

NBER WORKING PAPER SERIES

THE COLONIAL ORIGINS OF COMPARATIVE DEVELOPMENT:  
AN INVESTIGATION OF THE SETTLER MORTALITY DATA

David Y. Albouy

Working Paper 14130  
<http://www.nber.org/papers/w14130>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2008

I thank Raj Arunachalam, Raphael Auer, Pranab Bardhan, Christina Berkley, Chris Blattman, David Card, Brad DeLong, Gregory Clark, William Easterly, Rob Gillezeau, Tarek Hassan, Jim Hines, Chang-Tai Hsieh, Michael Jansson, Chad Jones, Annalisa Leibold, Ian McLean, Ted Miguel, Kris Mitchener, Robert Moffitt, Marcelo Moreira, Maurice Obstfeld, Rohini Pande, Gerard Roland, Christina Romer, David Romer, Emmanuel Saez, Andrei Shleifer, Francesco Trebbi, four anonymous referees, and the participants at the Berkeley Development Lunch and the Economic History and Macroeconomics Seminars for their help, input, and advice. I am particularly grateful to Daron Acemoglu, Simon Johnson, and James Robinson for providing me with data, and for sharing with me a preliminary response and later formal responses to my work. Any mistakes are my own. Please e-mail any comments to [albouy@umich.edu](mailto:albouy@umich.edu). The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2008 by David Y. Albouy. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality  
Data

David Y. Albouy

NBER Working Paper No. 14130

June 2008

JEL No. I12,N10,O11,O57,P16,P51

**ABSTRACT**

In a seminal contribution, Acemoglu, Johnson, and Robinson (2001) argue property-rights institutions powerfully affect national income, using estimated mortality rates of early European settlers to instrument capital expropriation risk. However 36 of the 64 countries in their sample are assigned mortality rates from other countries, typically based on mistaken or conflicting evidence. Also, incomparable mortality rates from populations of laborers, bishops, and soldiers - often on campaign - are combined in a manner favoring their hypothesis. When these data issues are controlled for, the relationship between mortality and expropriation risk lacks robustness, and instrumental-variable estimates become unreliable, often with infinite confidence intervals.

David Y. Albouy

Department of Economics

University of Michigan

611 Tappan Street

351C Lorch Hall

Ann Arbor, MI 48109-1220

and NBER

albouy@umich.edu

Acemoglu, Johnson, and Robinson's seminal paper (2001) – henceforth AJR – has reinvigorated debate over the relationship between property rights and economic growth. Following research by Knack and Keefer (1995), Mauro (1995), La Porta et al. (1998), Hall and Jones (1999), Rodrik (1999) and others, AJR endeavor to determine the causal effect of institutions that protect property rights, measured by risk of capital expropriation, on economic performance. This endeavor is complicated by the fact that the correlation between institutional and economic measures may reflect the reverse influence of economic growth on institutions or the simultaneous influence of omitted variables on both economic output and institutions. To circumvent these problems, AJR use an instrumental variable (IV) for expropriation risk in an equation determining GDP per capita across previously colonized countries.

AJR argue that during the colonial era, Europeans were more likely to settle in places where they had a lower risk of dying from disease. Colonies in which Europeans settled developed institutions that protect property better than colonies where Europeans did not settle. The authors argue that, in the long run, the direct effects of mortality and European settlement on national income faded, while the indirect effect through property-rights institutions persisted. Their argument motivates the use of potential European settler mortality rates as an instrument for the risk of capital expropriation. AJR's IV estimates of the effect of expropriation risk on GDP per capita are large, explaining much of the variation in income across countries.

The historical sources containing information on mortality rates during colonial times are thin, which makes constructing a series of potential European settler mortality rates challenging. AJR construct their series by combining the mortality rates of soldiers (Curtin 1989, 1998), laborers (Curtin 1995), and bishops (Gutierrez 1986). Researchers have been eager to use this new series, particularly given its promise as an instrumental variable for institutions. Currently, over twenty published articles, and many more working papers, use AJR's settler mortality data.

This paper argues that despite AJR's ingenuity and diligence, there are a number of reasons to doubt the reliability and comparability of their European settler mortality

rates and the conclusions which depend on them. First, out of 64 countries in their sample, only 28 countries have mortality rates that originate from within their own borders. The other 36 countries in the sample are assigned rates based on AJR's conjectures as to which countries have similar disease environments. These assignments are based on weak and sometimes inaccurate foundations. Six assignments are based upon AJR's misunderstanding of former names of countries in Africa. Another sixteen assignments are based on a questionable use of bishop mortality data in Latin America from Gutierrez (1986), which are based on 19 deaths. Additionally, AJR use the bishop rates multiplied by a factor of 4.25, a procedure that appears to contradict evidence in their own sources. At a minimum, the sharing of mortality rates across countries requires that statistics be corrected for clustering (Moulton, 1990). This correction noticeably reduces the significance of AJR's results. If, in the hope of reducing measurement error, AJR's 36 conjectured mortality rates are dropped from the sample, the empirical relationship between expropriation risk and mortality rates weakens substantially, particularly in the presence of additional covariates.

Second, AJR's mortality rates never come from actual European settlers, although some settler rates are available in their sources. Instead, AJR's rates come primarily from European and American soldiers in the nineteenth century. In some countries, AJR use rates from soldiers at peace in barracks, while in others, they use rates from soldiers on campaign. Soldiers on campaign typically have higher mortality from disease, and AJR use campaign rates more often in countries with greater expropriation risk and lower GDP. Thus, AJR's measures of mortality artificially favor their hypothesis. In a few countries, AJR use the maximum mortality rates of African laborers, although these do not appear comparable with average soldier mortality rates. Controlling for the source of the mortality rates weakens the empirical relationship between expropriation risk and mortality rates substantially. Furthermore, if these controls are added and the conjectured data are removed, the relationship virtually disappears. Additional data provided by AJR in their Response (2005) do not restore this relationship.

Without a robust relationship between expropriation risk and mortality rates,

AJR's IV estimates of the effect of expropriation risk on GDP per capita suffer from weak instrument problems: point estimates are unstable, and corrected confidence intervals are often infinite.

Lastly, AJR's (2006) defense that their results hold when African observations are removed is not reassuring. Without conjectured mortality rates, the sample without Africa contains only 13 observations, and the relationship between mortality and expropriation risk rests entirely on the inclusion of the "Neo-Europes," which do not seem to belong in the sample.

### **I. Problems with the Settler Mortality Data**

AJR construct their mortality rates in four steps, as described in their data appendix. In their first step they take average mortality rates from a table in Curtin (1989, pp. 7-8) of European soldiers from disease (not combat) in the early to mid-nineteenth century. In step two, AJR add new countries to their sample by using average mortality rates from a selection of military campaigns in Curtin (1998). AJR state that when more than one rate is available, they take the earliest rate. In step three, they add peak mortality rates from Curtin et al. (1995) of African laborers who were moved to foreign disease environments. Also in step three, AJR assign mortality rates to neighboring countries which they believe to have similar disease environments. Finally in the fourth step, AJR take the mortality rates of Latin American bishops in the seventeenth and eighteenth centuries from Gutierrez (1986), multiply them by a factor of 4.25 to conform to a rate taken from a campaign in Mexico, and apply them to sixteen countries.

Mortality rates are expressed in the number of deaths per year per thousand at risk, and are catalogued in Table A1. In order to keep the discussion here brief, considerable detail is left to my Appendix, available on the *NBER* website.

#### *A. The Matching of Mortality Rates to Neighboring Countries*

AJR extend their limited data to 64 countries. They state they assign "a mortality number to a country if it neighbors a country for which we have data and has the same disease environment," (Data Appendix, p. 3). However, AJR provide little explanation of how

they determine whether countries share similar disease environments. In regions such as sub-Saharan Africa and Southeast Asia, neighboring countries in AJR's data have widely differing mortality rates, so their series is sensitive to how they choose neighboring countries.

AJR argue that large differences in mortality occur between neighboring countries "because there exists substantial variation in disease environment, particularly for malaria, even in neighboring areas" citing differences in microclimates (Data Appendix, p.1).<sup>1</sup> Yet, substantial differences in disease environments undermine AJR's strategy of assigning mortality rates to neighboring countries. With the paucity of information they present, AJR cannot reasonably defend how they assign such different rates to some neighboring countries, and then share the same rates across others. If disease environments vary little across neighboring countries, then much of the variation in AJR's data is due to measurement error, and true mortality rates are likely collinear with other variables suspected to affect institutions or GDP.

One set of mortality assignments, illustrated in Figure 1, comes from mortality rates which are all from French campaigns in western Mali, reported in Curtin (1998). These assignments are difficult to explain, but appear to originate from a misunderstanding of changing geographic names for Mali, as explained in my Appendix. Summarizing briefly,

- Mali is assigned a rate of 2940, which is annualized from a severe yellow fever epidemic that killed 49 percent of an expeditionary force over two months in 1878 (p.81).
- Niger is assigned a rate of 400, from 1880 to 1883 (p. 85),
- Burkina Faso, Cameroon, Gabon, Angola and Uganda are assigned a rate of 280, from 1883 to 1884 (p. 238).

There are two fundamental problems with these assignments. First, since all three rates come from western Mali, there is no basis for assigning each of these rates to different

---

<sup>1</sup> This comment arises when AJR assign a rate of 17.7 to Malaysia and 170 to neighboring Indonesia. In fact, Curtin (1989, pp.17-18) does not ascribe this difference to microclimates, but rather to the fact that soldiers were at war in Indonesia.

countries. Second, and more fundamentally, there is no justification for assigning rates from Mali to countries as far away as Angola and Uganda. The six countries with rates taken from Mali have neighbors with widely varying rates, from 78.2, in Algeria (which borders Niger) to 2004 in Nigeria (which borders Niger and Cameroon). This large variation implies that assigning mortality rates from neighboring countries is very sensitive to choice.

The differing rates from Mali raise the question of what rate properly represents it. According to Curtin (p. 81), the rate of 2940 is an overestimate: because of acquired immunity, “the annual rate and the rate of loss over two months [490] would have been about the same.” The second rate of 400 is not representative either as it is “unusually high because it included the deaths from yellow fever of soldiers who stayed in Saint Louis [on the coast of Senegal]” (Curtin p. 84). Thus, the third rate of 280 seems to be the first available rate that represents Mali.<sup>2</sup>

AJR’s assignment of mortality rates to sixteen Latin American countries based on thin data from bishops in Gutierrez (1986) is also worrisome. Gutierrez does not provide mortality rates by country: rather, he categorizes cities with bishops into low, medium, and high temperature regions, admitting that he only assumes that cities with similar temperatures have similar disease environments.<sup>3</sup> It is AJR who assign the countries to the three regions.

The bishop rates AJR use (p. 39) are based on 4, 5, and 10 deaths out of at-risk populations of 24, 28.5, and 30.5 bishops in each region over ten years, implying mortality rates of 16.7, 17.5 and 32.8. These rates are not significantly different from each other, or from mortality rates of similarly-aged contemporary males in Sweden of 18.32 (Sundärg, 1905), or from soldiers in barracks (15.3) in England or France (20.17)

---

<sup>2</sup> Curtin (p. 87) singles out a lower rate of 200.24 (1883 to 1888) as campaign rate useful for comparison with a barracks rate from Senegal. Using that rate instead of 280 only strengthens the results below.

<sup>3</sup> Gutierrez states (p. 33, my translation) “we cannot study in a profound way the influence of climate on the mortality of Latin-American bishops in the seventeenth and eighteenth centuries, given the small number of observations, the diversity of environmental situations of which we do not know well the characteristics, and finally the lack of knowledge of the diseases which could affect adults having survived the perils of diseases in infancy and youth.”

(Curtin 1989, p. 7).<sup>4</sup> In his abstract, Gutierrez (1986) writes that the life expectancy at age 40 for bishops was 20.3 years in Latin America relative to 29 years in France, implying mortality was about 43 percent higher than in Europe, with the difference due mainly to the high mortality region. Also, bishops in Latin America born in Europe died at rates slightly lower than those born in the New World.

Although this evidence suggests that mortality in Latin America was not much higher than in Europe, AJR scale up all of the bishop rates by 325 percent. AJR's justification for this adjustment is that campaigning French soldiers in Mexico from 1862 to 1863 incurred a mortality rate of 71, 4.25 times the low-temperature bishop rate of 16.7.<sup>5</sup> In defense of this "benchmarking" method, AJR (2001, p. 1383) claim that "alternative methods produce remarkably similar results." However, as I document in my Appendix, using similar assumptions, alternative benchmarking methods produce remarkably dissimilar results. Of the many methods possible, AJR report those that produce relatively high rates.<sup>6</sup>

The countries with mortality rates inferred from Mali and Mexico account for 22 of the 36 countries with conjectured rates. There are other problems with the remaining 14. For Hong Kong, once called "an unhealthy, pestilential, unprofitable and barren rock" (Cantlie, 1974, p. 480), AJR use a rate of 14.9, belonging to a British force that in the summer of 1860 campaigned close to Beijing.<sup>7</sup> Also, this rate applies only to the duration of the campaign, not the year. As AJR report in their Response (2005, p. 32), British soldiers in Hong Kong in peacetime died at a rate of 285 from 1842 to 1845 (Tulloch, 1847, p. 254), 19 times AJR's original assignment.<sup>8</sup>

---

<sup>4</sup> An *F*-test that all three regions have the same mortality rate is not rejected at a level of 12 percent.

<sup>5</sup> AJR's extrapolation appears incorrect given their assumptions. First, the rate from Mexico is not annualized; a more accurate rate, based on the annualized troop strengths in Mexico reported in Reynaud (1898), is 61. Also, the French soldiers spent more time in Veracruz, a high temperature area, than in Mexico City, a low temperature area (Reynaud, 1898, pp. 102-22). Benchmarking the annualized rate to the high temperature area lowers the scaling factor from 4.25 to 1.86.

<sup>6</sup> In their Response (2005, p. 35), AJR propose a benchmarking system which produces a mortality rate for low-temperature regions of 15.4, close to the original bishop mortality rate of 16.7.

<sup>7</sup> The soldiers did assemble briefly in Hong Kong, but left before the pestilential summer (Elleman, 2001).

<sup>8</sup> AJR cite many valuable additional sources in their Response (2005), including Tulloch (1847), Cantlie (1974), and others mentioned in my Appendix.

## B. *Campaigning Soldiers and African Laborers*

The cited works by Curtin are concerned primarily with the health and mortality of soldiers during the European conquests of the nineteenth century.<sup>9</sup> Accordingly, he took as given the current circumstances and living conditions of the soldiers when comparing their mortality rates. These rates do not necessarily provide a good proxy for potential European settler mortality, which would ideally compare settlers with similar living conditions, subject to the constraints imposed by their environments. Living conditions have a large effect on mortality rates from disease. Curtin (1989, pp. 40-61) discusses how clean water and adequate sewage disposal can drastically lower mortality rates from waterborne diseases, such as typhoid and other gastrointestinal infections. Adequate shelter, nutritious food, and quinine prophylactics – long known to protect against malaria – also lower mortality from disease.

Variation in disease due to living conditions seriously affects AJR's mortality data. One reason for this is that they use the mortality rates (from disease alone) of soldiers in barracks from some countries, and rates from soldiers on campaign from others, without adjustment. Yet Curtin carefully distinguishes between what he terms "barracks rates" and "campaign rates," asserting (1989, p. 4) that "one of the fundamental facts of military medical experience [is that] troops in barracks are much healthier than troops on campaign, even disregarding losses from combat." Soldiers on campaign took fewer precautions against disease and were less likely to have safe water, fresh food, decent shelter, or sewage disposal. Consequently, "The disease toll for soldiers on campaign was inevitably higher than it was in peacetime," (Curtin, 1998, p. xi).

In his writing, Curtin usually discusses whether a mortality rate is from a campaign or not, making it possible to code a variable indicating which of AJR's rates are from a campaign, as discussed in my Appendix. For a given country, campaign rates tend to be higher than barracks rates, although there is no stable relationship between the two. Curtin (1998, pp. 221-4) documents how during campaigns mortality from malaria

---

<sup>9</sup> This is evident in Curtin (1989, p. xiii): "This book is a quantitative study of the relocation costs among European soldiers in the tropics between about 1815 and 1914," and the title of Curtin (1998), *Disease and Empire: The Health of European Troops in the Conquest of Africa*.

typically increases by more than 100 percent, from gastrointestinal infections by more than 200 percent, and from typhoid by more than 600 percent, resulting in mortality rates 66 to 2000 percent higher than barracks rates.<sup>10</sup> Even in Europe, where barracks rates are usually below 25 (Curtin, 1989, p. 5), campaign rates rose as high as 332, seen by the British in the Netherlands in 1809 (Balfour, 1845, p. 198).<sup>11</sup>

The distinction between barracks and campaign rates affects the analysis as AJR use campaign rates more often in countries with high risk of capital expropriation and low GDP per capita.<sup>12</sup> Thus, measured mortality rates are endogenous: places with lower future security of property rights and lower output per capita essentially suffer from positive measurement error in their mortality rates. This creates artificial support for AJR's hypothesis that mortality is negatively correlated with expropriation risk and GDP per capita.<sup>13</sup>

The effects of campaigning on mortality are evident in North Africa, where according to Curtin (1989, p. 17) mortality is similar to Southern Europe in more peaceful conditions, as seen in AJR's rate of 16.3 for Malta.<sup>14</sup> Instead, AJR use campaign rates about four times as high: 63 for Tunisia, 67.8 for Egypt, and 78.2 for Algeria and Morocco. Most of these deaths were from typhoid and other digestive diseases, with malaria playing a minor role (Curtin, 1989, p. 36; 1998, pp. 152, 158,

---

<sup>10</sup> Curtin's distinction is only two-fold: he uses the terms "peacetime" and "barracks" interchangeably, as he does with the terms "campaign" and "expedition." AJR's three-fold distinction in their later (2005, 2006) work between what they call "peacetime," "expedition," and "wartime" rates is their own, not Curtin's. AJR claim that peacetime and expedition rates are comparable, contrary to Curtin's views, but not with wartime rates. AJR's distinction seems inappropriate since higher mortality rates during expeditions and wartime are primarily due to living conditions which differ from those in barracks. Furthermore, the rates AJR use for Algeria, Indonesia, Mexico, and Sudan are from violent conflicts, which seem worthy of the term "wartime," despite AJR's claims that they do not use wartime rates.

<sup>11</sup> This source is used in AJR (2005), although they do not mention these rates.

<sup>12</sup> At a 2 percent size one rejects the null hypotheses that either expropriation risk or log GDP per capita is unrelated to variable indicating when a country's rate is taken from a campaign.

<sup>13</sup> AJR (footnote 17) admit that their data contain measurement error, but state that "this measurement error does not lead to inconsistent estimates of the effect of institutions on performance." This is true only if measurement error is uncorrelated with the error term in the equation determining log GDP per capita.

<sup>14</sup> "Climatically the south shore of the Mediterranean was much like the north shore in Italy or southern France...The high Algerian figure [78.2] in the 1830s was certainly the result of campaigning in the conquest period. Within a decade or so, the Algerian death rate was close to the rates of the Mediterranean islands." AJR (2005, p. 22) disagree with my interpretation of this passage.

169).<sup>15</sup>

A related difficulty arises as AJR inconsistently combine campaign and barracks rates in the second step of their data construction. AJR state two different rules for how they select their data: in their original paper, AJR (Data Appendix, p. 2) state that, “Whenever Curtin provides more than one estimate, we use the earliest available number.” In their Response (2005), AJR state they take the earliest peacetime rate if a peacetime rate is available, otherwise they use the earliest “expedition” rate. Yet, as discussed further in the Appendix, for Sudan, Egypt, and Madagascar AJR choose rates from Curtin (1998) which violate both of these stated selection rules, as they are from campaigns and are not the earliest rates available – although they are the highest rates available.<sup>16</sup> As documented in the Appendix, these inconsistent choices strengthen the empirical relationship between measured mortality and expropriation risk, further justifying the need to control for the effects of campaigning on measured mortality rates.

Another source of incomparability comes from AJR’s use of mortality rates from African laborers, coerced to move to foreign environments under harsh conditions (Curtin et al., 1995, pp. 463, 491). Comparing rates in Curtin (1968), AJR argue that the laborer rates provide a lower bound for soldier rates, as black soldiers in Africa had lower average mortality rates than white soldiers. There are two problems with this argument. First, it uses the mortality of black soldiers, not black laborers. Second, the rates referred to are average rates, but AJR instead use *maximum* rates available for laborers: for the Congo they choose a maximum rate of 240 over an average rate of 100; for Kenya they use a maximum rate of 145, as no average is reported.<sup>17</sup>

---

<sup>15</sup> Deaths from digestive diseases also play a large role in the rates for Mexico, India, and Vietnam. This may have more to do with preexisting poverty than with climate: Curtin (1998, p. 113) writes “Typhoid had become a ‘tropical disease’ – because the tropical world is poor, not because of climate.” Earle (1979, p. 119) estimates that in Virginia from 1618 to 1624, British settlers suffered a mortality rate of 283, primarily from dysentery and typhoid. This is far less than the later barracks rate of 15 AJR use for the United States. Because of these diseases, AJR disregard actual settler mortality rates, mentioned in Curtin (1998, p. 116), but use similarly impacted mortality rates from campaigning soldiers in poorer countries.

<sup>16</sup> AJR (2005, n. 16) contend that they never changed their stated rule of choosing the first available rate, as this always meant the first available peacetime rate, stating “We thought this was obvious.” However, neither rule is applied consistently. Furthermore, AJR use a campaign rate to benchmark the bishop rates.

<sup>17</sup> Quotations from Curtin, (1995) in my Appendix make it clear that the mortality rates are maxima.

## II. Sensitivity of AJR’s Empirical Results

The above discussion raises questions about any empirical results based on AJR’s mortality data. For the sake of brevity, only results from AJR’s original article (2001) are examined here.

AJR’s econometric model can be written as the combination of a first-stage equation  $r_i = \beta m_i + v_i$  and a second-stage equation  $y_i = \alpha m_i + \varepsilon_i$ , where  $i$  indexes colonial countries,  $y_i$  is log GDP per capita,  $r_i$  is expropriation risk,  $m_i$  is log potential settler mortality, and  $v_i$  and  $\varepsilon_i$  are error terms, with  $E[m_i v_i] = 0$  by construction.<sup>18</sup> IV estimates require an instrument which is *relevant* ( $\beta \neq 0$ ) and *excludable* ( $E[m_i \varepsilon_i] = 0$ ). Letting  $\pi = \alpha\beta$  and  $\zeta_i = \alpha v_i + \varepsilon_i$ , the reduced form of the second stage equation is given by  $y_i = \pi m_i + \zeta_i$ . By the principle of indirect least squares, the IV estimator of  $\alpha$  is the ratio of the OLS estimates of  $\pi$  and  $\beta$ , i.e.  $\hat{\alpha}_{IV} = \hat{\pi}_{OLS} / \hat{\beta}_{OLS}$ . The analysis here first considers the OLS estimate of  $\beta$ , and afterwards the IV estimate of  $\alpha$ .

Because mortality rates are shared by countries, the residuals are correlated because of clustering effects (see Moulton, 1990). This invalidates the conventional standard errors and test statistics used by AJR. The standard procedure used to correct for these clustering effects, as well as heteroscedasticity (Wooldridge, 2001, pp. 152, 191), is applied below.<sup>19</sup>

More fundamentally, it is worthwhile to examine how sensitive AJR’s results are to robustness checks that account for the weaknesses in the data documented above. One check is to drop countries with conjectured mortality rates that originate from outside their own borders – including the “benchmarked” Latin American data – and to replace Mali’s rate of 2940 with the more representative rate of 280.<sup>20</sup> If AJR’s theory is true,

---

<sup>18</sup> Control variables may be accounted for by having all of the above variables refer to the residual projections of the original variables, after being regressed on the control variables.

<sup>19</sup> AJR do not report clustered standard errors although they mention in their footnote 18 that clustering has “little effect on the standard errors.” See Table 1, Panel A, for the differences.

<sup>20</sup> The countries kept in this check do not correspond to the countries kept in columns (3) and (4) of AJR’s Appendix Table A5 labeled “Earliest Available Data”, with 30 observations (31 in their NBER Working Paper), and which is supposed to correspond to the rates derived from their first two steps. AJR’s sample retains Niger, Burkina Faso, Guyana, and Singapore although their rates are from elsewhere, while they omit Ghana and Nigeria, whose rates are native. I also retain Congo and Kenya, since the African laborer

their results should continue to hold in the smaller sample without the conjectured rates.

A second robustness check, to deal with AJR's use of different data sources, is to add two control variables which indicate when mortality rates are taken from campaigning soldiers or from imported African laborers. These controls weaken AJR's results, indicating that comparability problems in AJR's data indeed bias their results towards their conclusion.

#### A. *First-stage Estimates*

Table 1 presents the first-stage estimates obtained when one applies the two checks described above, using the types of controls found in AJR's original paper. The first five columns use geographic controls: latitude (measured in absolute degrees), continent dummies (Asia, Africa, and "Other," with the Americas as the reference), and omitting "Neo-Europes" (Australia, Canada, New Zealand, and the United States). These correspond to Columns (1), (2), (3), (7), and (8) in Table 4 of AJR's paper. The specification in column (6) adds climate controls from Parker (1997), similar to AJR's Table 6, column (1), except that it is more parsimonious, using only mean temperature and minimum monthly rain, rather than four temperature and four humidity variables. Column (7) controls for the percentage of the population of European descent in 1975, like AJR's Table 6, column (3). Column (8) controls for the percentage of the population living where *falciporum* malaria is endemic in 1994, as in AJR's Table 7, column (1).

The first-stage results with the original data in Panel A report that log mortality is usually a significant predictor of expropriation risk, although the clustered standard errors are larger than the homoscedastic ones, making  $\beta$  insignificant at the 10 percent level in columns (5) and (6).

In Panel B, the first robustness check is applied, dropping conjectured rates and correcting the Mali rate. Normally, using a more accurate sample should reduce measurement error, counteracting the effects of attenuation, and raising the point estimate of  $\beta$ . The opposite occurs here as the estimate of  $\beta$  falls, which should not occur unless

---

rates are derived directly from these countries. These countries should be omitted from AJR's check, since they are added in their third step, yet they retain Congo. Gabon is not in AJR's Appendix Data Table (A2).

the relationship between mortality and expropriation risk is stronger in countries with conjectured rates.<sup>21</sup> The standard errors also widen, but not drastically with the clustering correction. Altogether,  $\beta$  is only significant at the 18 percent level in all of the specifications with controls in Panel B.

With their original sample, AJR find that most control variables, with the exception of latitude, are not significant and do not affect their estimates of  $\beta$ . Accordingly, AJR only consistently use latitude as a control variable. Yet, when the conjectured mortality rates are dropped, all of the control variables grow appreciably in significance, while the point estimates of  $\beta$  are smaller with the controls. AJR's conjectured mortality rates diminish the importance of the control variables, which, in the more reliable subsample, appear collinear with mortality.

Using the original sample again, Panel C demonstrates that controlling for whether a mortality rate comes from soldiers on campaign or from African laborers also reduces the estimate for  $\beta$ , which is insignificant at the 5 percent level in all specifications with controls. However, the campaign and laborer dummies themselves are generally insignificant.

As shown in panel D, without conjectured rates, these dummies become significant, as do several other control variables. With both data checks in place, the estimates of  $\beta$  fall to very low levels, becoming insignificant even in column (1) without controls, and switching signs in columns (5), (6) and (8).<sup>22</sup> In conclusion, when either robustness check is applied, the relationship between expropriation risk and mortality loses robustness with control variables; with both checks combined, it loses robustness even without controls.

Data revisions using new rates from AJR's Response (2005), discussed in my Appendix, do not restore their hypothesis in the presence of these data checks, as seen in

---

<sup>21</sup> Results without the Mali correction, given in Table A3, are still not highly significant. Also, first-stage significance is greatly reduced if Mali is corrected and only other countries with Mali-based rates, shown in Figure 1, are dropped. Results in Table A4 reveal that if unadjusted bishop mortality rates are used in Latin America, first-stage significance falls more than if the countries are simply dropped.

<sup>22</sup> To ensure that results are not dependent on using expropriation risk as the measure of institutions, my Appendix Tables A5 and A6 show results using alternate measures – Constraint on Executive in 1990 and Law and Order Tradition in 1995. These estimates reveal a similar lack of robustness and significance.

Panel E.<sup>23</sup>

*B. Instrumental Variable Estimates*

When the first-stage estimate of  $\beta$  is not significantly different from zero – a common occurrence in the results seen so far – the relevance assumption needed for IV estimates ( $\beta \neq 0$ ) is not guaranteed, causing a *weak instrument* problem. This introduces a number of statistical pathologies to the IV estimates. Most importantly, inference based on the IV estimate using conventional asymptotic confidence regions (point estimate  $\pm t \times$  standard error), based on the Wald statistic, can be grossly incorrect (Dufour, 1997). Confidence regions for  $\alpha$  of the correct size can be built by inverting the AR statistic proposed by Anderson and Rubin (1949). While using the AR statistic seems unorthodox – producing asymmetric, and sometimes disjointed and unbounded confidence regions – it provides an exact test as appropriate as  $t$ -statistics in OLS, and provides correct inference in the presence of a weak instrument. When an instrument is strong, AR and Wald confidence regions are similar, as the latter is not grossly incorrect.<sup>24</sup>

Table 2 presents the IV estimates and confidence regions corresponding to the first-stage results in Table 1. In Panel A with the original data, weak-instrument problems appear despite the stability of the point estimates. In columns (1) and (2), where the first stage is fairly strong, the AR and Wald 95 percent confidence regions are fairly similar. However, as the instrument weakens in columns (3) and (4), the AR confidence regions widen, until in columns (5), (6), and (8) they become unbounded: as the indirect least squares formula  $\alpha = \pi/\beta$  implies, once zero cannot be rejected for  $\beta$ , infinity cannot be rejected for  $\alpha$ .

As the robustness checks are applied in panels B through D, these weak

---

<sup>23</sup> New data shown in Table A7; results without the data checks, or one at a time, are in Table A8.

<sup>24</sup> Moreira (2003) proves that, in the exactly identified case, AR tests are the uniformly most powerful amongst unbiased tests. The AR confidence regions are said to have “95 percent confidence” because they have 5 percent size. It does not mean that the true  $\alpha$  is within this region 95 percent of the time, but that the AR statistic computed is within the first 95 percent of the cumulative distribution of the statistic under the null hypothesis. With a weak instrument, Staiger and Stock (1997) show that conventional  $F$ -tests of significance for exogenous variables and over-identification tests (e.g. Sargan, 1958) for the second stage are invalid. Correctly specified tests depend on parameters which cannot be estimated. Since mortality is a weak instrument in most cases, these test statistics are not reported to save space.

instrument problems are aggravated: point estimates become unstable and the confidence regions expand until most of the regions in Panels D and E equal the entire real line. Furthermore, the estimates of  $\alpha$  are sometimes implausibly large. A value of  $\alpha$  equal to two implies some incredible conclusions: e.g. if Mexico and the United States had the same property rights (a 2.5 point difference) then the GDP per capita ratios of the two countries would go from less than one third to over 40 in Mexico's favor. In other cases, the estimate of  $\alpha$  becomes large and negative, as the estimate of  $\beta$  becomes small and positive, while the reduced-form estimate of  $\pi$  remains negative.

The volatile estimates and unbounded confidence regions for  $\alpha$  reveal how instrumental variable inference is frustrated when the first-stage estimate of  $\beta$  is not highly significant. This occurs even with AJR's original data using controls, albeit much more strongly when problems with the mortality data are accounted for.

### C. *Special Treatment of Africa*

AJR (2006) claim their results are highly robust if Africa is excluded from the sample. This claim is addressed in Table 3, which reports estimates of  $\beta$  and  $\alpha$ , and cumulatively applies the two data robustness checks to three samples: one without Africa, one with only Africa, and one without Africa or the Neo-Europes. These results reveal several problems with this defense.

First, Africa provides a large fraction of AJR's data. Without Africa there are only 37 rates, of which only 13 are not conjectured, and the rates outside of Africa appear no less problematic than the rates in Africa. Second, AJR provide no compelling reason for why their theory should not be tested in Africa. In fact, North Africa, with a hospitable Mediterranean climate but disappointing performance, provides an important counterexample to their theory. As seen in column (2), within Africa  $\beta$  is insignificant, especially with the robustness checks.<sup>25</sup>

Third, as seen in Panels B and C of column (3), results without Africa or

---

<sup>25</sup> Note also, that even using the original data in column (1) of Panel A, excluding Africa lowers the IV estimate of  $\alpha$  to 0.61, putting it close to the OLS estimate of about 0.52, which AJR (2001) had originally rejected as being too small an estimate, motivating their IV approach.

conjectured rates, based on 13 countries, are driven by the Neo-Europes – Canada, New Zealand, and the United States. AJR’s IV model assumes that European settlers changed property-rights institutions and nothing else which affected growth, an assumption which is clearly violated by these countries, where Europeans imported their entire civilization. The Neo-Europes should be excluded from the sample as they cannot support AJR’s theory.<sup>26</sup>

### **III. Conclusion**

Given the paucity of plausible instruments in the cross-country growth literature it is regrettable that AJR’s mortality series suffers from severe measurement issues. While AJR are right to point out that regions like West Africa and the Caribbean were unhealthy for Europeans, the mortality differences between neighboring countries are largely unreliable. Much of the mortality variation is due to AJR’s questionable assignments, which often reflect transitory fluctuations or living conditions of the populations observed rather than actual permanent differences among these countries. Given the limited data sources currently available, it seems unlikely that a convincing set of settler mortality rates can be constructed. As such, cross-country growth regressions cannot disentangle the effect of settler mortality from that of other variables which may explain institutions and growth, such as geography, climate, culture, and pre-existing development. This leaves AJR’s theoretical hypotheses without a strong empirical foundation.

### **References**

Acemoglu, Daron; Johnson, Simon; and Robinson, James A. “The Colonial Origins of Comparative Development.” National Bureau of Economic Research (Cambridge,

---

<sup>26</sup> AJR (2006) also claim that their results are robust to capping mortality rates at 250, which primarily affects mortality rates in Africa. This ad hoc adjustment cancels out much of the variation within Africa which is unfavorable to their hypothesis. If AJR stand by these results, then they should explicitly retract their original estimates with rates over 250. Furthermore, rather than making piecemeal adjustments, AJR should also consider eliminating other sources of variation which appear specious, such as the mortality difference between the United States and Argentina, which become more important once this first adjustment is made.

- MA) Working Paper No. 7771, June 2000.
- \_\_\_\_\_. “The Colonial Origins of Comparative Development.” *American Economic Review*, December 2001, 91(5), pp. 1369-1401.
- \_\_\_\_\_. “A Response to Albouy’s ‘A Reexamination Based on Improved Settler Mortality Data.’” MIT mimeo, March 2005.
- \_\_\_\_\_. “Reply to the Revised (May 2006) version of David Albouy’s ‘The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data.’” MIT mimeo, September 2006.
- Anderson, T.W. and Rubin, Herman. “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations.” *Annals of Mathematical Statistics*, March 1949, 20(1), pp. 46-63.
- Bolton, Herbert E. and Marshall, Thomas M. *The Colonization of North America: 1492-1783*. New York: Hafner, 1971.
- Cantlie, Sir Neil. *A History of the Army Medical Department, Volume I*. Edinburgh and London: Churchill Livingstone, 1974.
- Curtin, Philip D. *The Image of Africa*. Madison, WI: University of Wisconsin Press, 1964.
- \_\_\_\_\_. “Epidemiology and the Slave Trade.” *Political Science Quarterly*, June 1968, 83(2), pp. 181-216.
- \_\_\_\_\_. *Death by Migration: Europe’s Encounter with the Tropical World in the 19<sup>th</sup> Century*. New York: Cambridge University Press, 1989.
- \_\_\_\_\_. *Disease and Empire: the Health of European Troops in the Conquest of Africa*. New York: Cambridge University Press, 1998.
- Curtin, Philip D.; Feierman, Steven; Thompson, Leonard and Vansina, Jan. *African History: From Earliest Times to Independence*, 2<sup>nd</sup> Ed. London: Longman, 1995.
- Dufour, Jean-Marie. “Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models.” *Econometrica*, November 1997, 65(6), pp. 1365-87.
- Earle, Carville V. “Environment, Disease, and Mortality in Early Virginia.” in Tate, Thad

- W. and David L. Ammerman. *The Chesapeake in the Seventeenth Century*. Chapel Hill: University of North Carolina Press, 1979, pp. 96-125.
- Elleman, Bruce A. *Modern Chinese Warfare, 1795-1989*. New York: Routledge, 2001.
- Gallup, John L. and Sachs, Jeffrey D. "The Economic Burden of Malaria." The Supplement to *The American Journal of Tropical Medicine & Hygiene*, January/February 2001, 64(1,2), pp. 85-96.
- Gutierrez, Hector. "La Mortalite des Eveques Latino-Americains aux XVIIe et XVIIIe Siecles." *Annales de Demographie Historique*, 1986, pp. 29-39.
- Hall, Robert E. and Jones, Charles I. "Why Do Some Countries Produce So Much More Output Per Worker than Others?" *Quarterly Journal of Economics*, February 1999, 114(1), pp. 83-116.
- Knack, Stephen and Keefer, Philip. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures." *Economics and Politics*, November 1995, 7(3), pp. 207-27.
- La Porta, Rafael; Lopez-de-Silanes, Florencio; Shleifer, Andrei and Vishny, Robert W. "Law and Finance." *Journal of Political Economy*, December 1998, 106(6), pp. 1113-55.
- Mauro, Paulo. "Corruption and Growth." *Quarterly Journal of Economics*, August 1995, 110(3), pp. 681-712.
- Moreira, Marcelo. "A General Theory of Hypothesis Testing in the Simultaneous Equations Model." Unpublished Manuscript, Harvard University, 2003.
- Moulton, Brent R. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units" *The Review of Economics and Statistics*, May 1990, 72(2), pp. 334-38.
- Parker, Philip M. *National cultures of the world: A statistical reference*, Cross-cultural Statistical Encyclopedia of the World, Vol. 4. Westport, CT: Greenwood Press, 1997.
- Reynaud, Gustave A. *Considerations sanitaires sure l'expedition de Madagascar et quelques autres expeditions coloniales, francaises et anglais*. Paris: Societe francaise d'edition d'art, 1898.

- Rodrik, Dani. "Where Did All the Growth Go?" *Journal of Economic Growth*, December 1999, 4(4), pp. 385-412.
- Royle, Charles *The Egyptian Campaigns: 1882 to 1885*. London: Hurst and Blackett, 1900.
- Sargan, J.D. "The Estimation of Economic Relationships Using Instrumental Variables." *Econometrica*, July 1958, 29(3), pp. 393-415.
- Showers, Victor. *World Facts and Figures: A Unique, Authoritative Collection of Comparative Information about Cities, Countries, and Geographic Features of the World*. New York: Wiley, 1979.
- Staiger, Douglas and Stock, James H. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, May 1997, 65(3), pp. 557-86.
- Sundbärg, G. "Tab 1: Döde efter alder och kön, 1751-1900" *Statistisk Tidskrift*, 1905, pp. 109-32.
- Tulloch, A.M. "On the Mortality of Her Majesty's Troops Serving in the Colonies during the years 1844 and 1845," *Journal of the Statistical Society of London*, September (1847), 10(3), pp. 252-259.
- Wooldridge, Jeffrey. *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, 2001.

TABLE 1: FIRST STAGE ESTIMATES  
(Dependent Variable: Expropriation Risk)

Control Variables	No Controls	Latitude Control	Without Neo-Europes	Continent Dummies	Continent Dummies & Latitude	Mean Temp and Min Rain	Percent European, 1975	Malaria in 1994
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Original Data (64 countries, 36 mortality rates)</i>								
Log mortality ( $\beta$ )	-0.61	-0.52	-0.40	-0.44	-0.35	-0.29	-0.42	-0.44
{homoscedastic s.e.}	{0.13}	{0.14}	{0.13}	{0.17}	{0.18}	{0.15}	{0.14}	{0.19}
(heteroscedastic-clustered s.e.)	(0.17)	(0.19)	(0.17)	(0.20)	(0.21)	(0.19)	(0.19)	(0.25)
<i>p</i> -value of log mortality	0.001	0.01	0.03	0.04	0.11	0.13	0.03	0.08
<i>p</i> -value of controls	-	0.17	-	0.40	0.34	0.001	0.02	0.20
<i>Panel B: Removing conjectured mortality rates and correcting Mali (28 countries and mortality rates)</i>								
Log mortality ( $\beta$ )	-0.59	-0.37	-0.26	-0.25	-0.12	-0.15	-0.21	-0.17
(heteroscedastic s.e.)	(0.24)	(0.26)	(0.21)	(0.23)	(0.27)	(0.26)	(0.24)	(0.32)
<i>p</i> -value of log mortality	0.02	0.18	0.22	0.28	0.65	0.57	0.39	0.59
<i>p</i> -value of controls	-	0.05	-	0.01	0.001	0.002	0.010	0.02
<i>Panel C: Original data, adding campaign and laborer dummies (64 countries, 36 mortality rates)</i>								
Log mortality ( $\beta$ )	-0.45	-0.39	-0.31	-0.37	-0.30	-0.12	-0.27	-0.26
(heteroscedastic-clustered s.e.)	(0.18)	(0.20)	(0.17)	(0.22)	(0.23)	(0.21)	(0.19)	(0.24)
<i>p</i> -value of log mortality	0.02	0.06	0.09	0.09	0.20	0.58	0.17	0.29
<i>p</i> -value of dummies	0.16	0.22	0.31	0.26	0.35	0.12	0.19	0.24
<i>p</i> -value of controls	-	0.27	-	0.75	0.66	0.001	0.02	0.11
<i>Panel D: Removing conjectured mortality, correcting Mali, adding campaign and laborer dummies (28 countries and mortality rates)</i>								
Log mortality ( $\beta$ )	-0.29	-0.08	-0.06	-0.16	0.01	0.07	-0.08	0.04
(heteroscedastic s.e.)	(0.25)	(0.27)	(0.22)	(0.26)	(0.29)	(0.29)	(0.23)	(0.32)
<i>p</i> -value of log mortality	0.03	0.03	0.05	0.30	0.29	0.01	0.11	0.06
<i>p</i> -value of dummies	0.03	0.04	0.05	0.32	0.31	0.01	0.11	0.06
<i>p</i> -value of controls	-	0.05	-	0.03	0.01	0.004	0.04	0.04
<i>Panel E: Removing conjectured rates, correcting Mali, adding campaign and laborer dummies, and revising with new data (34 countries and mortality rates)</i>								
Log mortality ( $\beta$ )	-0.36	-0.22	-0.10	-0.25	-0.10	0.02	-0.15	-0.14
(heteroscedastic s.e.)	(0.22)	(0.24)	(0.21)	(0.25)	(0.24)	(0.24)	(0.23)	(0.27)
<i>p</i> -value of log mortality	0.11	0.35	0.66	0.32	0.69	0.93	0.53	0.61
<i>p</i> -value of dummies	0.01	0.02	0.02	0.28	0.30	0.00	0.06	0.10
<i>p</i> -value of controls	-	0.11	-	0.14	0.15	0.001	0.04	0.03

Expropriation Risk is "Average protection against expropriation risk 1985-1995" as measured on a scale from 0 to 10, where a higher score represents greater protection, by Political Risk Services. The original Log Mortality is the logarithm of European settler mortality rates from AJR (Acemoglu, Johnson, and Robinson, 2001). Standard errors, assuming uncorrelated homoscedastic errors, are shown in braces {} in Panel A. All other standard errors and tests adjust for heteroscedasticity and clustering effects, where clusters are defined by countries sharing the same mortality rate. *p*-value of controls are probability values from standard F-tests of whether the controls are significant in the regression. *p*-value of dummies refers to an F-test of the joint significance of the campaign and laborer dummies. See Appendix Table A1 for indicators of whether a country's data is conjectured or is a rate from campaigning soldiers or laborers. "Correcting Mali" involves replacing AJR's mortality rate of 2940 with 280. "Neo-Europes" consist of Australia, Canada, New Zealand, and the United States, and are based off of three mortality rates. The three continent variables included are Africa, Asia, and Other, taken from AJR, consists of Australia, Malta, and New Zealand. Minimum monthly rainfall and mean temperature are taken from Parker (1997). Percent of European Descent in 1975 is the percent of the population of European descent in 1975 from AJR. Malaria in 1994 refers to percent of the population with endemic malaria in 1994 in Gallup and Sachs (2001) which does not contain data for Malta and the Bahamas. Revisions with new data from AJR (2005) are discussed in the Appendix and given in Table A7. See the text for more detail.

TABLE 2: INSTRUMENTAL VARIABLE ESTIMATES AND CONFIDENCE REGIONS  
(First Stage Dependent Variable: Expropriation Risk; Second Stage Dependent Variable, Log GDP per Capita, 1995, PPP basis)

Control Variables	No Controls (1)	Latitude Control (2)	Without Neo- Europes (3)	Continent Dummies (4)	Continents & Latitude (5)	Mean Temp & Min Rain (6)	Percent European, 1975 (7)	Malaria in 1994 (8)
<i>Panel A: Original Mortality (64 countries, 36 mortality rates)</i>								
Expropriation Risk ( $\alpha$ )	0.93	0.96	1.24	0.97	1.07	1.34	0.92	0.59
Wald 95% Conf. Region	[0.52, 1.34]	[0.42, 1.50]	[0.35, 2.14]	[0.25, 1.70]	[-0.01, 2.16]	[-0.18, 2.86]	[0.28, 1.56]	[0.07, 1.11]
AR "95%" Conf. Region	[0.66, 1.83]	[0.64, 2.39]	[0.73, 7.04]	[0.50, 9.02]	( $-\infty$ , -3.08] U [0.41, $+\infty$ )	( $-\infty$ , -4.34] U [0.65, $+\infty$ )	[0.51, 6.45]	( $-\infty$ , -1.62] U [0.05, $+\infty$ )
<i>Panel B: Removing conjectured mortality rates and correcting Mali (28 countries and mortality rates)</i>								
Expropriation Risk ( $\alpha$ )	0.95	0.98	1.51	1.46	2.26	2.36	1.33	1.21
Wald 95% Conf. Region	[0.42, 1.48]	[-0.09, 2.04]	[-0.59, 3.61]	[-0.81, 3.74]	[-6.76, 11.28]	[-5.20, 9.92]	[-1.33, 3.98]	[-2.41, 4.83]
AR "95%" Conf. Region	[0.63, 3.36]	( $-\infty$ , -0.71] U [0.40, $+\infty$ )	( $-\infty$ , -1.42] U [0.60, $+\infty$ )	( $-\infty$ , -0.58] U [0.45, $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , -0.36] U [0.59, $+\infty$ )	( $-\infty$ , -0.16] U [0.25, $+\infty$ )	( $-\infty$ , $+\infty$ )
<i>Panel C: Original data, adding campaign and laborer dummies (64 countries, 36 mortality rates)</i>								
Expropriation Risk ( $\alpha$ )	1.09	1.15	1.45	1.06	1.19	2.60	1.18	0.66
Wald 95% Conf. Region	[0.32, 1.87]	[0.12, 2.18]	[-0.01, 2.91]	[0.07, 2.05]	[-0.30, 2.67]	[-6.07, 11.26]	[-0.29, 2.66]	[-0.50, 1.81]
AR "95%" Conf. Region	[0.62, 5.07]	( $-\infty$ , -17.59] U [0.60, $+\infty$ )	( $-\infty$ , -8.05] U [0.69, $+\infty$ )	( $-\infty$ , -3.28] U [0.45, $+\infty$ )	( $-\infty$ , -0.67] U [0.29, $+\infty$ )	( $-\infty$ , -0.53] U [0.64, $+\infty$ )	( $-\infty$ , -1.67] U [0.44, $+\infty$ )	( $-\infty$ , $+\infty$ )
<i>Panel D: Removing conjectured mortality, correcting Mali, campaign and laborer dummies (28 countries and mortality rates)</i>								
Expropriation Risk ( $\alpha$ )	1.34	2.19	4.49	1.96	-26.33	-2.92	2.66	-2.18
Wald 95% Conf. Region	[-0.54, 3.22]	[-12.2, 16.6]	[-26.8, 35.7]	[-3.66, 7.59]	[-2234, 2181]	[-29.2, 23.3]	[-11.4, 16.7]	[-49.9, 45.5]
AR "95%" Conf. Region	( $-\infty$ , -0.44] U [0.46, $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )
<i>Panel E: Removing conjectured rates, correcting Mali, adding campaign and laborer dummies, and revising with new data (34 countries and rates)</i>								
Expropriation Risk ( $\alpha$ )	0.84	0.71	1.99	0.99	1.16	-4.74	1.11	0.35
Wald 95% Conf. Region	[0.06, 1.62]	[-0.56, 1.99]	[-5.59, 9.56]	[-0.64, 2.61]	[-4.29, 6.62]	[-118, 109]	[-1.78, 4.00]	[-1.86, 2.55]
AR "95%" Conf. Region	( $-\infty$ , -0.70] U [0.07, $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )	( $-\infty$ , $+\infty$ )

Panels present the IV estimates of *Expropriation Risk* on *Log GDP per Capita, 1995, PPP basis*, using *Log Mortality* as an instrument, and the control variables and sample selection described in Table 1. *Wald 95% Conf Region* are the standard (erroneous) IV confidence regions based on the Wald statistic. *AR "95%" Conf Region* are the confidence regions calculated from the Anderson-Rubin (1949) statistic as described in the text. Heteroscedasticity and clustering effects are corrected for all confidence regions. See text and Table 1 for more details.

TABLE 3: THE ROLE OF AFRICA  
(Dependent Variable: Expropriation Risk)

Control Variables	Without Africa (1)	Only Africa (2)	Without Africa or Neo- Europes (3)
<i>Panel A: Original data</i>			
Log mortality ( $\beta$ ) (heteroscedastic-clustered s.e.)	-1.21 (0.18)	-0.12 (0.21)	-0.83 (0.27)
<i>p</i> -value of log mortality	0.001	0.57	0.01
Expropriation Risk ( $\alpha$ )	0.61	2.00	0.77
Wald 95% Conf. Region	[0.39,0.82]	[-4.57,8.57]	[0.20,1.33]
AR "95%" Conf. Region	[0.43,0.89]	$(-\infty,+\infty)$	[0.37,2.19]
Countries	37	27	33
Mortality Rates	19	17	16
<i>Panel B: Removing conjectured mortality rates, correcting Mali</i>			
Log mortality ( $\beta$ ) (heteroscedastic s.e.)	-1.00 (0.28)	-0.03 (0.25)	-0.32 (0.23)
<i>p</i> -value of log mortality	0.004	0.90	0.21
Expropriation Risk ( $\alpha$ )	0.900	8.69	2.11
Wald 95% Conf. Region	[0.44,1.36]	[-134, 152]	[-1.86,6.07]
AR "95%" Conf. Region	[0.59,1.89]	$(-\infty,+\infty)$	$(-\infty,-3.96)$ U [0.55, $+\infty$ )
Countries and mortality rates	13	15	10
<i>Panel C: Removing conjectured mortality, correcting Mali, and adding campaign and laborer dummies</i>			
Log mortality ( $\beta$ ) (heteroscedastic s.e.)	-0.88 (0.32)	0.03 (0.27)	-0.12 (0.22)
<i>p</i> -value of log mortality	0.02	1.00	0.71
<i>p</i> -value of dummies	0.63	0.87	0.49
Expropriation Risk ( $\alpha$ )	0.92	-6.20	4.55
Wald 95% Conf. Region	[0.27,1.57]	[-115, 103]	[-21.3,30.4]
AR "95%" Conf. Region	[0.48,2.92]	$(-\infty,+\infty)$	$(-\infty,+\infty)$
Countries and mortality rates	13	15	10

See Table 1 for details. "Neo-Europes" consists of Australia, Canada, New Zealand, and the United States, and are based on three mortality rates.

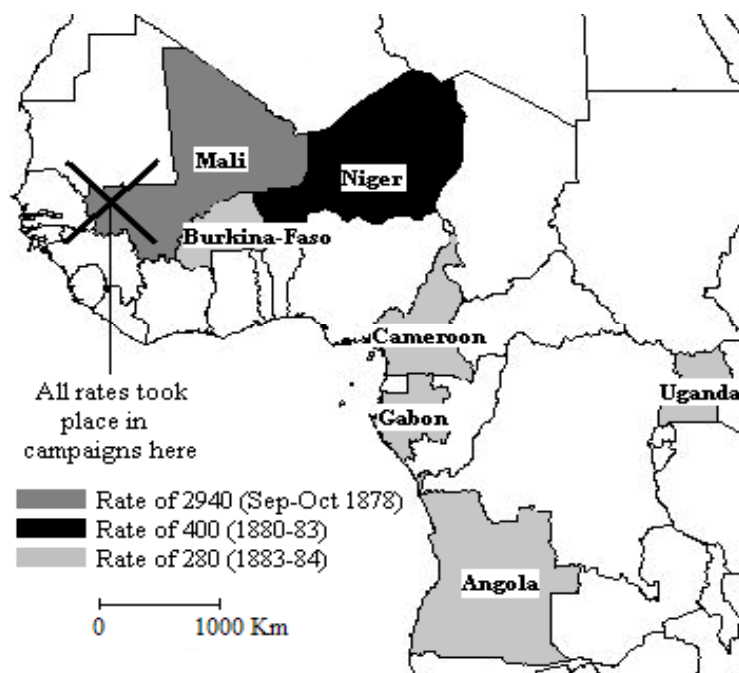


FIGURE 1: ASSIGNMENT OF MORTALITY RATES FROM MALI

APPENDIX TABLE A1: ORIGINAL MORTALITY RATES AND DATA INDICATORS

Country Name	Original Mortality	Campaign Rate	Laborer Rate	Rate From Within Country	"Benchmarked" Latin American Data
Angola	280	✓			
Argentina	68.9	✓			✓
Australia	8.55				
Burkina Faso	280	✓			
Bangladesh	71.41	✓		✓	
Bahamas	85				
Bolivia	71	✓			✓
Brazil	71	✓			✓
Canada	16.1			✓	
Chile	68.9	✓			✓
Cote d'Ivoire	668	✓			
Cameroon	280	✓			
Congo	240		✓	✓	
Colombia	71	✓			✓
Costa Rica	78.1	✓			✓
Dominican Republic	130				
Algeria	78.2	✓		✓	
Ecuador	71	✓			✓
Egypt	67.8	✓		✓	
Ethiopia	26	✓		✓	
Gabon	280	✓			
Ghana	668	✓		✓	
Guinea	483	✓			
Gambia	1470	✓		✓	
Guatemala	71	✓			✓
Guyana	32.18				
Hong Kong	14.9				
Honduras	78.1	✓			✓
Haiti	130				
Indonesia	170	✓		✓	
India	48.63			✓	
Jamaica	130			✓	
Kenya	145		✓	✓	
Sri Lanka	69.8			✓	
Morocco	78.2	✓			
Madagascar	536.04	✓		✓	
Mexico	71	✓		✓	
Mali	2940	✓		✓	
Malta	16.3			✓	
Malaysia	17.7			✓	
Niger	400	✓			
Nigeria	2004	✓		✓	
Nicaragua	163.3	✓			✓
New Zealand	8.55			✓	
Pakistan	36.99	✓			
Panama	163.3	✓			✓
Peru	71	✓			✓
Paraguay	78.1	✓			✓
Sudan	88.2	✓		✓	
Senegal	164.66			✓	
Singapore	17.7				
Sierra Leone	483	✓		✓	
El Salvador	78.1	✓			✓
Togo	668	✓			
Trinidad and Tobago	85			✓	
Tunisia	63	✓		✓	
Tanzania	145		✓		
Uganda	280	✓			
Uruguay	71	✓			✓
USA	15			✓	
Venezuela	78.1	✓			✓
Vietnam	140	✓		✓	
South Africa	15.5			✓	
Zaire	240		✓		

See the text and Appendix for further details.