The Colonial Origins of Comparative Development: An Empirical Investigation

Daron Acemoglu* **Simon Johnson†** **James A. Robinson‡**

March 19, 2000

Abstract

We exploit differences in early colonial experience to estimate the effect of institutions on economic performance. Our argument is that Europeans adopted very different colonization policies in different colonies, with different associated institutions. The choice of colonization strategy was, at least in part, determined by the feasibility of whether Europeans could settle in the colony. In places where Europeans faced high mortality rates, they could not settle and they were more likely to set up worse (extractive) institutions. These early institutions persisted to the present. We document these hypotheses in the data. Exploiting differences in mortality rates faced by soldiers, bishops and sailors in the colonies during the 18th and 19th centuries as an instrument for current institutions, we estimate large effects of institutions on income per capita. Our estimates imply that a change from the worst (Zaire) to the best (US or New Zealand) institutions in our sample would be associated with a five fold increase in income per capita.

*Massachusetts Institute of Technology, Department of Economics, E52-371, Cambridge, MA 02319, and Canadian Institute of Advanced Research; e-mail: daron@mit.edu

†Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA 02319; e-mail: sjohnson@mit.edu.

‡University of California, Department of Political Science, 210 Barrows Hall, Berkeley CA94720; e-mail: jamesar@socrates.berkeley.edu.

1 Introduction

What are the fundamental causes of the large differences in income per capita across countries? Although there is still little consensus on the answer to this question, differences in institutions and property rights have received considerable attention in recent years. The idea is not new. Adam Smith, for example, argued:

> "little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism, but peace, easy taxes, and a tolerable administration of justice; all the rest being brought about by the natural course of things." (Quoted in Jones, 1981, p 235).

More explicitly, countries with better institutions, more secure property rights, and less distortionary policies will invest more in physical and human capital, and will use these factors more efficiently to achieve a greater level of income. This view receives some support from cross-country correlations between measures of property rights and economic development (e.g., Knack and Keefer, 1995, Mauro, 1995, Barro, 1998, Hall and Jones, 1999, Rodrik, 1999), and from a few micro-studies that investigate the relationship between property rights and investment or output (e.g., Besley, 1995, and Mazingo, 1999).

At some level, it is obvious that institutions matter. Witness, for example, the divergent paths of North and South Korea, or East and West Germany, where one part of the country stagnated under central planning and collective ownership, while the other prospered in a market economy. Nevertheless, we still lack convincing evidence that institutional differences can have a large enough effect to explain the phenomenal differences in output per capita across countries. It is quite likely that economies that are rich choose or can afford better institutions. Perhaps more important, economies that are different for a variety of reasons will differ both in their institutions and in their income per capita, biasing the cross-country relationship between institutions and performance.

To estimate the impact of institutions on performance, we need a source of exogenous variation in institutions. In this paper, we propose the differences in mortality rates faced by European settlers at the time of colonization as a possible source of such exogenous

variation.[1] We focus on economies that were colonized by European countries since the 16th century. These provide us with a set of economies that had relatively similar income levels 300 years ago and still exhibit large differences in per capita income. More important for our purposes, institutions in all of these countries were shaped, at least in part, by their colonization experience.

Our argument rests on three premises, which will be substantiated in more detail below:

1. There were different types of colonization policies which created different sets of institutions. At one extreme, European powers set up extractive states, exemplified by the Belgian colonization of the Congo (Zaire). These institutions did not introduce much protection for private property, nor did they provide checks and balances against government expropriation. In fact, the main purpose of the extractive state was to transfer as much of the resources of the colony to the colonizer, with the minimum investment possible.

   At the other extreme, many Europeans went and settled in a number of colonies, creating what the historian Alfred Crosby (1986) calls "Neo-Europes". The settlers tried to replicate European institutions, with great emphasis on private property, and checks against government power. Primary examples of this include Australia, New Zealand, Canada, and the United States.

2. The colonization strategy was influenced by the feasibility of settlements. In particular, in places where the disease environment was not favorable to European settlement, the cards were stacked against the type of colonization policies that led to creation of Neo-Europes. As a result, in these colonies the formation of the extractive state was more likely.

3. The colonial state and institutions persisted even after independence.

Based on these three premises, we use the mortality rates expected by the first settlers in colonies as an instrument for institutions. More specifically, our theory can be schematically summarized as

$$\text{(potential) settler mortality} \implies \text{settlements} \implies \text{early institutions}$$

[1] We do not argue that differences in mortality rates are the only, or even the main, cause of variation in institutions. For our empirical approach to work, all we need is that they are *a source of* exogenous variation.

$\Longrightarrow$ current institutions $\Longrightarrow$ current performance

We put together data on the mortality rates of soldiers, bishops, and sailors stationed in the colonies during the 18th and 19th centuries, largely based on the work of the historian Philip Curtin. These give a good indication of the mortality rates faced by settlers, and in fact, Europeans were well informed about these mortality rates at the time,[2] though they did not know how to control the diseases that caused these high mortality rates. Furthermore, since these mortality rates refer to fairly homogeneous groups, they are comparable across countries. We document that (potential) settler mortality rates were a major determinant of settlements; that settlements were a major determinant of early institutions (in practice, institutions in 1900); that there is a strong correlation between early institutions and institutions today; and finally that current institutions have a first-order effect on current performance.

Our most parsimonious specification is to regress current performance on current institutions, and instrument the latter by settler mortality rates. Since our focus is on property rights and checks against government power, we use the "risk of expropriation" index from Political Risk Services as the proxy for institutions, which measures the differences in institutions originating from different types of states and state policies.[3] The first-stage relationship between this measure of institutions and settler mortality is strong. For example, settler mortality alone explains 25 percent of the variation in this index of institutions. Using this specification, we find that institutions are a major determinant of per capita income. Our estimates imply that a change from the best to the worst institutions in our sample would lead to a decline in per capita income by approximately a factor of 5. The estimates are also quite precise, and in fact somewhat larger than the OLS estimates. They suggest that a significant part of the cross-country income differences we observe are due to variations in institutions.

We document that this relationship is not driven by outliers. For example, excluding Australia, New Zealand, Canada, and the United States does not change the result,

[2]Pilgrim fathers provide an interesting example of this. They first considered migrating to Guyana, but did not because of high expected mortality rates, and chose the U.S. instead (see Crosby, 1986, pp. 143–144).

[3]We do not mean to imply that expropriation by the government is the only institutional feature that matters. Our view is that there is a variety of institutional guarantees, including constraints on government expropriation, independent judiciary, property rights enforcement, equal access to education, and respect for civil liberties, that are important to encourage investment and growth. Expropriation risk is related to all these institutional features. We also check our results using alternative measures of institutions.

nor does excluding Africa. The validity of our approach will also be threatened if other factors correlated with the mortality rates of the settlers affect income per capita. We adopt two strategies to substantiate that our results are not driven by omitted factors. First, we investigate whether institutions have an important effect once we control for a number of variables potentially correlated with mortality and economic outcomes. We find that none of these overturn our results; the estimates change remarkably little when we include controls for climate, temperature, humidity, geography, religion, legal origins, first colonizer, natural resources, and soil quality. Furthermore, the results are also robust to the inclusion of controls for the current disease environment (e.g., the prevalence of malaria), the current fraction of the population with European descent, and measures of ethnolinguistic fragmentation.

Naturally, it is impossible to control for all possible variables that might be correlated with settler mortality and economic outcomes. Furthermore, our empirical approach might capture the effect of settler mortality on economic performance, but working through other factors. For example, Europeans settlers of the 18th and 19th centuries may have brought a "culture" conducive to economic progress, which could still have an effect on income per capita today. Our instrumental variables strategy would then incorrectly assign this effect to institutions. We deal with this problem by using a simple overidentification test. Since our hypothesis is that settler mortality affected settlements; settlements affected early institutions; and early institutions persisted and formed the basis of current day institutions, we can test the validity of our approach by using measures of European migration to the colonies and of early institutions as additional instruments. We then use overidentification tests to detect whether any of these variables has a direct effect, and whether the effect of settler mortality is working through settlements and early institutions, or through other, omitted, variables. The results are encouraging for our approach. In all cases, we fail to reject our hypothesis and we never find evidence for a direct effect of settler mortality on economic outcomes. We conclude that potential mortality rates faced by settlers do not appear to have an effect on economic performance through variables other than institutions, and therefore are a valid instrument for current institutions.

We are not aware of others who have pointed out the link between settler mortality rates and institutions, or have used this observation in conjunction with a theory of institutional persistence to account for comparative development. Scholars such as McNeill

(1976), Crosby (1986) and Diamond (1997) have discussed the important impact of diseases on human history and the historical pattern of imperial conquest and migration. Although Diamond does emphasize comparative development, his theory is based on the geographical determinants of the incidence of the neolithic revolution, not institutions. A large historical literature, including work by Gann and Duignan (1962), Robinson and Gallagher (1961), Denoon (1983), and Cain and Hopkins (1993), emphasizes that settler colonies such as the U.S. and New Zealand are different than other colonies, and point out that these differences were important for their economic success. Nevertheless, they do not develop the link between mortality, settlements and institutions.

Our argument that the profit maximizing strategy for the colonists in places where they did not settle may harm long-term growth is similar to the arguments made by North (1981), and also Robinson (1999), and Acemoglu and Robinson (2000). These works argue that political elites (colonists in our case) will often choose policies that maximize their own income at the expense of the total income of the society. Other related work emphasizing institutional sources of comparative development includes Engerman and Sokoloff (1997) who argue that institutions were shaped primarily by factor endowments and technology, and North, Summerhill and Weingast (1998) who maintain that the reason the U.S. prospered relative to Latin America was because of the good economic and political institutions it inherited from Britain.[4] Our theory agrees with North et al. that different colonial policies were crucial for institutions, and yet differs in focusing on the impact of settler mortality as the key to understanding colonial institutions. In our view, there was not one British or Spanish colonial policy; policies varied depending on settlement patterns.

Empirically, our work is related to a number of other attempts to uncover the link between institutions and development. Studies that look at the correlation between institutions and development, for example Knack and Keefer (1995), are subject to the biases noted above. Two papers try to solve this problem by using an instrumental variables approach as we do here. The first is Mauro (1995) who instruments for corruption using enthnolinguistic fragmentation, and the second is Hall and Jones (1999) who use distance from the equator as an instrument for social infrastructure because they argue that the distance from the equator is correlated with 'Western influence', which leads to

[4]As with Engerman and Sokoloff, Landes (1998, chapters 19 and 20) also discusses factor endowments in his comparison of North and Latin America. However, he puts little stress on institutions and attributes importance to different cultural influences emanating from Britain and Iberia.

good institutions. The theoretical reasoning for these instruments is not entirely convincing. It is not easy to argue that the Belgian influence in the Congo, or Western influence in the Gold Coast during the era of slavery promoted good institutions or governance. Enthnolinguistic fragmentation, on the other hand, seems endogenous, especially since such fragmentation has almost completely disappeared in Europe during the era of growth when a centralized market emerged (e.g., see the classic historical work by Anderson, 1983, on Europe). Econometrically, the problem with both studies is that their instruments can plausibly have a direct effect on performance. For example, Easterly and Levine (1997) argue that enthnolinguistic fragmentation can affect performance by creating political instability (see also Alesina and Perotti, 1997). Sachs and his coauthors (e.g., Bloom and Sachs, 1998, Gallup, Mellinger, and Sachs, 1998) have argued for the direct importance of climate variables on performance, a thesis that goes back at least to Montesquieu [1748](1989), who suggested that low income and despotism are more likely in warmer climates. If, indeed, these variables have a direct effect, they are invalid instruments and do not establish that it is institutions that matter. The advantage of our approach is that conditional on the variables we already control for, settler-mortality more than 100 years ago should have no effect on output today, other than through its effect on institutions.[5]

The paper proceeds as follows. The next section outlines the key hypothesis of this paper, and provides historical evidence in support of this hypothesis. Section 3 presents OLS regressions of GDP per capita on the index of institutions in a variety of samples. Section 4 describes our key instrument for institutions, the mortality rates faced by potential settlers at the time of colonization. Section 5 presents our main results. It demonstrates the correlation between settler mortality and settlements of Europeans in the colonies, the correlation between settlements and early institutions, and the correlation between early and current institutions. It then presents two-stage least squares estimates of the impact of institutions on income per capita using potential settler mortality at the time of colonization as an instrument for institutions. Section 6 investigates the robustness of our results, and Section 7 concludes.

[5]Another closely related paper is La Porta, Florencio de-Silanes, Shleifer and Vishny (1999). They also emphasize the importance of colonial origins. But they focus on the identity of colonial power and legal system, and do not investigate the impact of institutions on economic performance. Finally, Bertocchi and Canova (1996) investigate the effect of being a colony on postwar growth.

2 The Hypothesis and Historical Background

The hypothesis that underlies our empirical work is the following: settler mortality affected settlements; settlements affected early institutions; and early institutions persisted and formed the basis of current day institutions. Or schematically,

$$\begin{aligned} \text{settler mortality} &\Longrightarrow \text{settlements} \Longrightarrow \text{early institutions} \\ &\Longrightarrow \text{current institutions} \Longrightarrow \text{current performance} \end{aligned}$$

In this section, we discuss and substantiate this hypothesis. The next subsection discusses the link between mortality rates of settlers and settlement decisions, then we discuss differences in colonization policies, and finally, we turn to the causes of institutional persistence.

2.1 Mortality and settlements

There is little doubt that mortality rates were a key determinant of European settlements. As mentioned above, the Pilgrim fathers decided to migrate to the U.S. rather than Guyana because of the very high mortality rates in Guyana. Curtin (1964 and 1998) documents how both the British and French press informed the public of the mortality rates in the colonies. For example, early European attempts to settle in West Africa foundered due to high mortality from disease. In the "Province of Freedom" European mortality in the first year was 46%, in Bulama (April 1792-April 1793) there was 61% mortality among Europeans, and in the first year of the Sierra Leone Company (1792-1793) 72% of the European settlers died. On Mungo Park's Second Expedition (May-November 1805), on the overland trip from Gambia to the Niger, 87% of Europeans died, and all the Europeans died before completing this trip. Such rates of mortality were shockingly high for Europeans at the time.[6]

Another interesting example of the awareness of the disease environment comes from the Beauchamp Committee in 1795 which was set up to decide where to send British

[6]The problem was that most mortality in tropics was from (1) malaria (particularly Plasmodium Falciporum), and (2) yellow fever (with devastating periodic epidemics). In the first half of the nineteenth century there was almost a complete misunderstanding of the nature of malaria ("miasma" from swamps was the prevailing view). Quinine was available but not understood nor widely. The role of hygiene was also not properly understood. In the second half of the nineteenth century there developed improved heuristic rules about how to control disease (e.g., an understanding that flight from yellow fever often works, and that mortality from malaria is often less at higher altitudes.) Low mortality for Europeans in areas with tropical diseases only arrived in the early twentieth century. But throughout the whole nineteenth century, areas without malaria were more healthy than Europe (e.g., New Zealand, Mauritius).

convicts, who had previously been sent to the U.S.. One of leading proposals was the island of Lemane, which was 400 miles up Gambia river. However, the committee rejected this possibility precisely because they decided mortality rates would be too high even for the convicts. South-West Africa was also rejected for health reasons; the final decision was to send convicts to Australia.

The eventual expansion of many of the colonies was also related to the living conditions there. In places where the early settlers faced high mortality rates, there would be less incentive for new settlers to come. Curtin (1964), for example, documents how early British expectations for settlement in West Africa were dashed by very high mortality among attempted settlers, about half of whom could be expected to die in the first year.

2.2 Types of colonization and settlements

The historical evidence supports both the notion that there was a wide range of different types of colonizations and that the presence or absence on European settlers was a key determinant of the form colonialism took. Historians, including Gann and Duignan (1962), Robinson and Gallagher (1961), Denoon (1983), and Cain and Hopkins (1993), have documented the development of *"settler colonies"*, where Europeans settled in large numbers, and life were modeled after the home country. Denoon (1983), for example, emphasizes that settler colonies had representative institutions which promoted what the settlers wanted and what they wanted was freedom and the ability to get rich by engaging in trade. He argues that "there was undeniably something capitalist in the structure of these colonies. Private ownership of land and livestock was well established very early..." (p. 35). The classic account by the historians Robinson and Gallagher also agrees with this conclusion. They write "all the most successful trading associations, with the exception of the Indian Empire, were the Europeans transplanted abroad... Here there were stable governments and superior guarantees for life, capital and property." (1981, p. 7). Similarly, the historian Crawford Young (1994, pp. 119-120) notes that in most colonies, the colonial state was completely dominant in society, yet "settlers, where numerous, were a different matter. They carried the perquisites and expectations of citizenship with them and believed that the basic purpose of the colonial state was to assure conditions for their prosperity....Settlers, where they were dominant therefore imposed major limits to the internal autonomy of the colonial state."

In many cases, when the establishment of European-like institutions did not arise

naturally, the settlers were ready to fight for them against the wishes of the home country. Australia is an interesting example here. Most of the early settlers in Australia were ex-convicts, but the land was owned largely by ex-jailors. As a result, initial property rights were unequal, and there was no legal protection against the arbitrary power of landowners. However, the majority of settlers wanted institutions and political rights very similar to those prevailing in England at the time. They demanded jury trials, freedom from arbitrary arrest, and electoral representation. Although the British government resisted at first, the settlers argued that they were British and deserved the same rights as in the home country (see Hughes, 1987). Cain and Hopkins agree with this conclusion and write (1993, p. 237) "from the late 1840's the British bowed to local pressures and, in line with observed constitutional changes taking place in Britain herself, accepted the idea that, in mature colonies, governors should in future form ministries from the majority elements in elected legislatures." They also suggest that "the enormous boom in public investment after 1870 [in New Zealand]... was an attempt to build up an infrastructure... to maintain high living standards in a country where voters expected politicians actively to promote their economic welfare." (p. 225).

The policies that colonial powers adopted in non-settler colonies were very different. European powers adopted many exploitative policies in Latin America during the 17th and 18th centuries, and in Asia and Africa during the 19th and early 20th centuries. The main objective of the Spanish and the Portuguese colonization was to obtain gold and other valuables from America. For example, soon after the conquest the Spanish crown granted rights to land and labor (the *encomienda*) and set up a complex mercantilist system of monopolies and trade regulations to extract resources from the colonies (see Lockhart and Schwartz, 1983, and Lang, 1975).[7] Other European powers were attracted to colonialism because of the success of this strategy (see for example, Young, 1994, p. 64).

European powers developed the slave trade in Africa for the same reasons. Before the mid-nineteenth century, colonial powers were mostly restricted to the African coast and concentrated on monopolizing trade in slaves, gold and other valuable commodities.[8] Thereafter, colonial policy was determined both by an element of superpower rivalry, but

[7]Interestingly, migration to Spanish America was limited by the Spanish Crown, in part because of a desire to keep control of the colonists and limit their independence (for example Coatsworth, 1982), which gives support to our notion that settlers were able to influence the type of institutions set-up in the colonies, even against the wishes of the home country government.

[8]Witness the names used to descibe West African countries; the Gold Coast, the Ivory Coast.

mostly by economic motives. Crowder (1968, p. 50), for example, notes "it is significant that Britain's largest colony on the West Coast [Nigeria] should have been the one where her traders were most active and bears out the contention that, for Britain....flag followed trade." Although in almost all cases the main aim was to cement trading privileges, protect economic interests, and obtain profits, the recipients of these profits varied. In the Portuguese case, it was the state, in the Belgian case, it was King Leopold, and in the British case, it was often private enterprises who obtained concessions in Africa or monopoly trading rights (Crowder, 1968, Part III).

Davis and Huttenback (1986, p. 307) conclude that "the colonial Empire provides strong evidence for the belief that government was attuned to the interests of business and willing to divert resources to ends that the business community would have found profitable." They find that before 1885 investment in the British empire had a return 25% higher than that on domestic investment, though afterwards the two converged. Patrick Manning (1982) estimates that between 1905 and 1914 50% of GDP in Dahomey was extracted by the French, and Young (1994, p. 125) notes that taxation rates in Tunisia were four times as high as those in metropolitan France. Probably the most extreme case of this policy was that of King Leopold of Belgium in the Congo. Gann and Duignan (1979, p. 30) argue that following the example of the Dutch in Indonesia, Leopold's philosophy was that "the colonies should be exploited, not by the operation of a market economy, but by state intervention and compulsory cultivation of cash crops to be sold to and distributed by the state at controlled prices." Peemans (1975) documents the amount of resources extracted from the Belgian Congo and calculates that tax rates on Africans approached 60% of their income during the 1920's and 1930's. Jewsiewicki (1983) writes that during the period when Leopold was directly in charge, policy was "based on the violent exploitation of natural and human resources," with a consequent "destruction of economic and social life...[and].. dismemberment of political structures."

It is important for our hypothesis that contrary to the situation in settler societies, in non-settler colonies, there were few constraints on state power. The colonial powers set-up authoritarian and absolutist states, with the purpose of solidifying their control and facilitating the extraction of resources. Young (1994, p. 101) quotes a French official in Africa as "the European commandant is not posted to observe nature,... He has a mission... to impose regulations, to limit individual liberties...., to collect taxes." Manning (1988, p. 84) summarizes this in the following way: "In Europe the theorists of represen-

tative democracy won out over the theorists of absolutism...But in Africa, the European conquerors set up absolutist governments, based on reasoning similar to that of Louis XIV." These policies included forced labor (both for the construction of public works and on European owned plantations), particularly in French, Belgian and Portuguese colonies (see Cooper, 1996), compulsory cultivation of crops, and heavy rates of taxation. Another interesting example is the marketing boards in Africa (see Fieldhouse, 1994). Marketing boards were first created during the first world war to buy crops from farmers when world trade collapsed. They were re-introduced in the late 1930's both because of the effects of the depression on agricultural prices and because there were 'cocoa strikes' in Ghana against the private companies. During the second world war, these marketing boards were used to extract rents from the farmers in order to help the war effort.

With a strategy of exploitation extraction in mind, European powers had little incentive to invest in institutions or in infrastructure in Africa. Young notes this by writing " [the Belgian companies] brought little capital–a mere 8000 pounds...[to the Congo basin]–and instituted a reign of terror sufficient to provoke an embarrassing public-protest campaign in Britain and the United States at a time when the threshold of toleration for colonial brutality was high." (1994, p. 104). Another famous example, the Indian textile industry was rundown heavily, with no investment (e.g., Fieldhouse, 1999). In fact, despite apparently very high rates of return almost no investment went to Africa (except from South Africa). Instead the British invested heavily in the southern cone of Latin America, Australia, and New Zealand (see Freiden, 1994).

2.3 Institutional persistence

There is a variety of evidence suggesting that the control structures set-up in the non-settler colonies during the colonial era persisted, while there is little doubt that the institutions of law and order and private property established during the early phases of colonialism in Australia, Canada, New Zealand, and the U.S. have been the basis of the current day institutions of these countries.[9] Young emphasizes that the institutions that were set-up by the colonists persisted long after the colonial regime ended. He writes that (1994, p. 283) "although we commonly described the independent polities as 'new

[9]The thesis that institutions persist for a long time goes back at least to Wittfogel (1957), who argued that the control structures set up by the large "hydraulic" empires such as China, Russia, and the Ottoman Empire persisted for more than 500 years to the 20th century. Engerman and Sokoloff (1997), North, Summerhill and Weingast (1998) and Coatsworth (1999) also argue that colonial institutions persisted. Engerman, Mariscal and Sokoloff (1998) provide further evidence in support of this point.

states', in reality they were successors to the colonial regime, inheriting its structures, its quotidian routines and practices, and its more hidden normative theories of governance." Arthur Lewis gives a succinct statement of the issues (1965, p. 32-33)

> "...for most of them [the governments of newly independent African states] independence means merely that they have succeeded to the autocracy vacated by British and French civil servants. They model themselves on the arrogant and arbitrary pattern set by Governors and district commissioners."

There are a number of intuitive economic mechanisms that will lead to institutional persistence of this type. Here, we discuss three possibilities.

1. It is reasonable to suppose that setting up functioning institutions, which place restrictions on government power and respect property rights, is costly. If these costs have been sunk by the colonial powers, then it may not pay the elites at independence to switch from this set of institutions to extractive institutions. In contrast, when the new elites inherit extractive institutions, they may not want to incur the costs of setting up the better institutions, and may prefer to exploit the extractive institutions for their own benefits.

   An interesting case where functioning institutions proved to be very durable is Mauritius. The French settled in relatively large numbers in Mauritius, and the institutions were less extractive than in many other colonies.[10] After independence, Paul Berenger and his party Mouvement Militant Mauricien, which were viewed as Communists at the time, came to power. But in contrast to other African regimes, they continued to support property rights and businesses, and in fact significantly expanded the export processing zones, which then became the foundation of the very rapid growth experience of Mauritius (see, for example, Bowman, 1991).

2. The gains to an extractive strategy may depend on the size of the ruling elite. When this elite is small, each member would have a larger share of the revenues, so the elite may have a greater incentive to be extractive (see Acemoglu and Robinson, 2000). In many cases where European powers set-up authoritarian institutions,

[10]Unfortunately, Mauritius is not in our data set as we do not have information on current institutions. Interestingly, however, Mauritius had very low mortality according to our sources, and is a relatively rich country, with GDP per capita similar to that of Argentina.

they delegated the day-to-day running of the state to a small domestic elite. This narrow group often was the one to control the state after independence and favored extractive institutions. Reno (1995), for example, argues that the governments of post-independence Sierra Leone adopted the tactics and institutions of the British colonizers to cement their political power and extract resources from the rest of society. Boone (1992) provides a similar analysis of the evolution of the modern state in Senegal. Perhaps the most extensively studied case is the Congo. Most scholars view the roots of authoritarianism under Mobutu in the colonial state practices (e.g., Callaghy, 1984). Turner and Young (1985, p. 43) argue that in the period after independence in the Congo, "the ideological model and institutional heritage of the centralized, authoritarian state were at hand in the machinery bequeathed by Bula Matari."[11] Interestingly, Botswana, where ethnic homogeneity led to a large ruling group, has performed well economically, and has not suffered similar institutional problems to other African countries (unfortunately, Bostwana is not in our data set).

The situation in Latin America is similar. Independence of most Latin American countries came in the early nineteenth centuries as domestic elites took advantage of the invasion of Spain by Napoleon to capture control of the state. However, the only thing that changed was the identity of the recipients of the rents. Coatsworth (1978, p. 95) notes "Mexican independence came through a virtual coup d'état by the colony's Creole elite, carried out largely to separate Mexico from the liberalizing process under way in the Mother country... The principal proponents of these conservative efforts was a limited social group of major landowners and industrialists in the center of the country...who had been the principal beneficiaries in the colony of the crown's interventionism or who, like the large merchant houses of the capital, sought to regain privileges the crown itself had abolished in the reforms of the late Bourbon era." The experience of many other Latin American countries is also similar (see, for example, Lynch, 1986).

3. Finally, if agents make irreversible investments that are complementary to a particular set of institutions, they will then be more willing to support them, making

[11] "Bula Matari"; literally "he who breaks rocks." This was the name given to the explorer Henry Morton Stanley by the Congolese to indicate his unrelenting and murderous missions. It is used by Crawford Young as a general term for the colonial state to capture its essence.

these institutions persist. For example, agents who have invested a lot in human and physical capital will be in favor of spending money to enforce property rights, while those who have less to lose may not be.

Further evidence for the persistence of extractive state institutions into the independence era is provided by the persistence of the most prominent extractive policies. In Latin America the full panoply of monopolies and regulations which had been created by Spain remained intact for most of the nineteenth century. Moreover, forced labor policies persisted and were even intensified or re-introduced with the expansion of export agriculture in the latter nineteenth century. Slavery persisted in Brazil until 1886, and during the sisal boom in Mexico, forced labor was reintroduced and persisted up to the start of the revolution in 1910. Forced labor was also reintroducedand in Guatemala and El Salvador to provide labor for coffee growing. In the Guatemalan case, forced labor lasted until the creation of democracy in 1945. Similarly, forced labor was re-instated in many independent African countries (e.g. by Mobutu in Zaire).

The situation is similar with other extractive policies. For example, after independence, colonial marketing boards were a ready made instrument for the newly independent governments to tax agricultural producers (the classic work is Bates, 1981). The elites who controlled the government had little to gain from abolishing these marketing boards and from setting up more efficient institutions. Very often these boards fitted well with the structure of political power in the post independence era. For example, in Ghana, Nkrumah's support was primarily urban and amongst the Fanti, while the cocoa farmers were inland and Ashanti, so he did not find it beneficial to set-up institutions that would restrict his taxation powers, and profit groups outside the ruling coalition. Thus the marketing board provided a perfect instrument of raising taxes from a group that was not going to support him politically in any case (see Rimmer, 1992). Fieldhouse (1999, p. 96) writes "once these boards came under the control of local politicians they were used to extract surplus from the rural producer, notionally for development purposes, in practice largely for party and personal advantage." He also argues (p. 146) that the fact there were no marketing boards was important in the development experience of Australia.

3 Institutions and Performance: Ordinary Least Squares Estimates

3.1 Data and Descriptive Statistics

Table 1 provides descriptive statistics for the key variables of interest. The first column is for the whole sample, and column 2 for our basic sample, limited to the 64 countries that were ex-colonies and for which we have settler mortality data. The GDP per capita in 1995 is PPP adjusted and is taken from the World Bank's 1999 World Development Indicators CD-Rom (a more detailed discussion of all data sources is provided in the appendix Table A1). Income (GDP) per capita will be our measure of economic outcome. Since all the ex-colonies in our sample started with relatively low levels of income 300 years ago, income per capita today is a good measure of long run economic performance. The standard deviation of log per capita income both in the world sample and in our basic sample are similar (1.1 in both cases), indicating that the large income differences are present in our basic sample.

We will use a variety of variables to capture institutional differences. Our main variable, reported in the second row, is a measure of the risk of expropriation faced by individuals in that country. These data are from Political Risk Services, and were first used in the economics literature by Knack and Keefer (1995). Political Risk Services reports a value between 1 and 10 for each country and year, with 1 corresponding to the highest expropriation risk. We use the average value for each country between 1985 and 1995 (values are missing for many countries before 1985). This expropriation measure is appropriate for our purposes since our focus is on differences in institutions originating from different types of states and state policies. We expect our notion of extractive state to correspond to a low value of this index, while the tradition of rule of law and well enforced property rights should correspond to high values. The next row gives an alternative measure, coded from the Polity III data set of Gurr and associates, which is constraints on the executive in 1990 (see Appendix Table A3). This variable is also closely related to state policies and institutions, so is relevant for our purposes. To economize on space, we will focus on the expropriation risk variable, but in the appendix we will report robustness results using the Constraint on the executive variable.

The next three rows give measures of early institutions from the same Gurr data set. The first is a measure of constraints on the executive in 1900 and the second is an index of democracy in 1900. This information is not available for countries that were still colonies in 1900, so we assign these countries the lowest possible score in each index. In the third

row, we report an alternative measure, constraints on the executive in the first year of independence, which does not use this assumption. The final row gives the fraction of the population of European descent in 1900, which is our measure of European settlement in the colonies.

The remaining columns in Table 1 give descriptive statistics for groups of countries at different quartiles of the settler mortality distribution. This is useful since settler mortality is our instrument for institutions (this variable is described in more detail in the next section).

3.2 Ordinary Least Squares Regressions

Table 2 reports Ordinary Least Squares (OLS) regressions of log per capita income on the expropriation risk variable in a variety of samples. The top panel presents linear regressions of the equation

$$\log y_i = \mu + \alpha R_i + X_i'\gamma + \varepsilon_i, \tag{1}$$

where y_i is income per capita in country i, R_i is the expropriation risk measure, X_i is a vector of other covariates (in Table 2, only latitude), and ε_i is a random error term. The coefficient of interest throughout the paper is α, the effect of institutions on income per capita. The first two columns of Table 2 are for the whole world sample, while the second two are for our basic sample. Since a priori we have no reason to believe that the relationship between the institutional index and GDP per capita is linear, in the bottom panel we regress GDP per capita on a set of dummy variables for different quartiles of the institutional index distribution. The omitted group is the set of countries in the lowest quartile (with the worst institutions), and three dummies for the other three quartiles are included in the regression.[12]

Column 1 shows that in the whole world sample there is a strong correlation between our measure of institutions and income per capita. The bottom panel indicates that the linear specification is appropriate, as the dummies are ranked in the expected order, and the gap between the various dummies is not very different from that implied by the linear specification. Columns 3 and 4 show that the impact of the institutions variable on income per capita is quite similar in our basic sample to that in the whole world. For example,

[12]More generally, it may therefore be more appropriate to interpret the OLS and IV estimates from linear specifications as local average treatment effects giving the weighted average of the effects at different points (as suggested in Imbens and Angrist, 1996).

the estimate in column 3, 0.52, indicates that on average, an economy with an institutions index of 10 such as the U.S. or New Zealand, has approximately 3 times as much income per capita as an economy with an index of 3.5 like Zaire. This is a very large effect, and if it were causal, it would imply that differences institutions can account for a very large fraction of the income per capita differences across countries.

Many social scientists, including Montesquieu [1784](1989), Diamond (1997) and Jeffrey Sachs and coauthors have argued for the direct importance of climate variables on performance, and Gallup, Mellinger, and Sachs (1998) and Hall and Jones (1999) document the correlation between distance from the equator (latitude) and economic performance. To control for this, in even numbered columns, we also add the absolute value of the distance from the equator in degrees (latitude) as an additional regressor (we follow the literature in using a measure of latitude that is scaled between 0 and 1). This changes the coefficient of the index of institutions little, and latitude itself is also significant and has the sign found by previous studies. The coefficient 1.80 in this column 4 indicates that a ten degree (0.10) distance gap from the equator (such as the gap between Mauritania and Algeria) is associated with approximately 18 percent difference in income per capita.

Figure 1 shows the relationship between the index of institutions and income per capita diagrammatically in our basic sample. It is apparent from this figure that the relationship is not driven by outliers. The rest of Table 2 further demonstrates this point. Columns 5 and 6 drop the "Neo-Europes", the U.S., Canada, Australia and New Zealand, which are the richest ex-colonies. Columns 7 and 8 drop all African countries. These changes in sample affect the estimates little; the expropriation risk variable is highly significant with a similar coefficient to the basic sample. Columns 9 and 10 limit the sample to the 26 African countries for which we have data. Here the expropriation index is still significant, but with a smaller coefficient than the rest of the table. This suggests that other dimensions of institutions or other factors may be more important in accounting for income differences within Africa.

Overall, the results in Table 2 show a strong correlation between institutions and economic performance. There are a number of important reasons, however, for not interpreting this relationship as causal. First, our measure refers to current institutions, so it is plausible that rich economies are able to afford, or perhaps prefer, better institutions. Arguably more important than this reverse causality problem, there are many omitted determinants of income differences that will naturally be correlated with institutions. Fi-

nally, the measures of institutions are constructed ex post, and the analysts may have had a natural bias in seeing better institutions in richer places. As well as these problems introducing positive bias in the OLS estimates, the fact that the institutions variable is measured with considerable error creates attenuation and may bias the OLS estimates downwards. All of these problems can be solved if we have an instrument for institutions. Such an instrument must be an important factor in accounting for the institutional variation that we observe, but not have a direct effect on performance. Our discussion in Section 2 suggests that the mortality rates faced by the settlers during the time of colonization is a plausible instrument. The next section will describe the construction of this variable in more detail.

4 Data on the Mortality of Early Settlers

Our data on the mortality of European settlers comes largely from the work of Philip Curtin. In a series of books and articles over a period of nearly 40 years, Curtin has used a variety of sources to document the mortality of Europeans in new colonies.

Our main sources, Curtin (1989 and 1998), focus on the mortality and disease rates of European soldiers posted to colonies in the early nineteenth century. Systematic military medical record keeping began only after 1815, as an attempt to understand why so many soldiers were dying in some places. The first detailed studies were retrospective and dealt with British forces between 1817 and 1836. The U.S. and French governments quickly adopted similar methods (Curtin 1989, p. 3 and p. 5), and some early data are also available for the Dutch East Indies. By the 1870s, most European countries published regular reports on the health of their soldiers.

The standard measure adopted was annualized deaths per thousand mean strength. This measure reports the death rate among 1,000 soldiers with each death is replaced with a new soldier. Curtin (1989 and 1998) reviews in detail the construction of these estimates for particular places and campaigns, and assesses which data should be considered reliable.

The data can be divided into two parts. Curtin (1989), *Death by Migration*, deals primarily with the mortality of European troops from 1817 to 1848. At this time modern medicine was still in its infancy and none of the European militaries yet understood how to design their campaigns so as to reduce mortality. This period is well before the control of malaria and yellow fever was understood, and these were the major causes of death from disease for both soldiers and settlers in places where mortality was high (the third

main cause of death was gastrointestinal diseases). These mortality rates can therefore be interpreted as reasonable estimates of settler mortality. They are consistent with substantial evidence from other sources (see, for example, Curtin 1964 and Curtin 1968).

Curtin (1998), *Disease and Empire*, adds similar data on the mortality of soldiers in the second half of the nineteenth century. These numbers have to be used with more care, because there was a growing awareness of how to avoid epidemics of the worst tropical diseases, at least during short military campaigns. For example, the campaigns in Egypt and Ethiopia at the end of the nineteenth century had very low mortality rates (see Table A2 and Figure 1). We did not exclude these two countries because excluding them would have helped our hypothesis. In all case, we use the earliest available number for each country, reasoning that this is the best estimate of the mortality that settlers would have faced, at least until the very end of the nineteenth century.

Appendix 2 reviews in detail how our data are constructed, and describes alternative assumptions that check the robustness of our results. The main gap in the Curtin data is for South America; unfortunately, the Spanish and Portuguese militaries did not keep good records of mortality. However, Gutierrez (1986) used Vatican records to construct estimates for the mortality rates of bishops in Latin America from 1604 to 1876.[13] Because this data overlaps with the Curtin estimates for several countries, we were able to construct a data series for South America (details are in the Appendix). Curtin (1964) also provides estimates of mortality in naval squadrons for different regions which we can use to generate alternative estimates of mortality in South America. These alternative methods produce remarkably similar results.

5 Institutions and Performance: Instrumental Variables Results

5.1 Determinants of Current Institutions

Mathematically, our theory can be expressed in the following way. Equation (1), which we rewrite here,

$$\log y_i = \mu + \alpha R_i + X_i'\gamma + \varepsilon_i, \qquad ((1))$$

describes the relationship between current institutions and GDP. In addition we have

$$R_i = \lambda_R + \beta_R C_i + X_i'\gamma_R + \nu_{Ri} \qquad (2)$$

$$C_i = \lambda_C + \beta_C S_i + X_i'\gamma_C + \nu_{Ci} \qquad (3)$$

[13] We are grateful to Robert McCaa for guiding us to this source.

$$S_i = \lambda_S + \beta_S \log M_i + X_i' \gamma_S + \nu_{Si} \tag{4}$$

where R is the measure of current institutions (expropriation risk between 1985 and 1995), C is our measure of early (circa 1900) institutions, S is the measure of European settlements in the colony (fraction of the population with European decent in 1900), and M is mortality rates faced by settlers. X is a vector of covariates that affect all variables.

The simplest identification strategy is to use $\log M_i$ directly as an instrument for R_i in equation (1). This identification strategy will be valid as long as $\log M_i$ is uncorrelated with ε_i—even if C_i and S_i were correlated with ε_i. For example, if Europeans were more likely to migrate to places with better resources and soil quality, this would create a correlation between S_i and ε_i, so European migration patterns in the 19th-century would not be a good instrument for current institutions. Nevertheless, our identification strategy of using $\log M_i$ directly as an instrument would be valid. We start with this identification strategy, but will then use the other equations to derive overidentifying restrictions.

Figure 2 illustrates the relationship between the (potential) settler mortality rates and the index of institutions. We chose to use the logarithm of the settler mortality rates, since there are no theoretical reasons to prefer the level as a determinant of institutions rather than the log, and as the figure demonstrates, there is a closer association between the log of the mortality rate and our measure of institutions.[14]

In Table 3, we document that this relationship works through the channels hypothesized in Section 2. In particular, we present OLS regressions of equations (2), (3), and (4). In column 1, we regress the expropriation risk on the constraint faced by the executive in 1900. The top panel reports the OLS relationship and shows that there is a close association between early institutions and institutions today. For example, past institutions alone explain 21 percent of variation in the index of current institutions. The second column adds the latitude variable, with little effect on the estimate.

Nevertheless, this regression may be driven by some omitted factors. For example, counties with adverse geographical conditions not captured by our latitude variable, or other unobserved adverse characteristics, may have chosen bad institutions early on, and continue to choose bad institutions today. We deal with this problem in the second panel

[14]Notice two outliers in this figure, Egypt and Ethiopia, which have very low mortality rates relative to their institutions. This is because mortality rates for these two countries come from very short, and well-managed, late 19th-century British campaigns. Excluding these two countries would strengthen our results further. Although the mortality rates from the successful campaigns certainly underestimate the mortality rates faced by the settlers in Egypt and Ethiopia, we did not exclude these two countries in order to stack the cards against us.

by using an instrumental variables approach. We use the (log) mortality of early settlers as an instrument for early institutions. The first stage is reported at the bottom panel, and the second panel reports the two-stage least squares (2SLS) estimate. The relationship documented in columns 1 and 2 continues to hold when we look at the arguably exogenous component of past institutions, driven by settler mortality rates at the time. In fact, if anything, the estimated effect of early institutions is somewhat more pronounced (the point estimate is about 60 percent larger, but so is the standard error). Columns 3 and 4 repeat this exercise using our index of democracy in 1900, and yield similar results.

Both the constraint on the executive and the democracy indices assign low scores to countries that were colonies in 1900. In columns 5 and 6, we adopt an alternative approach and use the constraint on the executive in the first year of independence (and also control separately for time since independence). The results are once again similar, and indicate both with OLS and two-stage least squares that early institutions persist.

Columns 7-10 provide evidence in support of the hypothesis that early institutions were shaped, at least in part, by whether colonists settled in the country or not. They relate our measure of constraint on the executive and democracy in 1900 to the measure of European settlements in 1900 (fraction of the population of European decent). Again, the top panel reports OLS, and the rest of the table reports two-stage least squares estimates using log mortality as an instrument, this time for European settlements. The results indicate that Europeans settled in fewer numbers in counties where they face higher mortality rates, and the colonies where they did not settle ended up with worse institutions.

Overall, Table 3 provides support for our hypothesis that mortality rates faced by potential settlers affected settlement behavior, and that settlement behavior affected the type of institutions that the colonists brought to the country, and finally that early institutions persisted to the present. Next, we use settler mortality directly as an instrument for current institutions, and look at the impact of current institutions on performance.

5.2 Institutions and Economic Performance

The basic results are presented in Table 4. The main equation is (1). We treat the expropriation risk variable, R_i, as endogenous, and model it as

$$R_i = \zeta + \beta \log M_i + X_i'\delta + v_i \tag{5}$$

where M_i is our settler mortality rate. The exclusion restriction is that this variable does not appear in (1), so we estimate equations (1) and (5) jointly with two stage least squares

(2SLS), using $\log M_i$ as an instrument for R_i.

The top panel of Table 4 reports 2SLS estimates of the coefficient of interest, α, and the bottom panel gives the corresponding first-stages.[15] Column 1 shows that in our basic sample, there is a strong first-stage relationship between (log) settler mortality and current institutions. This variable alone explains 25 percent of the differences in institutions we observe today. The corresponding two-stage least squares (2SLS) estimate of the impact of institutions on income per capita is 0.87. This estimate is highly significant with a standard error of 0.15, and is in fact larger than the OLS estimates reported in Table 2. This suggests that measurement error in the institutions variables that creates attenuation bias is likely to be more important than reverse causality and omitted variables biases. The point estimate of 0.87 suggests that an improvement in institutions from 3.5 (Zaire) to 9.5 (Australia or Singapore) will increase income per capita approximately 5 folds, which is a very large effect.

Column 2 shows that adding latitude does not change the relationship; the institutions coefficient is 0.86 (s.e.=0.20). Interestingly, the latitude variable is now highly insignificant. This is suggests that many previous studies may have found latitude to be a significant determinant of economic performance because it is correlated with institutions (or with their exogenous component caused by early colonial experience).

Columns 3 and 4 document that this relationship is not driven by the Neo-Europes. Remarkably, when we exclude the U.S., Canada, Australia, and New Zealand, the estimates remain highly significant, and in fact increase a little. For example, the coefficient for institutions is now 1.1 (s.e.=0.34) without the latitude control, and 1.0 (s.e.=0.31) when we control for latitude.

Columns 5 and 6 show that the result is also robust to dropping all the African countries from our sample. The estimates without Africa are somewhat smaller, but also more precise. For example, the coefficient for institutions is 0.65 (s.e.=0.11) without the latitude control, and 0.63 (s.e.=0.13) when we control for latitude. In contrast, the impact of institutions on performance is much weaker when we limit the sample to African countries only. Without the latitude control, the estimate in column 7 is 1.4, but with a very large standard error. In column 8, when we control for latitude, the standard error becomes enormous. This is because the first-stage relationship between mortality and institutions is nonexistent in the African sample. This result is not too surprising, since

[15]Appendix Table A4 reports similar regressions using constraint on the executive in 1990 as the relevantinstitutions variable, and shows very similar results.

the African sample only contains 26 countries.

Overall, the results in Table 4 show a large effect of institutions on economic performance, though the relationship is considerably weaker within Africa alone. In the rest of the paper, we investigate the robustness of these results.

6 Robustness

6.1 Alternative samples

In Table 5, we investigate the robustness of our results in different subsamples. In Table 4 we reported results in samples which differed by region. Here our objective is to see whether the results are robust in samples with varying degrees of data quality. In columns 1 and 2, we only use data from Curtin (1989), *Death by Migration*, which is for 17 countries and refers always to pre-1840 data. In this smaller sample, we find an even stronger relationship both in the first stage and the second stage estimation. The 2SLS estimate of the effect of institutions on income per capita is 1.0 (s.e.=0.25) without controlling for latitude, and 1.1 (s.e.=0.33) when latitude is included in the regression (compared to 0.87 and 0.86 in Table 4). In columns 3 and 4, we add data from the second book by Curtin (1998), *Disease and Empire,* for the earliest date available for each country. This increases the sample to 31 countries, and reduces both the estimate and the standard error a little; the results are now much more similar to those reported in Table 4. In columns 5 and 6, we use other information (described in detail in the Appendix), to assign mortality rates to neighboring countries with the same disease environment. This increases the sample to 44, and changes the estimates little. In columns 7 and 8, we add data for Latin America, but this time from naval stations instead of bishops. The results are identical to those obtained in Table 4. In columns 9 and 10, we drop Australia from the base sample in columns 9 and 10, with little effect on the results. Finally, in columns 11 and 12, we report regressions using mortality rates relative to home country, which give somewhat larger effects.

Overall we conclude that the results hold in different samples with varying degrees of data reliability.

6.2 Additional controls

The validity of our 2SLS results in Table 4 depends on the assumption that settler mortality in the past has no direct effect on current economic performance. Although this

presumption appears reasonable, we substantiate it further here by directly controlling for many of the variables that could plausibly be correlated both with settler mortality and economic outcomes.

In Table 6, we control for a variety of variables that previous researchers have emphasized, including enthnolinguistic fragmentation (see Mauro, 1995, Easterly and Levine, 1997), and religion, legal origin, and the identity of the colonial power (see LaPorta et al., 1999). In column 1, we add ethnolinguistic fragmentation to our regressions. The top panel gives the 2SLS estimates, the second panel gives the first stage, and the bottom panel gives the corresponding OLS coefficient on the expropriation risk. The even numbered columns add latitude. With or without latitude and in both OLS and IV regressions, the effect of expropriation risk on income per capita is robust to the inclusion of ethnolinguistic fragmentation. For example, in column 1 the coefficient of expropriation risk is 0.64 (s.e.= 0.12), which is only slightly smaller than our baseline estimate. Since ethnolinguistic fragmentation is likely to be endogenous with respect to development (that is, ethnolinguistic fragmentation tends to disappear during periods of growth and formation of centralized markets, see Andersen, 1983) and is correlated with settler mortality, the estimate of 0.64, if anything, understates the effect of institutions on income.[16]

An argument dating back to Max Weber, and recently emphasized by David Landes (1998), views culture as a key determinant of economic performance. If for some reason culture is correlated with settler mortality, the estimates in Table 4 may be biased upwards. To control for this, in columns 3 and 4 we add dummies for main religion in the country. These are Catholic, Muslim, Protestant and other religions. In the table we report the significance level (p-value) of the corresponding F-statistic for these dummies as well as the 2SLS estimate of the effect of institutions. Religion appears highly correlated with income per capita, but the 2SLS estimate of the effect of institutions hardly changes.

LaPorta et al (1999), in turn, argue for the importance of legal origin and colonial origin (main colonizer). In our sample, all countries have either French, British, or German legal origin, so in columns 5 and 6 we add dummies for these origins. The estimate of the expropriation risk changes little, and remains highly significant.[17] In columns 7 and 8 of the table, we add dummies for colonial origin, and finally in column 9 we add all the

[16]We demonstrate in the Appendix that, under plausible conditions, if we add an endogenous variable to our regression, the estimated effect of institutions on performance will be biased downwards.

[17]Interestingly, in contrast to the results of LaPorta et al, the results also indicate that, once we control for the effect of institutions on performance, countries with French legal origins in our sample have higher income per capita.

variables in this table simultaneously. Again, these controls have very little effect on our main findings; both the OLS and IV estimates of the effect of institutions on income are very similar to our baseline estimates.

Another concern is that settler mortality is correlated with climate variables and current disease environment. Our instrument may therefore be picking up the direct effect of these variables. In Table 7, we add a set of temperature and humidity variables (from Philip Parker, 1997). These are: average, minimum and maximum monthly high, and minimum and maximum monthly low temperatures, and morning minimum and maximum, and afternoon minimum and maximum humidity. In the table we report joint significance levels for these variables. Again, they have little effect on our estimates.

Sachs and a series of coauthors have argued for the importance of malaria in explaining African poverty (see, for example, Bloom and Sachs, 1998, and Gallup et al., 1998). Since malaria was one of the main, but not the only, cause of settler mortality, our estimate may be capturing some of the direct effects of malaria. We are skeptical of this argument since the prevalence of malaria is highly endogenous; it is the poorer countries with worse institutions that have been unable to eradicate malaria.[18] Moreover, it is hard to imagine how malaria could have such a large effect on economic performance: although Sachs and coauthors argue that this works through poor health, high mortality and absenteeism, most people who live in high malaria areas have developed some immunity to the disease; if they survive to the age of five, they can fall ill periodically but malaria is not fatal to them (see Curtin, 1998, Bruce-Chwatt, 1980). We might therefore expect that malaria should not have a very large effect on economic performance (though obviously it will have very high social costs). In contrast, for Europeans or anyone else who has not been exposed to malaria as a very young child, malaria was likely to be fatal.[19] So variations in malaria were a key determinant of whether Europeans could settle in a colony.

In any case, controlling for malaria does not eliminate our result. We do this in the second part of Table 7 by controlling for the fraction of the population who live in an area

[18]For example, the U.S. eliminated malaria from the Panama Canal zone, and Australia eliminated it from Queensland (see Crosby, 1986, p141-142). Even in Africa, there are very successful campaign against malaria, including those in Algeria and that conducted by the Rio-Tinto Zinc mining company in Zambia (then Northern Rhodesia).

[19]Some types of malaria are quite local. It is therefore quite possible for a person to have immunity to the local version of malaria, but to be highly vulnerable to malaria a short distance away. This is probably one explanation why Africans had such high mortality when they were forced to move by colonial powers. For example, African labor on the Congo-Ocean railroad had mortality of 240 per thousand at its peak, and 100 per thousand on average.

where falciporum malaria is endemic in 1994 (as constructed and used by Gallup et al, 1998). Since, as argued above, this malaria measure is highly endogenous, controlling for it directly will bias the effect of institutions on performance downwards (see the Appendix). As a result, the coefficient on expropriation risk is now estimated to be smaller, about two-thirds of our baseline estimate. Nevertheless, it remains highly significant, so we conclude that the effect of settler mortality on performance is not through its correlation with the current disease environment, but likely through its effect on institutions.

A related concern is that in colonies where Europeans settled, the current population consists of a higher fraction of Europeans. One might be worried that we are capturing the direct effect of having more Europeans. To control for this we add the fraction of the population of European descent in Table 7. Interestingly, in columns 7 and 8 this variable is insignificant, while the effect of institutions remains highly significant (though the parameter estimate is about 25% lower, which is not surprising given the correlation between our instrument and fraction of European descent today). In column 9 of Table 6, we add all these variables, for temperature, humidity, malaria, and European descent, together, and the estimate of expropriation risk on income per capita remains statistically significant.

Finally, in Table 8 we add a variety of controls for soil quality, natural resources, and whether the country is landlocked (see the Table A1 in the Appendix for definitions). The estimates change relatively little when these controls are added. We therefore conclude that the effect of variations institutions caused by early colonial experience on economic performance is robust, and likely captures the causal effect of institutions and government policies on economic well-being.

6.3 Overidentification tests

We can also investigate the validity of our approach by using overidentification tests. Recall that according to our theory settler mortality affected settlements; settlements affected early institutions; and early institutions affect current institutions. Mathematically, we expressed these links by equations (2), (3), and (4). We can test whether any of these variables, C, S, and M, has a direct effect on income per capita, $\log y$, by using measures of C and S as additional instruments.

The overidentification test will reject if any of the following is true:

1. The equation of interest (1) does not have a constant coefficient, i.e., $\log y_i = \mu +$

$\alpha_i R_i + \varepsilon_i$.

2. Either ν_{Ri} or ν_{Ci} is correlated with ε_i.

3. C or S has a direct effect on income per capita, $\log y$.

4. Settler mortality, M, has an effect on $\log y$ that works through another variable, such as culture.

We will see that the overidentifying restrictions are never rejected. This implies that, subject to the caveats related to the power of the overidentification test, we can rule out all four of the above possibilities. This gives us additional confidence that settler mortality is a valid instrument and that we are estimating the effect of institutions on current performance with our instrumental variable strategy (i.e., not capturing the effect of some other variable correlated with settler mortality).

The results of the overidentification tests are reported in Table 9. The top panel reports the 2SLS estimates of the effect of expropriation risk on GDP per capita using a variety of instruments. The second panel reports the corresponding first stages. The bottom panel gives the p-value for the overidentification test that the estimated 2SLS coefficient is different from that using only settler mortality (from Table 4). For example, in column 1, we add European settlement in 1900 as an additional instrument. This results in an estimated effect of 0.9, as compared to our baseline estimate of 0.87. The p-value of .55 indicates that we fail to reject the equality of these coefficients comfortably. The other columns add latitude, and other instruments, including our measures of early institutions used in Table 3 (constraint on the executive and democracy in 1900 and in the first year of independence). In all cases, we fail to reject the overidentifying restriction. We also repeated the same exercise, treating European settlement and early institutions variables as the base regressors, and added log mortality as an additional regressor. Once again, in all cases, we passed the overidentifying restrictions comfortably. There is, therefore, little evidence suggesting that either any of these variables have a direct effect or that the effect of settler mortality is working through other, omitted, variables.

7 Concluding Remarks

Many economists and social scientists believe that differences in institutions and state policies are at the root of large differences in income per capita across countries. There is

little agreement, however, about what determines institutions and the attitude of governments towards economic progress. This makes it difficult to isolate exogenous sources of variation in institutions to estimate their effect on performance. In this paper we argued that differences in colonial experience could be a viable source of exogenous differences in institutions.

Our argument rests on the following premises. First, Europeans adopted very different colonization strategies, with different associated institutions. In one extreme, as in the case of the U.S., Australia and New Zealand, they went and settled in the colonies and set up institutions that enforced the rule of law and encouraged investment. In the other extreme, as in the Congo or the Gold Coast, they set up extractive states with the intention of transferring resources rapidly to the metropole. The slave trade, perhaps, is the most extreme example of this behavior. These sets of institutions were very detrimental to investment and economic progress. Second, which colonization strategy was adopted was, at least in part, determined by the feasibility of settlement. In places where Europeans faced very high mortality rates, so they could naturally not go and settle, they were more likely to set up the extractive state. Finally, we argue that these early institutions persisted to the present. Therefore, determinants of whether Europeans could go and settle in the colonies have an important effect on institutions today. We exploit these differences as a source of exogenous variation to estimate the effect of institutions on economic performance.

We document these hypotheses in the data by showing a high correlation between mortality rates faced by soldiers, bishops, and sailors in the colonies and European settlements; between European settlements and early measures of institutions; and between early institutions and institutions today. We estimate large effects of institutions on income per capita using this source of variation. Our estimates imply that a change from the worst (Congo) to the best (the U.S. or New Zealand) institutions in our sample would be associated by a five fold increase in income per capita. We also document that this relationship is not driven by outliers, and is robust to controlling for climate, current disease environment, religion, natural resources, and current race composition. These results indicate a strong effect of institutions on performance.

It is useful to point out that our findings do not imply that institutions today are predetermined by colonial policies, and cannot be changed. We emphasize colonial experience as one of the many factors affecting institutions. Since mortality rates faced by

settlers are arguably exogenous, they are useful as an instrument to isolate the effect of institutions on performance. In fact, our reading is that these results suggest substantial income gains from improving institutions in poor countries.

8 Appendix A:Data Sources

Appendix Table A1 shows the complete list of variables that we use, together with their definitions and sources.

9 Appendix B:Construction of Settler Mortality Data

Appendix Table A2 reports our estimates for settler mortality in each country. The first column shows data from Curtin (1989) for countries located in the exact place for which we have a mortality estimate from the first half of the nineteenth century. These data are used in the first two columns of Table 5.

We have direct estimates of settler mortality for Algeria (1831-38), Canada (1817-36), Guyana (1819-36), Jamaica (1817-36), Malta (1817-36), New Zealand (1846-55), Senegal (1819-38), and Sierra Leone (1819-36). We also have comparable estimates for Britain (1830-36) and France (1820-22 and 1824-26).

In some cases there has been a simple change of name, but the country is the same geographically. We can therefore use the estimate of settler mortality from Ceylon (1817-36) for Sri Lanka, from Coastal Burma (1829-38) for Myanmar, from Cape Colony (1818-36) for South Africa, and from the Windwards and Leewards islands (1817-36) for Barbados and for Trinidad and Tobago. It is also reasonable to use the mortality estimate from Bengal (1830-38) for Bangladesh, from the Dutch East Indies (1819-28) for Indonesia, and from the Straits Settlements (1829-38) for Malaysia and Singapore.

For large countries where we have more than one regional estimate, we need to make a choice. For the U.S. we use the estimate for the Northern United States (1829-38). For India we use the estimate from Madras (1829-38); this is lower than for Bengal but higher than for Bombay. None of our regression results are significantly affected by using alternative regional data for these countries.

The second column in Appendix Table A2 adds data from Curtin (1998), which covers the second half of the nineteenth century. Whenever Curtin provides more than one estimate, we use the earliest available number for each country. We have mortality data directly from Mexico in 1862-63 and Tunisia in 1881. As we note in the text, data from the second half of the nineteenth century have to be used with care, because there was some improvement in the practice of military medicine around the mid-nineteenth century, but most of this would not help reduce the mortality of civilian settlers. For example, in the British campaigns of 1882 in Egypt and 1867-68 in Ethiopia had extremely

low mortality because the army hurried in and out carrying their own water. We use these data but note that they will tend to bias the results against our hypothesis. On the other hand, the very high mortality from disease estimate for Madagascar in 1895 is a reasonable estimate of what settlers would have expected because this was not a well-organized military campaign (Curtin 1998).

The second column of Appendix Table A2 only assigns a settler mortality number to a modern country if the country name today is the same as in Curtin (1989 and 1998), or if the colony included several modern countries, or if the named nineteenth century country encompasses another country and the disease environment is definitely the same (i.e., Senegal and Gambia.) For example, Burkina Fasu, Central African Federation, Chad, French Congo, Gabon, Mali, Mauritania, Niger were part of French Soudan, the estimate for Hong Kong is from the 1860 China Field Force, Sudan is from Nilotic Sudan, and Vietnam is from Cochin China. These data are used in the third and fourth columns of Table 5.

The third column of Appendix Table A2 assumes that some neighbouring countries have the same settler mortality as countries for which we have direct data, but only if it is reasonable to assume these neighbours have same disease environment. We use this for a number of countries within West Africa (in particular, Benin, Cote d'Ivoire, Ghana, Guinea, Nigeria, Togo are from Sierra Leone), and also for some neighbours of French Soudan (i.e., Angola, Burundi, Cameroon, Rwanda, Kenya, Uganda, Tanzania, and Zaire).[20] The estimate for Morocco is from Algeria and Pakistan is from Bombay. Surinam is a neighbor of French Guiana, for which we have a direct estimate for the period 1819-36. These data are used in the fifth and sixth columns.

The fourth column of Appendix Table A2 provides the data that we use in most of our analysis. We use the data from the third column and add settler mortality estimates for a few countries that are close by and have the same disease environment as places for which we have direct data. We use the New Zealand estimate for Australia, the New Caledonia and Tahiti estimates for Fiji, and the Windwards and Leewards estimate for the Bahamas, Jamaica for Haiti and the Dominican Republic (information from Gutierrez 1986 indicates that these were similar disease environments). Given that Australia is a

[20]French Soudan was obviously a very large area and about half our data on Africa comes from this one estimate. However, the other fragmentary evidence indicates that mortality was high throughout this region, although not as high as in West Africa. Table 4 shows that dropping Africa does not affect our results.

potentially important country for our estimates, the last two columns of Table 5 report results if we drop this data completely.

We also use a combination of the Gutierrez (1986) and Curtin (1989) numbers to construct estimates for South America. Gutierrez calculates the mortality of bishops aged 40-49 for three disease environments in the Caribbean, Central and South America: 10 (low), 11 (medium), and 23 (high) per 1000. We assume the ratio between bishop mortality levels in different disease environments was the same as the ratio between soldier mortality. We also assume that the type of disease environment (low, medium or high) was the same for bishops and soldiers. We then use Gutierrez's ratios with the Curtin (1998) estimate of mortality of 71 per 1000 for Mexico 1862-63 (a low mortality region in Gutierrez's classification), to generate estimates of mortality for Central and South America.

Note that the estimates we obtain would be virtually the same if we used the Curtin (1989) estimate for Jamaica as the benchmark (a high disease environment by Gutierrez's classification.) The implied mortality in Jamaica, using the Gutierrez data, would be 163.3. The direct estimate for mortality in Jamaica is 130 per 1000 in 1817-36 (Curtin 1989). Running our regressions using data that takes Jamaica as the base for applying the Gutierrez's ratios actually strengthens our results.

Gutierrez provides considerable detail on the location of low, medium, and high mortality areas for bishops. High mortality regions are the Caribbean, low lying parts of Central America, and tropical South America. This includes Nicaragua, and Panama. Medium mortality locations are Costa Rica, El Salvador, Honduras, Paraguay, and Venezuela. Low mortality locations are Bolivia, Brazil, Colombia, Ecuador, Guatemala, Peru, and Uruguay.[21]

Note that some of these countries have a large variation in their disease environment, even for a relatively small country. We know, for example, that the Colombian coast was very unhealthy to Europeans. Mortality in the British siege of Cartegena in 1742 was very high— the force was there for only two months, but deaths from disease were between 2/3 and 3/4 of the army (Curtin 1989, p. 2). But most Europeans who went to Colombia were able to settle at a higher altitude, e.g., in Bogota, where the disease environment was

[21]It is particularly difficult to assign an estimate to Brazil because we do not know the precise difference in mortality between the North-East, where much of the early settlement was attempted, and more southern locations that later attracted population. We have adopted a low mortality estimate as this is less favorable to our hypothesis.

much more favorable. It therefore seems reasonable to follow Gutierrez's classification of disease environments, as the bishops tended to live where European population was concentrated.

The last column of Appendix Table A2 uses an alternative source to calculate mortality in Latin America. Average annual mortality, 1825-1845, from disease in the British naval forces off West Africa was 54.4 per hundred (Curtin 1964, p. 486). Comparable mortality from disease at other naval stations over the same period was 7.7 in South America, 9.3 in the Mediteranean, 9.8 in the Home Station, 15.1 in the East Indies, and 18.1 in the West Indies. The level of mortality from disease for British sailors was lower than for British soldiers in all locations, but it is reasonable to assume that the ratio of mortalities between the various regions was approximately the same.

This change of data does not change our estimates much (see the seventh and eighth columns of Table 5). For example, Argentina and Chile move from a mortality of 71 per 1000 (using the data on bishops) to 68.9 (using the data on naval forces). Table 5 reports regression results using alternative classifications, with little effect on our findings.

Even using these alternative sources of data, there remain several significant gaps. Most important, we do not have any reliable estimates for much of Southern Africa, including Lesotho, Namibia, Swaziland, Zimbabwe, Zambia, Malawi, Mozambique, or Botswana. We also do not have any reasonable estimates for Eritrea, the Philippines, Cape Verde or Comoros. Liberia and Thailand were not colonies and are not included in our study.

Appendix Table A2 also reports mortality data for several countries for which we do not have data on institutions or GDP per capita (PPP basis): Mauritania, Myanmar and Mauritius. Qualitatively, of course, the GDP per capita in these countries is quite consistent with the mortality numbers and our regression results.

These estimates of mortality for soldiers are similar to the more fragmentary evidence available on civilian settlers. For example, we know that early European attempts to settle in West Africa foundered due to high mortality from disease (Curtin 1964). In the "Province of Freedom" European mortality in the first year was 46%, in Bulama (April 1792-April 1793) there was 61% mortality among Europeans, and in the first year of the Sierra Leone Company (1792-1793) 72% of the European settlers died. On Mungo Park's Second Expedition (May-November 1805), on the overland trip from Gambia to the Niger, 87% of Europeans died, and all the Europeans died before completing this trip.

10 Appendix C:Bias in The Effect of Institutions When Other Endogenous Variables Are Included

To simplify notation, suppose that R_i is exogenous, and another variable that is endogenous, z_i, such as prevalence of malaria or ethnolinguistic fragmentation, is added to the regression. Then, the simultaneous equations model becomes

$$\begin{aligned} Y_i &= \mu_0 + \alpha R_i + \pi z_i + \varepsilon_i \\ z_i &= \mu_1 + \phi Y_i + \eta_i, \end{aligned}$$

where $Y_i = \log y_i$. We presume that $\alpha \geq 0$, $\phi < 0$, and $\pi < 0$, which implies that we interpret z_i as a negative influence on income. Moreover, this naturally implies that $cov(\eta_i, \varepsilon_i) < 0$ and $cov(z_i, R_i) < 0$, that is the factor z_i is likely to be negatively correlated with positive influences on income.

Standard arguments imply that

$$\text{plim}\widehat{\alpha} = \alpha + \frac{cov\left(\widetilde{R}_i, \varepsilon_i\right)}{var(\widetilde{R}_i)} = \alpha - \kappa \cdot \frac{cov(z_i, \varepsilon_i)}{var(\widetilde{R}_i)},$$

where κ and are $\widetilde{R}_i$ are the coefficient and the residual from the auxiliary equation,

$$R_i = \kappa_0 + \kappa z_i + \widetilde{R}_i,$$

and so

$$\kappa = \frac{cov(z_i, R_i)}{var(z_i)} < 0,$$

which is negative due to the fact that $cov(R_i, z_i) < 0$. The reduced form for z_i is:

$$z_i = \frac{1}{1 - \phi\pi}\left((\mu + \phi\pi) + \phi\alpha R_i + \phi\varepsilon_i + \eta_i\right). \tag{7}$$

We impose the regularity condition $\phi \cdot \pi < 1$, so that an increase in the disturbance to the z-equation, η_i, actually increases z_i. Now using this reduced form, we can write

$$\text{plim } \widehat{\alpha} = \alpha - \kappa \cdot \frac{cov(z_i, \varepsilon_i)}{var(\widetilde{R}_i)} = \alpha - \kappa \cdot \frac{(\sigma_{\varepsilon\eta} + \phi\sigma_\varepsilon^2)}{(1 - \phi\pi) \cdot var(\widetilde{R}_i)} \tag{8}$$

where σ_ε^2 is the variance of ε, and $\sigma_{\varepsilon\eta}$ is the covariance of ε and η.

Substituting for κ in (8), we obtain:

$$\text{plim } \widehat{\alpha} = \alpha - \frac{(\sigma_{\varepsilon\eta} + \phi\sigma_\varepsilon^2)}{(1 - \phi\pi) \cdot var(\widetilde{R}_i)} \cdot \frac{cov(z_i, R_i)}{var(z_i)}$$

Recall that $\phi < 0$, $\sigma_{\varepsilon\eta} < 0$, and $cov(z_i, R_i) < 0$. Therefore, plim $\widehat{\alpha} < \alpha$, and when we control for the endogenous variable z_i, the coefficient on our institution variable will be biased downwards.

REFERENCES

Acemoglu, Daron and James A. Robinson (2000) "Political Origins of Poverty and Prosperity," Unpublished Working Paper, MIT.

Alesina, Alberto and Roberto Perotti (1997) "Income Distribution, Political Instability and Investment," *European Economic Review*, 40, 1203-1225.

Anderson, Benedict (1983) *Imagined Communities*, Verso, London

Bates, Robert H. (1981) *Markets and States in Tropical Africa*, University of California Press, Berkeley CA.

Bayart, Jean-Francois. (1993) *The State in Africa: The Politics of the Belly*, Longman, New York NY.

Bertocchi, Graziella and Fabio Canova (1996) "Did Colonization Matter for Growth? An Empirical Exploration into the Historical Causes of Africa's Underdevelopment," CEPR Discussion Paper #1444.

Besley, Timothy (1995) "Property Rights and Investment Incentives: Theory and Evidence from Ghana," *Journal of Political Economy*, 103, 903-937.

Bloom, David E. and Jeffrey D. Sachs (1998) "Geography, Demography, and Economic Growth in Africa," *Brookings Papers on Economic Activity*, 1998:2, 207-295.

Boone, Catherine (1992) *Merchant Capital and the Roots of State Power in Senegal, 1930-1985*, Cambridge University Press, New York NY.

Bowman, Larry W (1991) *Mauritius: democracy and development in the Indian Ocean*,

Boulder, Colo., Westview Press.

Bruce-Chwatt, Leonard J. (1980) *Essential Malariology*, London.

Cain, Philip J. and Anthony G. Hopkins (1993) *British Imperialism: Innovation and Expansion 1688-1914*, Longman, New York.

Callaghy, Thomas M. (1984) *The State-Society Struggle: Zaire in Comparative Perspective*, Columbia University Press, New York NY.

Coatsworth, John H. (1978) "Obstacles to Economic Growth in Nineteenth-Century Mexico," *American Historical Review*, 83, 80-100.

Coatsworth, John H. (1982) "The Limits of Colonial Absolutism: Mexico in the Eighteenth Century," in Karen Spalding ed. *Essays in the Political, Economic and Social History of Latin America*, University of Delaware Press, Newark DL.

Coatsworth, John H. (1999) "Economic and Institutional Trajectories in Nineteenth Century Latin America," in John H. Coatsworth and Alan M. Taylor eds. *Latin America and the World Economy since 1800*, Harvard University Press, Cambridge MA.

Cooper, Frederick (1996) *Decolonization and African Society: The Labor Question in French and British Africa*, Cambridge University Press, New York NY.

Crosby, Alfred (1986) *Ecological Imperialism: The Biological Expansion of Europe 900-1900*, Cambridge University Press, New York NY.

Crowder, Michael (1968) *West Africa under Colonial Rule*, Northwestern University Press, Evanston IL.

Curtin, Philip D. (1964) *The Image of Africa,* University of Wisconsin Press, Madison.

Curtin, Philip D. (1968) **"**Epidemiology and the Slave Trade," *Political Science Quarterly,* 83:181-216.

Curtin, Philip D. (1989) *Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century*, Cambridge University Press, New York NY.

Curtin, Philip D. (1998) *Disease and Empire: The Health of European Troops in the Conquest of Africa*, Cambridge University Press, New York NY.

Davis, Lance E. and Robert A. Huttenback (1987) *Mammon and the Pursuit of Empire: The Political Economy of British Imperialism, 1860-1912*, Cambridge University Press, Cambridge UK.

Denoon, Donald (1983) *Settler Capitalism: The Dynamics of Dependent Development in the Southern Hemisphere*, Clarendon Press, Oxford UK.

Diamond, Jared M. (1997) *Guns Germs and Steel: The Fate of Human Societies*, W.W. Norton & Co., New York NY.

Dunn, Richard S. (1972) *Sugar and Slaves: The Rise of the Planter Class in the English West Indies, 1624-1713*, University of North Carolina Press, Chapel Hill NC.

Easterly, William and Ross Levine (1997) "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 1203-1250.

Engerman, Stanley L. and Kenneth L. Sokoloff (1997) "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies," in S.H. Haber ed. *How Latin America Fell Behind*, Stanford University Press, Stanford CA.

Engerman, Stanley L., Elisa Mariscal and Kenneth L. Sokoloff (1998) "Schooling, Suffrage, and the Persistence of Inequality in the Americas, 1800-1945," Unpublished

Paper, Department of Economics, UCLA.

Fieldhouse, David K. (1994) *Merchant Capital and Economic Decolonization: The United Africa Company 1929-1987*, Oxford University Press, Oxford UK.

Fieldhouse, David K. (1999) *The West and the Third World*, Blackwell, Oxford UK.

Freiden, Jeffery (1994) "International Investment and Colonial Control: A New Interpretation," *International Organization*, 48, 559-593.

Gallup, John L., Andrew D. Mellinger and Jeffrey D. Sachs (1998) "Geography and Economic Development," NBER Working Paper #6849.

Gallup, John L. and Jeffrey D. Sachs (1998) "The Economic Burden of Malaria," Unpublished paper, Harvard Center for International Development, October.

Gann, Lewis. H. and Peter Duignan (1962) *White Settlers in Tropical Africa*, Penguin, Baltimore MD.

Gann, Lewis. H. and Peter Duignan (1979) *The Rulers of Belgian Africa*, Princeton University Press, Princeton NJ.

Gutierrez, Hector (1986) "La Mortalite des Eveques Latino-Americains aux XVIIe et XVII Siecles," *Annales de Demographie Historique*, 29-39.

Gurr, Ted Robert (1997) "Polity II: Political Structures and Regime Change, 1800-1986," Unpublished paper, University of Colorado, Boulder.

Hall, Robert E. and Charles I. Jones (1999) "Why Do Some Countries Produce so Much More Output per Worker than Others?," *Quarterly Journal of Economics*, CXIV, 83-116.

Hughes, Robert (1987) *The Fatal Shore*, Collins Harvill, London UK.

Imbens, Guido and Joshua Angrist (1994) "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467-475.

Jewsiewicki, B. (1983) "Rural Society and the Belgian Colonial Economy," in D. Birmingham and P.M. Martin eds. *The History of Central Africa, Volume II*, Longman, New York.

McEvedy, Colin and Richard Jones (1975), *Atlas of World Population History*, Facts on File, New York.

Knack, Steven and Philip Keefer (1995) "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Economics and Politics,*

Landes, David S. (1998) *The Wealth and Poverty of Nations: Why Some Are So*

Rich and Some So Poor, W.W. Norton & Co., New York.

Lang, James (1975) *Conquest and Commerce: Spain and England in the Americas*, Academic Press, New York NY.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny (1999) "The Quality of Government," *Journal of Law, Economics and Organization*, 15, 222-279.

Lewis, W. Arthur (1965) *Politics in West Africa*, Oxford University Press, New York NY.

Lockhart, James and Stuart B. Schwartz (1983) *Early Latin America*, Cambridge University Press, New York NY.

Lynch, John (1986) *The Spanish American Revolutions, 1808-1826*, W.W. Norton & Co. New York NY.

Manning, Patrick (1982) *Slavery, Colonialism, and Economic Growth in Dahomey, 1640-1980*, Cambridge University Press, New York NY.

Manning, Patrick (1988) *Francophone Sub-Saharan Africa, 1880-1995*, Cambridge University Press, New York NY.

Mauro, Paulo (1995) "Corruption and Growth," *Quarterly Journal of Economics*, 681-712.

Mazingo, Christopher (1999) "Effects of Property Rights on Economic Activity: Lessons from the Stolypin Land Reform," Unpublished Paper, Department of Economics, MIT.

McNeill, William H. (1976) *Plagues and Peoples*, Anchor Press, Garden City NJ.

Montesquieu, Charles de Secondat [1748] (1989) *The Spirit of the Laws*, Cambridge University Press, New York NY.

North, Douglas C. (1981) *Structure and Change in Economic History*, W.W. Norton & Co., New York.

North, Douglass C., William Summerhill and Barry Weingast (1998) "Order, Disorder and Economic Change: Latin America vs North America," Unpublished Manuscript, Hoover Institution, Stanford University.

Parker, Philip M. (1997) *National Cultures of the World: A Statistical Reference,* Cross-Cultural Statistical Encyclopedia of the World, Volume 4, Greenwood Press, Westport, CT.

Peemans, Jean-Philippe (1975) "Capital Accumulation in the Congo under Colo-

nialism: The Role of the State," in Lewis. H. Gann and Peter Duignan eds. *Colonialism in Africa 1870-1960, Volume 4 The Economics of Colonialism*, Hoover Institution Press, Stanford CA.

Reno, William (1995) *Corruption and State Politics in Sierra Leone*, Cambridge University Press, New York NY.

Rimmer, Douglas (1992) *Staying Poor: Ghana's Political Economy 1950-1990*, Pergamon Press, Oxford UK.

Robinson, James A. (1999) *A Political Theory of Underdevelopment*, Unpublished Book Manuscript, Department of Political Science, University of California, Berkeley.

Robinson, Ronald E. and John Gallagher (1961) *Africa and the Victorians: The Official Mind of Imperialism*, MacMillan, London.

Rodrik, Dani (1999) "Where has all the Growth Gone?" forthcoming in the *Journal of Economic Growth*.

Sheridan, Richard B. (1974) *Sugar and Slavery: An Economic History of the British West Indies, 1623-1775*, University of the West Indies Press, Bridgetown Barbados.

Stein, Barbara, H. and Stanley J. Stein (1970) *The Colonial Heritage of Latin America*, Oxford University Press, New York NY.

Turner, Thomas and Crawford Young (1985) *The Rise and Decline of the Zairian State*, University of Wisconsin Press, Madison WI.

Wittfogel, Karl A. (1957) *Oriental Despotism: A Comparative Study of Total Power*, Yale University Press, New Haven CT.

Young, Crawford (1994) *The African Colonial State in Comparative Perspective*, Yale University Press, New Haven CT.

Zolberg, Aristide R. (1966) *Creating Political Order; The Party-States of West Africa*, Rand McNally, Chicago IL.

Table 1
Descriptive Statistics

	Whole World	Base Sample	By Quartiles of Mortality			
			(1)	(2)	(3)	(4)
Log GDP per capita (PPP) in 1995	8.3	8.05	8.73	8.43	8	7.12
	(1.1)	(1.1)				
Average Expropriation Risk 1985-95	7	6.5	7.6	6.5	6.5	5.6
	(1.8)	(1.5)				
Constraint on Executive in 1970	2.8	3	4.4	2.7	3.1	2
	(2.1)	(2.2)				
Constraint on Executive in 1900	1.9	2.3	3.7	3.2	1.2	1
	(1.8)	(2.1)				
Democracy in 1900	1.1	1.6	3.9	2.5	0.25	0
	(2.6)	(3.0)				
Constraint on Executive in First Year of Independence	3.6	3.3	4.6	2.5	3.1	3.2
	(2.4)	(2.4)				
European Settlements in 1900	0.32	0.19	0.31	0.34	0.12	0.004
	(0.4)	(0.3)				
Log Mortality	4.5	4.5	2.9	4.3	4.8	5.8
	(1.1)	(1.1)				
N	163	64	16	17	13	18

Standard deviations are in parentheses

Quartiles of mortality are for our a base sample of 64 observations. These are
1: less than 65.4
2: greater than or equal to 65.4 and less than 78.1
3: greater than or equal to 78.1 and less than 280
4: greater than or equal to 280

Log GDP per capita for 1995 in PPP terms from World Development Indicators, 1999
Constraint on executive in 1900 and democracy in 1900 is set equal to 1 for countries that were colonies in that year
Number of observations in "Whole World"
Log GDP (163), Expropriation Risk (128), Log Mortality (88), Constraint on Executive in 1900 (89), Democracy in 1900 (89)
Constraint on Executive in First Year of Independence (89), European settlements in 1900 (155)
Number of observations in "Base Sample," where different from 64
Constraint on Executive in 1900 (59), Democracy in 1900 (59), Constraint on Executive in First Year of Independence (59)
European settlements in 1900 (62)
See Appendix Table A1 for more detailed variable definitions and sources.

Table 2
OLS Regressions

	Whole World (1)	Whole World (2)	Base Sample (3)	Base Sample (4)	Base Sample without neo-Europes (5)	Base Sample without neo-Europes (6)	Base Sample without Africa (7)	Base Sample without Africa (8)	Africa only (9)	Africa only (10)
				Linear Specification						
Average Expropriation Risk 1985-1995	0.53 (0.04)	0.46 (0.06)	0.52 (0.06)	0.46 (0.07)	0.48 (0.08)	0.46 (0.08)	0.49 (0.07)	0.45 (0.08)	0.30 (0.11)	0.25 (0.10)
Latitude		1.00 (0.49)		1.80 (0.70)		2.00 (0.86)		0.97 (0.77)		2.60 (1.20)
Adjusted R-Squared	0.60	0.62	0.52	0.55	0.38	0.43	0.56	0.56	0.22	0.33
				Specification with Dummy Variables						
dummy for expropriation index in second quartile	0.24 (0.21)	0.21 (0.21)	0.52 (0.25)	0.45 (0.23)	0.52 (0.24)	0.46 (0.24)	0.31 (0.34)	0.33 (0.33)	0.63 (0.29)	0.45 (0.28)
dummy for expropriation index in third quartile	1.46 (0.21)	1.30 (0.21)	1.50 (0.26)	1.40 (0.25)	1.50 (0.25)	1.40 (0.25)	1.20 (0.32)	1.14 (0.31)	1.20 (0.44)	1.10 (0.41)
dummy for expropriation index in fourth quartile	2.20 (0.21)	1.70 (0.27)	1.96 (0.35)	1.60 (0.35)	0.98 (0.47)	1.10 (0.47)	1.80 (0.39)	1.60 (0.40)	0.34 (0.71)	0.25 (0.65)
Latitude		1.50 (0.50)		2.40 (0.74)		1.80 (0.92)		1.5 (0.85)		2.70 (1.20)
Adjusted R-Squared	0.58	0.61	0.43	0.54	0.36	0.39	0.43	0.46	0.3	0.33
N	111	111	64	64	60	60	38	38	26	26

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices (from World Bank's World Development Indicators 1999)
Standard errors are in parentheses
The top panel reports OLS regressions, and the bottom panel reports regressions on dummies for the expropriation index, with the dummy for the first quartile --i.e., with highest risk of expropriation-- as the omitted category.

Table 3
Determinants of Current Institutions

	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample	Base Sample
	Dependent Variable is Expropriation Risk in 1995						Dependent variable is Constraint on Executive in 1900		Dependent Variable is Democracy in 1900	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Ordinary Least Squares									
Constraint on Executive in 1900	0.33	0.28								
	(0.08)	(0.09)								
Democracy in 1900			0.24	0.21						
			(0.06)	(0.07)						
Constraint on Executive in First Year of Independence					0.26	0.24				
					(0.08)	(0.07)				
European Settlements in 1900							4.80	4.40	7.20	6.20
							(0.71)	(0.89)	(1.00)	(1.20)
Latitude		1.80		1.40		2.50		1.60		3.40
		(1.40)		(1.50)		(1.40)		(1.90)		(2.60)
Adjusted R-Squared	0.21	0.21	0.22	0.22	0.19	0.22	0.42	0.42	0.47	0.48
	Second-Stage Instrumenting using log mortality									
Constraint on Executive in 1900	0.55	0.56								
	(0.14)	(0.18)								
Democracy in 1900			0.40	0.43						
			(0.10)	(0.14)						
Constraint on Executive in First Year of Independence					0.66	0.59				
					(0.22)	(0.22)				
European Settlements in 1900							9.30	14.10	12.90	18.10
							(1.80)	(5.10)	(2.40)	(6.40)
Latitude		-0.18		-1.1		1.5		-10.9		-11.8
		(1.90)		(2.10)		(1.80)		(7.00)		(8.90)
				First-Stage						
Log Mortality	-1.2	-1.0	-1.7	-1.3	-0.95	-0.96	-0.13	-0.07	-0.13	-0.07
	(0.20)	(0.20)	(0.30)	(0.30)	(0.27)	(0.29)	(0.03)	(0.03)	(0.03)	(0.03)

Standard errors are in parentheses. The top panel reports OLS estimates, the middle panel 2SLS estimates, and the bottom panel the corresponding first stages. There are 62 observations in each regression.
The coefficient for latitude in the first-stage regression in even numbered columns is not reported. Regressions with constraint on executive in first year of independence also include years since independence as a regressor.
See Appendix Table A1 for more detailed variable definitions and sources.

Table 4
IV Regressions

	Base Sample	Base Sample	Base Sample without neo-Europes	Base Sample without neo-Europes	Base Sample without Africa	Base Sample without Africa	Africa only	Africa only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expropriation Risk 1985-1995	0.87	0.86	1.1	1.0	0.65	0.63	1.4	3.5
	(0.15)	(0.20)	(0.34)	(0.31)	(0.11)	(0.13)	(1.90)	(58.9)
Latitude		0.1		1.4		0.3		-5.6
		(1.20)		(1.20)		(0.92)		(150.6)
				First-Stage for Expropriation Risk				
Log Mortality	-0.66	-0.56	-0.42	-0.44	-1.2	-1.1	-0.14	-0.01
	(0.14)	(0.16)	(0.15)	(0.16)	(0.21)	(0.23)	(0.20)	(0.25)
Latitude		1.70		-0.50		1.00		2.50
		(1.40)		(1.50)		(1.40)		(2.80)
Adjusted R-Squared	0.25	0.26	0.11	0.09	0.44	0.44	-0.02	-0.03
N	64	64	60	60	38	38	26	26

Dependent Variable is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS estimates, and the bottom panel gives the corresponding first stage.
Standard errors are in parentheses

Table 5
IV Regressions with Alternative Samples

	Pre-1855 data	Pre-1855 data	Earliest available data	Earliest available data	Including neighbours with same disease environment	Including neighbours with same disease environment	Including Latin America (from naval stations)	Including Latin America (from naval stations)	dropping Australia from Base Sample	dropping Australia from Base Sample	Using Ratio of Colonial to European Mortality	Using Ratio of Colonial to European Mortality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Expropriation Risk 1985-1995	1.0	1.1	0.8	0.75	0.85	0.86	0.87	0.86	0.9	0.9	0.93	0.96
	(0.25)	(0.33)	(0.18)	(0.22)	(0.16)	(0.24)	(0.15)	(0.20)	(0.17)	(0.23)	(0.15)	(0.25)
Latitude		-0.57		0.59		-0.11		0.1		0.03		-0.32
		(1.90)		(1.30)		(1.40)		(1.20)		(1.30)		(1.4)
					First-Stage for Expropriation Risk							
Log Mortality	-0.88	-0.77	-0.68	-0.59	-0.62	-0.5	-0.66	-0.56	-0.62	-0.52	-0.63	-0.52
	(0.27)	(0.30)	(0.21)	(0.23)	(0.16)	(0.18)	(0.14)	(0.16)	(0.15)	(0.16)	(0.15)	(0.17)
Latitude		1.60		1.80		2.10		1.70		1.70		1.9
		(1.80)		(1.90)		(1.70)		(1.40)		(1.40)		(1.4)
Adjusted R-Squared	0.38	0.37	0.24	0.24	0.25	0.26	0.25	0.26	0.22	0.23	0.22	0.23
N	17	17	31	31	44	44	64	64	63	63	64	64

Dependent Variable is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS estimates, and the bottom panel gives the corresponding first stage.
Standard errors are in parentheses
The last two columns use the log ratio of mortality in the colony to that the home country (using Britain for all British colonies and France for all other colonies) as the instrument.

Table 6
IV Regressions with Additional Controls

	Base Sample (1)	Base Sample (2)	Base Sample (3)	Base Sample (4)	Base Sample (5)	Base Sample (6)	Base Sample (7)	Base Sample (8)	Base Sample (9)
Expropriation Risk 1985-1995	0.64	0.66	0.89	0.92	1.00	1.10	1.00	1.1	0.97
	(0.12)	(0.15)	(0.14)	(0.22)	(0.18)	(0.28)	(0.19)	(0.28)	(0.23)
Latitude		-0.24		-0.29		-0.58		-0.54	-0.83
		(0.91)		(1.30)		1.50		(1.40)	(1.40)
Ethnolinguistic Fragmentation	-1.2	-1.2							-0.44
	(0.30)	(0.30)							(0.47)
P-value for Religion Variables			[0.001]	[0.002]					[0.52]
French legal origin dummy					0.91	0.96			0.23
					(0.31)	(0.38)			(0.75)
P-value for Colonial Dummies							[0.01]	[0.04]	[0.46]
				First-Stage for Expropriation Risk					
Log Mortality	-0.68	-0.59	-0.63	-0.5	-0.58	-0.47	-0.65	-0.56	-0.68
	(0.160)	(0.170)	(0.140)	(0.160)	(0.140)	(0.160)	(0.170)	(0.200)	(0.220)
Latitude		1.9		2.2		2.0		1.2	1.6
		(1.4)		(1.5)		(1.3)		(1.5)	(1.6)
Adjusted R-Squared	0.24	0.25	0.28	0.3	0.3	0.31	0.29	0.27	0.3
				Ordinary Least Squares					
Expropriation Risk 1985-1995	0.52	0.46	0.53	0.46	0.57	0.51	0.55	0.5	0.46
	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)

Standard errors are in parentheses
Dependent Variable is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS estimates, the middle panel gives the corresponding first stage, and the bottom panel gives the corresponding OLS.
p-values are for joint significance of the indicated variables.
Religion dummies are catholics, muslims, and other religions; protestant is the base case.
Our sample is all either French or British legal origin (as defined by LLSV 1999, JLEO)
Colony dummies are French, German, Spanish, Italian, Belgian, Dutch and Portuguese; British is the base case.
All regressions have 64 observations. See Table A1 for more detailed variable definitions.

Table 7
Robustness Checks for IV Regressions

	Base Sample (1)	Base Sample (2)	Base Sample (3)	Base Sample (4)	Base Sample (5)	Base Sample (6)	Base Sample (7)	Base Sample (8)	Base Sample (9)
Expropriation Risk in 1995	0.85	0.84	0.85	0.8	0.55	0.58	0.75	0.78	0.5
	(0.20)	(0.22)	(0.14)	(0.17)	(0.20)	(0.26)	(0.19)	(0.22)	(0.16)
Latitude		0.23		0.68		-0.4		-0.45	-0.9
		(1.50)		(1.10)		(0.92)		(1.1)	(1.04)
P-value for Temperature Variables	[0.85]	[0.85]							[.94]
P-value for Humidity Variables			[0.41]	[0.33]					[0.27]
Malaria in 1994					-0.83	-0.83			-0.7
					(0.38)	(0.40)			(0.25)
Percent of European descent in 1975							0.61	0.66	0.66
							(0.5)	(0.5)	-0.41
				First-Stage for Expropriation Risk					
Log Mortality	-0.59	-0.53	-0.8	-0.69	-0.44	-0.37	-0.52	-0.49	-0.67
	(0.16)	(0.17)	(0.15)	(0.16)	(0.20)	(0.20)	(0.15)	(0.16)	(0.21)
Latitude		2.3		2		1.6		0.67	2.3
		(1.8)		(1.3)		(1.5)		(1.5)	(2.0)
Adjusted R-Squared	0.26	0.26	0.34	0.35	0.27	0.27	0.29	0.28	0.36
				Ordinary Least Squares					
Expropriation Risk 1985-1995	0.44	0.41	0.49	0.41	0.37	0.37	0.40	0.40	0.30
	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)

Standard errors are in parentheses. Dependent Variable is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS estimates, the middle panel gives the corresponding first stage, and the bottom panel gives the corresponding OLS.
Temperature variables are average temperature, minimum monthly high, maximum monthly high, minimum monthly low, and maximum monthly low, all in centigrade.

Table 8
Further Robustness Checks for IV Regression

	Base Sample (1)	Base Sample (2)	Base Sample (3)	Base Sample (4)	Base Sample (5)	Base Sample (6)	Base Sample (7)
Expropriation Risk in 1995	0.78	0.77	1	0.98	0.91	0.91	0.83
	(0.15)	(0.19)	(0.25)	(0.30)	(0.17)	(0.25)	(0.31)
Latitude		0.1		0.33		-0.07	-0.56
		(1.1)		(1.3)		(1.3)	(1.3)
P-Value for Soil Quality	[0.32]	[0.41]					[0.87]
P-Value for Natural Resources			[0.61]	[0.62]			[0.63]
Dummy for being landlocked					0.5	0.51	0.34
					(0.48)	(0.55)	(0.52)
			First-Stage for Expropriation Risk				
Log Mortality	-0.67	-0.59	-0.5	-0.45	-0.6	-0.49	-0.41
	(0.15)	(0.17)	(0.15)	(0.17)	(0.14)	(0.16)	(0.19)
Latitude		1.6		0.97		1.9	0.86
		(1.4)		(1.4)		(1.4)	(1.6)
Adjusted R-Squared	0.27	0.27	0.24	0.23	0.29	0.3	0.25
				Ordinary Least Squares			
Expropriation Risk 1985-1995	0.50	0.46	0.50	0.46	0.51	0.44	0.39
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)

Standard errors are in parentheses. Dependent Variable is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS estimates, the middle panel gives the corresponding first stage, and the bottom panel gives the corresponding OLS.
Measures of natural resources are: percent of world gold reserves today, percent of world iron reserves today, percent of world zinc reserves today, number of minerals present in country, and oil resources (thousands of barrels per capita.)
Measures of soil quality/climate are steppe (low latitude), desert (low latitude), steppe (middle latitude)

Table 9
Overidentification Tests

	Base Sample (1)	Base Sample (2)	Base Sample (3)	Base Sample (4)	Base Sample (5)	Base Sample (6)	Base Sample (7)	Base Sample (8)	Base Sample (9)	Base Sample (10)
Expropriation Risk 1985-1995	0.9	0.91	0.75	0.73	0.75	0.72	0.59	0.6	0.58	0.59
	(0.13)	(0.16)	(0.13)	(0.17)	(0.13)	(0.16)	(0.11)	(0.13)	(0.11)	(0.13)
Latitude		-0.23		0.33		0.34		-0.17		0.12
		(1.10)		(1.00)		(1.00)		(0.83)		(0.81)
				First-Stage for Expropriation Risk						
Log Mortality	-0.41	-0.42	-0.46	-0.42	-0.44	-0.41	-0.5	-0.43	-0.43	-0.34
	(0.15)	(0.16)	(0.19)	(0.20)	(0.19)	(0.20)	(0.18)	(0.19)	(0.19)	(0.20)
Latitude		-0.4		1.2		0.9		1.7		1.80
		(1.5)		(1.5)		(1.5)		(1.5)		(1.50)
European settlements in 1900	2.10	2.2								
	(0.60)	(0.7)								
Constraint on executive in 1900			0.17	0.15						
			(0.1)	(0.1)						
Democracy in 1900					0.14	0.12				
					(0.07)	(0.08)				
Constraint on executive in first year of independence							0.17	0.17		
							(0.08)	(0.08)		
Democracy in first year of independence									0.13	0.13
									(0.06)	(0.06)
Adjusted R-Squared	0.37	0.36	0.27	0.26	0.28	0.27	0.27	0.27	0.28	0.29
					Results from Overidentification Test					
p-value (from chi-squared test)	0.55	0.54	0.68	0.67	0.74	0.68	0.92	0.89	0.8	0.77

Dependent Variable in Second-Stage Estimates is log GDP per capita (PPP basis) in 1995. The top panel reports 2SLS coefficients of expropriation risk.
The second panel gives the corresponding first stage. The bottom panel gives the p-value for that null hypothesis is that the coefficient on expropriation risk in the second-stage regression is the same as that without the additional instruments (as in Table 4).
Standard errors are in parentheses
All regressions with constraint on executive and democracy in first year of independence also include years since indepedence as a regressor.
The two covariance matricies are based on a common estimate of the disturbance variance (from the more fully efficient estimator)
All regressions have 59 observations, except those with European population in 1900, which have 62 observations

Appendix Table A1

Variable	Description	Source
Log GDP per capita (PPP) in 1995	Logarithm of GDP per capita, on Purchasing Power Parity Basis. In our base sample this ranges from 6.1 to 10.2.	World Bank, World Development Indicators, CD-Rom, 1999
Average expropriation Risk, 1985-95	Risk of expropriation of private investment by government, from 1 to 10, where a higher score means less risk. We calculated the mean value for the scores in all years from 1985 to 1995. In our base sample this ranges from 3.5 to 10.	Dataset obtained directly from Political Risk Services, September 1999. This data was previously used by Knack and Keefer (1995) and was organized in electronic form by the IRIS Center (University of Maryland). The original compilers of this data are Political Risk Services.
Constraint on Executive in 1970	A seven category scale, from 1 to 7, with a higher score indicating more constraints. Score of 1 indicates unlimited authority; score of 3 indicates slight to moderate limitations; score of 5 indicates substantial limitations; score of 7 indicates executive parity or subordination. Scores of 2, 4, and 6 indicate intermediate values.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Constraint on Executive in 1900	From 1 to 7. Coding as for Constraint on Executive in 1970.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Constraint on Executive in First year of independence	From 1 to 7. Coding as for Constraint on Executive in 1970. Date of independence is first year that country appears in Polity III dataset.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Democracy in 1900	An eleven category scale, from 0 to 10, with a higher score indicating more democracy. Points are awarded on three dimensions: Competitiveness of Political Participation (from 1 to 3); Competitiveness of Executive Recruitment (from 1 to 2, with a bonus of 1 point if there is an election); and Constraints on Chief Executive (from 1 to 4).	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Democracy in First Year of Independence	From 0 to 10. Coding as for Democracy in 1900. Date of independence is first year that country appears in Polity III dataset	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
European settlements in 1900	Percent of population that was European or of European descent in 1900. Ranges from 0 to 0.99 in our base sample.	McEvedy and Jones (1978).
Percent of European descent 1975	Percent of population that was European or of European descent in 1975. Ranges from 0 to 1 in our base sample.	McEvedy and Jones (1978).
Ethnolinguistic Fragmentation	Average of five different indices of ethnolinguistic fragmentation.	Easterly and Levine (1997), as used in La Porta et al (1999).
Religion Variables	Percentage of the population that belonged to the three most widely spread religions of the world in 1980 (or for 1990-95 for countries formed more recently). The four classifications are: Roman Catholic, Protestant, Muslim, and "other".	La Porta et al (1999)
French legal origin dummy	Legal origin of the company law or commercial code of each country. All the countries in our base sample are either of French Commercial Code Origin or English Common Origin.	La Porta et al (1999)
Colonial Dummies	Dummy variable indicating whether country was a British, French, German, Spanish, Italian, Belgian, Dutch or Portuguese colony.	La Porta et al (1999)
Temperature Variables	Temperature variables are average temperature, minimum monthly high, maximum monthly high, minimum monthly low, and maximum monthly low, all in centigrade.	Parker (1997)
Humidity Variables	Humidity variables are morning minimum, morning maximum, afternoon minimum, and afternoon maximum, all in percent.	Parker (1997)
Soil Quality	Measures of soil quality/climate are steppe (low latitude), desert (low latitude), steppe (middle latitude), desert (middle latitude), dry steppe wasteland, desert dry winter, and highland.	Parker (1997)
Natural Resources	Measures of natural resources are: percent of world gold reserves today, percent of world iron reserves today, percent of world zinc reserves today, number of minerals present in country, and oil resources (thousands of barrels per capita.)	Parker (1997)
Dummy for Landlocked	Dummy variable equal to 1 if country does not adjoin the sea.	Parker (1997)
Malaria in 1994	Malaria in 1994 is percent of people living in area where falciparum malaria is endemic	Gallup and Sachs 1998.
Latitude	Absolute value of the latitude of the country, scaled to take values between 0 and 1.	La Porta et al (1999)
Log Mortality	Log of estimated settler mortality. From 1.7 to 6.2.	See Appendix B for details on how this variable is constructed; Curtin (1989 and 1998); Gutierrez (1986)

Appendix Table A2
Data on Mortality

Former Colonies

	Abbreviated name used in graphs	Log GDP per capita (PPP) in 1995	Average expropriation 1985-95	First mortality estimate	Second mortality estimate	Third mortality estimate	Fourth mortality estimate (used in main analysis)	Fifth mortality estimate
Angola	AGO	7.77	5.36			280	280	280
Argentina	ARG	9.13	6.39				71	68.9
Australia	AUS	9.90	9.32				8.55	8.55
Burkina Faso	BFA	6.85	4.45		280	280	280	280
Bangladesh	BGD	6.88	5.14	71.41	71.41	71.41	71.41	71.41
Bahamas	BHS	9.29	7.50				85	85
Bolivia	BOL	7.93	5.64				71	71
Brazil	BRA	8.73	7.91				71	71
Barbados	BRB	9.27	.	85	85	85	85	85
Central African Fed.	CAF	7.19	.		280	280	280	280
Canada	CAN	9.99	9.73	16.1	16.1	16.1	16.1	16.1
Chile	CHL	9.34	7.82				71	68.9
Cote d'Ivoire	CIV	7.44	7.00			483	483	483
Cameroon	CMR	7.50	6.45			280	280	280
Congo (French)	COG	7.42	4.68		280	280	280	280
Colombia	COL	8.81	7.32				71	71
Costa Rica	CRI	8.79	7.05				78.1	78.1
Dominican Re	DOM	8.36	6.18			130	130	130
Algeria	DZA	8.39	6.50	78.2	78.2	78.2	78.2	78.2
Ecuador	ECU	8.47	6.55				71	71
Egypt	EGY	7.95	6.77		5.7	5.7	5.7	5.7
Ethiopia	ETH	6.11	5.73		12.1	12.1	12.1	12.1
Gabon	GAB	8.91	7.82		280	280	280	280
Ghana	GHA	7.37	6.27			483	483	483
Guinea	GIN	7.49	6.55			483	483	483
Gambia	GMB	7.27	8.27		164.66	164.66	164.66	164.66
Guatemala	GTM	8.29	5.14				71	71
Guyana	GUY	7.90	5.89	32.18	32.18	32.18	32.18	32.18
Hong Kong	HKG	10.05	8.14			14.9	14.9	14.9
Honduras	HND	7.69	5.32				78.1	78.1
Haiti	HTI	7.15	3.73				130	130
India	DNI	7.33	8.27	48.63	48.63	48.63	48.63	48.63
Indonesia	IND	7.33	7.59	170	170	170	170	170
Jamaica	JAM	8.19	7.09	130	130	130	130	130
Kenya	KEN	7.06	6.05			280	280	280
Sri Lanka	LKA	7.73	6.05	69.8	69.8	69.8	69.8	69.8
Morocco	MAR	8.04	7.09			78.2	78.2	78.2
Madagascar	MDG	6.84	4.45		302	302	302	302
Mexico	MEX	8.94	7.50		71	71	71	71
Mali	MLI	6.57	4.00		280	280	280	280
Malta	MLT	9.43	7.23	16.3	16.3	16.3	16.3	16.3
Myanmar	MMR	.	5.77	34.6	34.6	34.6	34.6	34.6
Mauritania	MRT	7.41	.		280	280	280	280
Mauritius	MUS	9.05	.	30.5	30.5	30.5	30.5	30.5
Malaysia	MYS	8.89	7.95	17.7	17.7	17.7	17.7	17.7
Niger	NER	6.73	5.00		280	280	280	280
Nigeria	NGA	6.81	5.55			483	483	483

	Abbreviated name used in graphs	Log GDP per capita (PPP) in 1995	Average expropriation 1985-95	First mortality estimate	Second mortality estimate	Third mortality estimate	Fourth mortality estimate (used in main analysis)	Fifth mortality estimate
Nicaragua	NIC	7.54	5.23				163.3	163.3
New Zealand	NZL	9.76	9.73	8.55	8.55	8.55	8.55	8.55
Pakistan	PAK	7.35	6.05			36.99	36.99	36.99
Panama	PAN	8.84	5.91				163.3	163.3
Peru	PER	8.40	5.77				71	71
Paraguay	PRY	8.21	6.95				78.1	78.1
Rwanda	RWA	6.48	.			280	280	280
Sudan	SDN	7.31	4.00		36.4	36.4	36.4	36.4
Senegal	SEN	7.40	6.00	164.66	164.66	164.66	164.66	164.66
Singapore	SGP	10.15	9.32	17.7	17.7	17.7	17.7	17.7
Surinam	SUR	.	4.68			32.12	32.12	32.12
Chad	TCD	6.84	.		280	280	280	280
Togo	TGO	7.22	6.91			483	483	483
Trinidad and Tobago	TTO	8.77	7.45	85	85	85	85	85
Tunisia	TUN	8.48	6.45		61	61	61	61
Tanzania	TZA	6.25	6.64			280	280	280
Uganda	UGA	6.97	4.45			280	280	280
Uruguary	URY	9.03	7.00				71	71
USA	USA	10.22	10.00	15	15	15	15	15
Venezuela	VEN	9.07	7.14				78.1	78.1
Vietnam	VNM	7.28	6.41		140	140	140	140
South Africa	ZAF	8.89	6.86	15.5	15.5	15.5	15.5	15.5
Zaire	ZAR	6.87	3.50			280	280	280
El Salvador	SLV	7.95	5.00				78.1	78.1
Sierra Leone	SLE	6.25	5.82	483	483	483	483	483
Memo								
France				20.17	20.17	20.17	20.17	20.17
Britain				15.3	15.3	15.3	15.3	15.3
A blank indicates missing data								

Appendix Table A3
Coding of Historical Institutions in Polity III dataset

	First Year in Polity III data set	Institutions in First Year of Independence		Institutions in 1900		Institutions in 1970		Institutions in 1990	
		Democracy	Constraint on Executive	Democracy	Constraint on Executive	Democracy	Constraint on Executive	Democracy	Constraint on Executive
Afgan	1800	1	1	1	1	0	3	0	2
Algeria	1962	0	2	-99	-99	0	1	3	2
Angola	1975	0	3	-99	-99	-99	-99	0	3
Argentina	1825	1	1	3	3	0	1	8	6
Australia	1901	10	7	10	7	10	7	10	7
Bangladesh	1972	8	7	-99	-99	-99	-99	0	2
Benin	1960	4	3	-99	-99	0	1	0	1
Bhutan	1907	0	1	-99	-99	0	2	0	2
Bolivia	1825	1	3	4	3	0	1	8	7
Botswana	1966	10	7	-99	-99	10	7	10	7
Brazil	1824	1	1	1	3	0	1	10	7
Burkina Faso	1960	0	3	-99	-99	0	2	0	1
Burundi	1961	3	5	-99	-99	0	1	0	1
Cambodia	1949	0	3	-99	-99	0	1	0	3
Cameroon	1960	0	3	-99	-99	0	1	0	2
Canada	1867	7	7	9	7	10	7	10	7
Central African Fed.	1960	0	1	-99	-99	0	1	0	1
Chad	1960	0	1	-99	-99	0	1	0	1
Chile	1818	1	1	5	7	6	5	9	7
Colombia	1832	4	3	3	5	7	2	8	6
Comoros	1975	5	5	-99	-99	-99	-99	5	5
Congo	1960	5	4	-99	-99	0	3	0	2
Costa Rica	1838	0	3	10	7	10	7	10	7
Cote d'Ivoire	1960	0	1	-99	-99	0	1	0	2
Cuba	1901	4	3	-99	-99	0	1	0	1
Dominican Re	1844	1	3	1	3	1	3	7	6
Ecuador	1830	1	3	1	3	4	1	9	7
Egypt	1811	0	1	7	7	0	3	0	3
El Salvador	1841	1	3	1	3	3	4	7	5
Ethiopia	1946	0	1	5	7	0	1	0	2
Fiji	1970	9	7	-99	-99	9	7	6	7
Gabon	1960	0	2	-99	-99	0	1	0	2
Gambia	1965	10	7	-99	-99	10	7	10	7
Ghana	1960	0	1	-99	-99	5	5	0	1
Guatemala	1839	1	1	0	1	3	3	5	4
Guinea	1958	0	1	-99	-99	0	1	0	1
Guinea-Bassu	1974	0	3	-99	-99	-99	-99	0	3
Guyana	1966	5	2	-99	-99	4	2	5	4
Haiti	1820	0	1	0	1	0	1	0	1
Honduras	1839	1	3	6	5	1	3	6	5
India	1950	9	7	-99	-99	9	7	8	7
Indonesia	1945	0	1	-99	-99	0	2	0	2
Jamaica	1959	10	7	-9	-9	10	7	10	7
Kenya	1963	5	5	-99	-99	0	3	0	3
Laos	1958	8	5	-99	-99	0	3	0	3
Lesotho	1966	9	7	-99	-99	0	1	0	1
Liberia	1847	7	7	1	4	0	3	0	1
Libya	1951	0	3	-99	-99	0	1	0	1
Madagascar	1960	3	4	-99	-99	2	3	0	3
Malawi	1964	0	1	-99	-99	0	1	0	1
Malaysia	1957	10	7	-99	-99	4	3	8	7
Mali	1960	0	3	-99	-99	0	1	0	1
Mauritania	1960	2	5	-99	-99	0	3	0	3

	First Year in Polity III data set	Institutions in First Year of Independence		Institutions in 1900		Institutions in 1970		Institutions in 1990	
		Democracy	Constraint on Executive	Democracy	Constraint on Executive	Democracy	Constraint on Executive	Democracy	Constraint on Executive
Mauritius	1968	9	7	-99	-99	9	7	10	7
Mexico	1822	1	3	0	1	0	3	1	3
Morocco	1800	1	1	1	1	0	1	1	2
Mozambique	1975	0	2	-99	-99	-99	-99	0	3
Myanmar	1948	8	7	-99	-99	0	1	0	1
Namibia	1990	9	3	-99	-99	-99	-99	9	3
Nepal	1800	1	1	1	1	0	2	8	7
New Zealand	1857	10	7	10	7	10	7	10	7
Nicaragua	1838	1	1	1	1	0	1	0	3
Niger	1958	0	3	-99	-99	0	3	0	3
Nigeria	1960	8	7	-99	-99	0	1	0	1
Pakistan	1947	1	1	-99	-99	4	3	3	3
Panama	1903	0	3	-99	-99	0	1	8	7
Papua New Guinea	1976	10	7	-99	-99	-99	-99	10	7
Paraguay	1811	0	1	1	2	0	1	6	5
Peru	1821	1	1	4	3	0	1	8	7
Philippines	1935	6	3	-99	-99	4	3	8	7
Rwanda	1961	1	1	-99	-99	1	1	0	1
Senegal	1960	3	3	-99	-99	0	3	3	3
Sierra Leone	1961	6	5	-99	-99	3	4	0	3
Singapore	1959	8	7	-99	-99	2	3	2	3
Somalia	1960	7	7	-99	-99	0	1	0	1
South Africa	1856	3	3	3	3	7	7	7	7
Sri Lanka	1948	7	7	-99	-99	8	7	5	5
Sudan	1954	8	7	-99	-99	0	3	0	1
SWAZILAND	1968	3	2	-99	-99	3	2	0	1
Tanzania	1962	0	3	-99	-99	0	3	0	3
Togo	1960	0	3	-99	-99	0	1	0	1
Trinidad and Tobago	1962	8	7	-99	-99	8	7	9	7
Tunisia	1959	0	1	-99	-99	0	2	0	3
Uganda	1962	7	7	-99	-99	0	3	0	3
Uruguary	1830	1	1	1	3	9	3	9	3
USA	1800	7	7	10	7	10	7	10	7
Venezuela	1830	1	1	1	3	9	3	9	3
Vietnam	1954	0	1	-99	-99	0	3	0	3
Zaire	1960	0	1	-99	-99	0	1	0	1
Zambia	1964	4	3	-99	-99	3	2	0	1
Zimbabwe	1923	7	7	-99	-99	7	7	0	3
-99 indicates that a country was not independent at this date (and therefore not included in the Polity III dataset).									
For our regressions, colonies were coded as having the lowest possible score on Democracy (0) and Constraint on the executive (1).									

Appendix Table A4
IV Regressions Using Alternative Measure of Institutions

	Base Sample	Base Sample	Base Sample without neo-Europes	Base Sample without neo-Europes	Base Sample without Africa	Base Sample without Africa	Africa only	Africa only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constraint on Executive in 1990	0.46	0.39	0.39	0.36	0.73	0.62	0.37	0.10
	(0.08)	(0.08)	(0.09)	(0.09)	(0.26)	(0.23)	(0.22)	(0.3)
Latitude		1.6		1.2		1.7		2.8
		(0.84)		(0.94)		(1.30)		(2.1)
				First-Stage for Expropriation Risk				
Log Mortality	-1.2	-1.2	-1.1	-1.1	-0.93	-1.0	-0.53	-0.36
	(0.20)	(0.24)	(0.24)	(0.25)	(0.35)	(0.38)	(0.23)	(0.27)
Latitude		-0.01		-1.30		-1.1		3.4
		(2.1)		(2.5)		(2.4)		(3.0)
Adjusted R-Squared	0.32	0.31	0.24	0.23	0.14	0.13	0.13	0.14
N	68	68	64	64	37	37	31	31

Dependent Variable in Second-Stage Estimates is log GDP per capita (PPP basis) in 1995
Standard errors are in parentheses

Figure 1

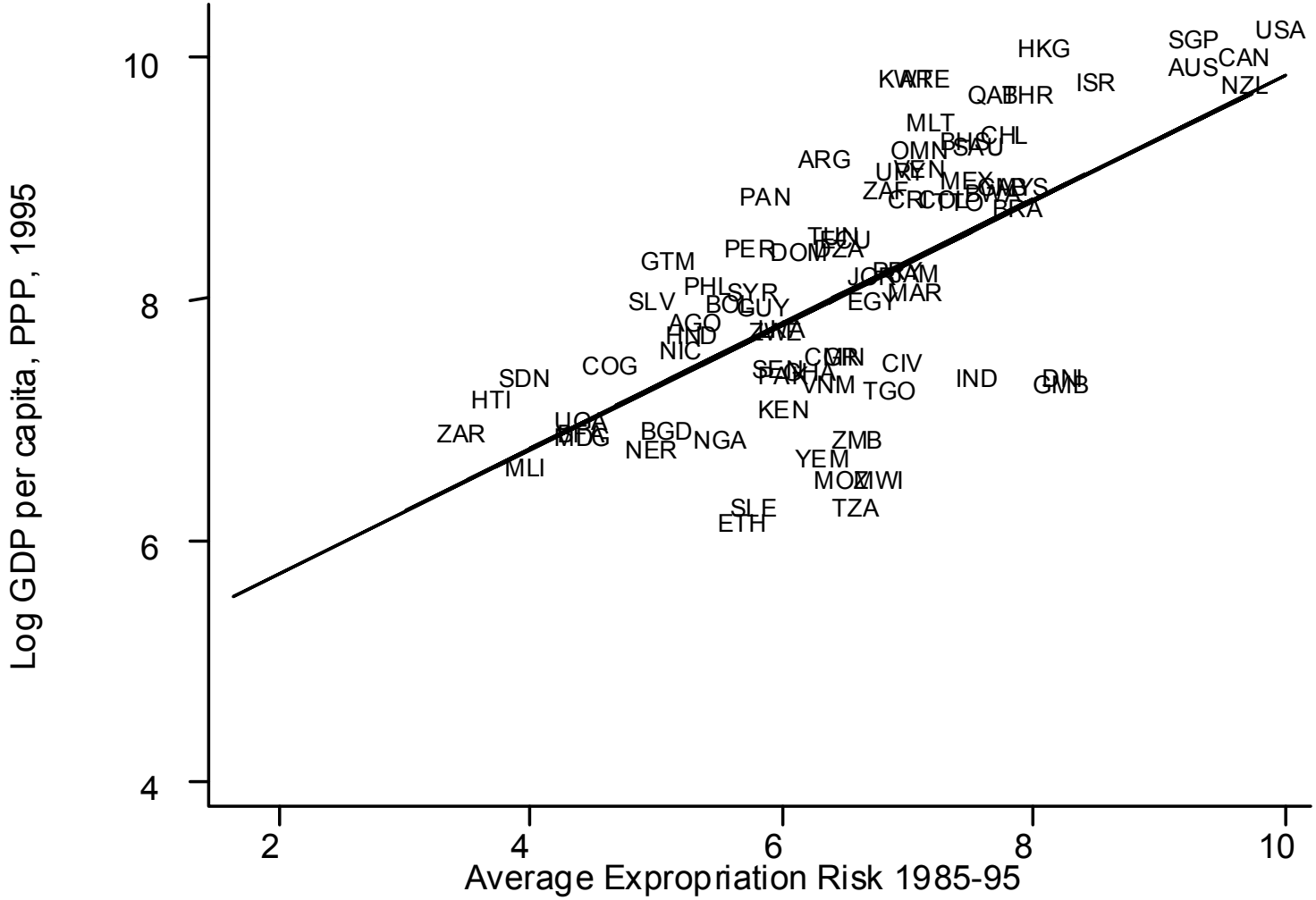
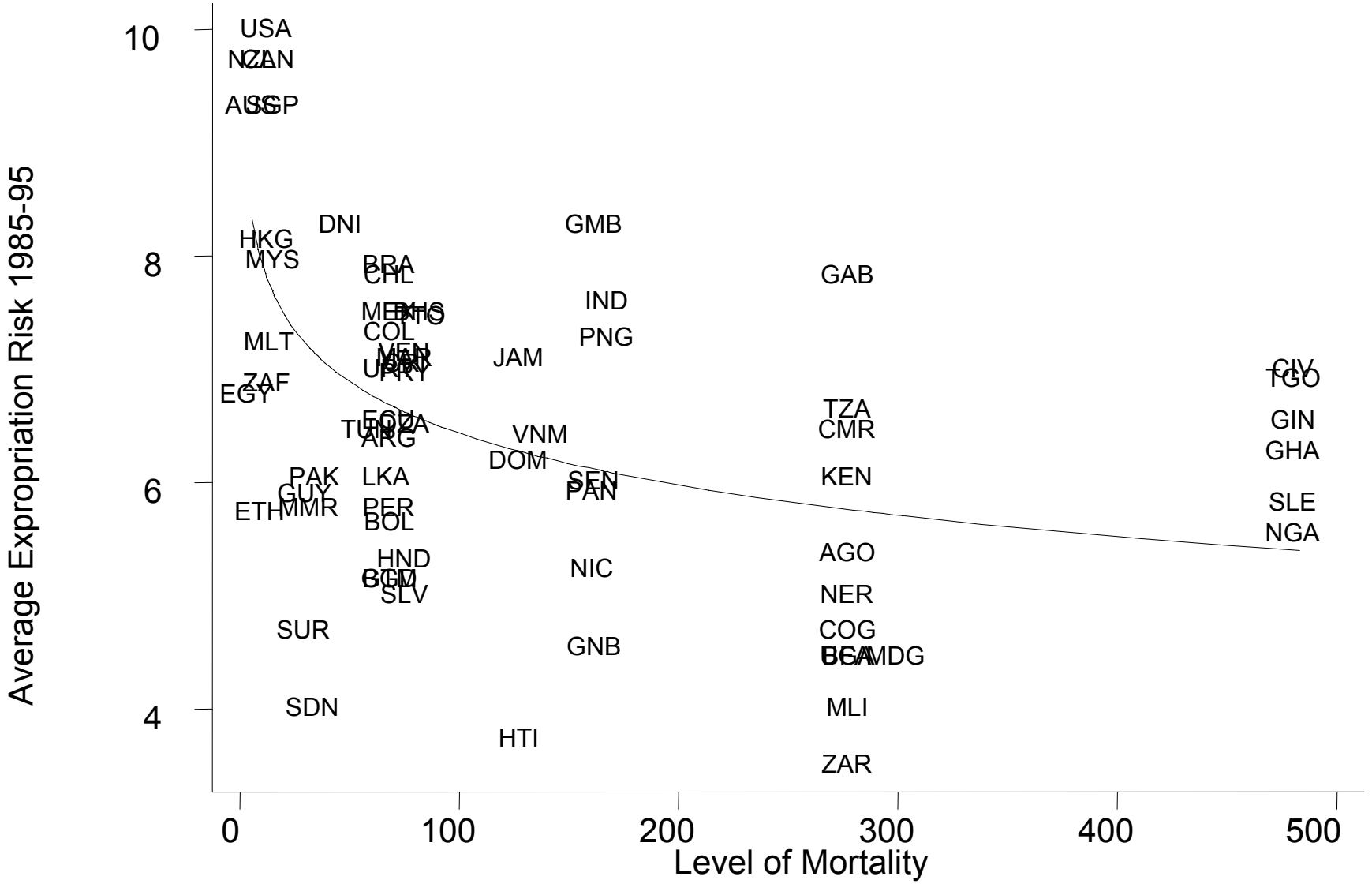

Log GDP per capita, PPP, 1995
10
8
6
4
2
4
6
8
10
Average Expropriation Risk 1985-95
USA
SGP
HKG
AUS
CAN
NZL
ISR
QAT
BHR
MLT
CHL
OMN
ARG
PAN
MEX
PER
DOM
GTM
PHL
SYR
SLV
BOL
GUY
EGY
MAR
AGO
NIC
COG
CIV
IND
SDN
HTI
TGO
KEN
ZAR
BGD
NGA
ZMB
YEM
MLI
TZA
ETH

Figure 2


Average Expropriation Risk 1985-95
10
8
6
4
0
100
200
300
400
500
Level of Mortality
USA
NZL
CAN
AUS
SGP
DNI
GMB
HKG
MYS
BRA
CHL
GAB
IND
COL
PNG
MLT
JAM
ZAF
EGY
TZA
CMR
VNM
ARG
DOM
PAK
LKA
KEN
GUY
PAN
SEN
ETH
MMR
PER
BOL
AGO
HND
NIC
SLV
NER
SUR
COG
GNB
MDG
SDN
MLI
HTI
ZAR
CIV
TGO
GIN
GHA
SLE
NGA