0.67) *</td>
<td>0.03 (-0.12)</td>
<td>(6.77) *</td>
<td>Germany</td>
<td></td>
</tr>
<tr>
<td>-0.36 (-0.96)</td>
<td>2.43 (2.04) **</td>
<td>-0.41 (-0.96)</td>
<td>(4.87) *</td>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td>-0.96 (-1.27)</td>
<td>2.22 (0.25)</td>
<td>-0.23 (-0.96)</td>
<td>(2.98) *</td>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td>-0.81 (-1.12)</td>
<td>2.13 (0.46)</td>
<td>-0.32 (-0.96)</td>
<td>(3.09) *</td>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.39) *</td>
<td>Australia</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The dependent variable is per capita GDP in 1995, PPP basis. Baseline 1 corresponds to the specification in column (6) of Table 2. Baseline 2 corresponds to the specification in column (9) of Table 2. 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 DFIITS 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 Table 2.

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

<table>
<thead>
<tr>
<th>Geography (DISTEQ)</th>
<th>Institutions (RULE)</th>
<th>Integration (LOOPEN)</th>
<th>Legal origin</th>
<th>Identity of colonizer</th>
<th>Religion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under frost (FROSTAREA)</td>
<td>-0.65 (0.65)</td>
<td>-0.24 (-1.48)</td>
<td>(-1.17)</td>
<td>0.24 (2.17) **</td>
<td></td>
</tr>
<tr>
<td>Access to sea (ACCESS)</td>
<td>-0.06 (0.06)</td>
<td>0.24 (2.17) **</td>
<td>-1.11</td>
<td>0.24 (2.17) **</td>
<td></td>
</tr>
<tr>
<td>Malaria (MALFAL94)</td>
<td>0.53 (1.29)</td>
<td>0.53 (1.29)</td>
<td>-0.24</td>
<td>0.53 (1.29)</td>
<td></td>
</tr>
<tr>
<td>Temperature (MEANTEMP)</td>
<td>0.65 (0.65)</td>
<td>0.65 (0.65)</td>
<td>-0.24</td>
<td>0.65 (0.65)</td>
<td></td>
</tr>
<tr>
<td>R-square</td>
<td>0.65 (0.65)</td>
<td>0.65 (0.65)</td>
<td>-0.24</td>
<td>0.65 (0.65)</td>
<td></td>
</tr>
<tr>
<td>No. of observations</td>
<td>80 80 77 77 68 72 72 72 68</td>
<td>80 80 77 77 68 72 72 72 68</td>
<td>80 80 77 77 68 72 72 72 68</td>
<td>80 80 77 77 68 72 72 72 68</td>
<td>80 80 77 77 68 72 72 72 68</td>
</tr>
</tbody>
</table>

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 ***. All regressors are scaled as described in the notes to Table 2.
Table 6. Determinants of Development: Robustness to Alternative Measures of and Instruments for Integration

<table>
<thead>
<tr>
<th>Baseline</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geography (DISTEQ)</td>
<td>-0.70</td>
<td>-0.49</td>
<td>-0.85</td>
<td>-1.01</td>
<td>-0.98</td>
</tr>
<tr>
<td>Institutions (RULE)</td>
<td>2.00</td>
<td>1.79</td>
<td>1.88</td>
<td>1.06</td>
<td>1.07</td>
</tr>
<tr>
<td>Integration (LCOPEN)</td>
<td>-0.30</td>
<td>0.20</td>
<td>0.04</td>
<td>-0.71</td>
<td>-0.70</td>
</tr>
<tr>
<td>Land area (AREA)</td>
<td>0.25</td>
<td>0.21</td>
<td>-0.32</td>
<td>-0.32</td>
<td>0.68</td>
</tr>
<tr>
<td>Population (POP)</td>
<td>0.21</td>
<td>0.13</td>
<td>-0.29</td>
<td>-0.28</td>
<td>0.29</td>
</tr>
<tr>
<td>&quot;Real openness&quot; (LNOPEN)</td>
<td>-0.74</td>
<td>-0.89</td>
<td>-0.81</td>
<td>-0.70</td>
<td></td>
</tr>
<tr>
<td>&quot;Policy openness&quot; (SW)</td>
<td>-2.09</td>
<td>-1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The dependent variable is log GDP per capita 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 2. 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. 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 7. Robustness of the Dollar-Kray Results

| Panel A: Two-stage least squares: Dependent variable is log GDP per capita in 1995 |
| (1) | (2) | (3) | (4) | (5) | (6) |
| Institutions (RULE) | 2.14 | 1.73 | 1.52 | 1.54 | 1.64 | 1.26 | 1.54 | 1.39 | 2.02 | 2.83 | 1.56 |
| "Real openness" (LNOPEN) | -1.17 | -0.52 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 |
| Institutions (RULE) | 2.14 | 1.73 | 1.52 | 1.54 | 1.64 | 1.26 | 1.54 | 1.39 | 2.02 | 2.83 | 1.56 |
| "Real openness" (LNOPEN) | -1.17 | -0.52 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 | -0.94 |

Notes: Panel A relates to the level regressions in Table 1 and Panel B to the decadal growth regressions in Table 4 of DK. All variables are as defined in that paper. In Panel A, equations (1), (2), and (3) without asterisks reproduce, respectively, equations (12), (7), and (6) in Table 1. Equations with asterisks represent variations on the equations without asterisks. All regressors, except RULE, are in logarithms and are not scaled as in Tables (2)-(5) to facilitate comparison with the original equations. In Panel B, equations (1) - (5) without asterisks correspond respectively to equations (4), (6), (8), (10), and (12) in Table 4 of DK. Equations with asterisks add time and region dummies to the un-asterisked equation. T-statistics are reported under coefficient estimates. Significance at the 1 percent, 5 percent, and 10 percent levels are denoted respectively by *, ***, and ****.
Appendix A: The Inappropriateness of “Real Openness”

Technology and trade. Imagine a symmetric world populated with a large number of small endowment economies. Each economy 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 $ON_i = 2T_i/(PN_i*N_i + T_i)$.

Preferences. 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.

Cross-national income differences. Now assume 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 $N$, but varying $T_i$. Countries with larger $T_i$ are richer.

Cross-sectional relationship between openness and income. 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.

The AC adjustment. 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.

Remarks. 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.
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 \(\frac{\text{abs(Latitude)}}{90}\). Source: World Bank (2002).

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).

FREEDOM = Political rights index. Freedom House, various issues.


Note: World Bank (2002) refers to the data set used in Dollar and Kraay (2002), which was kindly provided by Aart Kraay.


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


LCOPEN = Natural logarithm of openness. Openness is given by the ratio of (nominal) imports plus exports to GDP (in nominal US dollars). Source: Penn World Tables, Mark 6. Average over all 1950-98 available data.

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

LOPEN = 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).


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.


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

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.


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


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)


REFERENCES


