**Appendix for "The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data"**

This appendix has two sections. The first, A.I, presents a detailed discussion of the mortality rates used by AJR and how they assign them to countries. It discusses when mortality rates come from soldiers on campaign, soldiers in barracks, African laborers, or were ultimately derived from bishops. In most cases, this classification is made by Curtin himself, although a few cases require further investigation. The mortality rate classification, summarized in Appendix Table A1, is used for the various robustness checks in Section II.

The sub-section on the Caribbean and Latin America discusses how AJR's benchmarking system based on the bishop data is based on thin data and can produce widely varying results, supplementing the discussion in Section I.A. The sub-sections on Africa discusses how AJR do not consistently apply either the selection rule of

- choosing the earliest available rate, stated in AJR (2001),
- or of choosing the earliest available peacetime rate, when such a rate is available, or else using the earliest campaign rate, stated in AJR (2005, 2006),

This discussion provides support for claims made in Section I.B. In fact, for a number of countries AJR simply appear to be choosing the highest rate available. Mortality rates from new sources introduced by AJR (2005) are also discussed.

The second section, A.II, considers additional robustness checks of the relationship between mortality rates and property-rights institutions. This includes:

- Separately dropping countries with conjectured mortality rates, and revising the mortality rate in Mali from 2940 to 280.
- Using different methodologies to infer Latin American mortality rates from the bishop rates in Gutierrez (1986), such as applying the rates directly, or dropping them altogether.
- Testing alternative measures of property rights institutions other than expropriation risk.
- Incorporating additional data from sources introduced in AJR (2005).
- Correcting for obvious inconsistencies in AJR's choices of mortality rates in their second data-construction step, by applying their second selection rule consistently to

their original sources. This exercise illustrates how AJR's inconsistent data choices work to strengthen the first-stage relationship between mortality and property-rights institutions.

## A.I. Region-by-Region Summary of Problems and Revisions

*Sub-Saharan Africa*: AJR's assignments of three different mortality rates from western Mali to seven different countries are difficult to explain. They appear to be due to multiple misunderstandings of colonial names contained in Curtin (1998).

- Two tables in Curtin (1998) report mortality rates that belong to Mali, one titled "Haut-Senégal-Niger" on page 85, the other has rates marked "French Soudan" on page 238. The table with "Haut-Senégal-Niger" lists yearly rates from 1880 to 1892; the table with "French Soudan" rates contains only the years from 1884 to 1888. The yearly rates reported for the five overlapping years are very similar (i.e. 282, 225, 201, 222, and 117 versus 280, 225, 200, 221, and 116). This similarity suggests that the rates refer to the same campaigns; the original source of the "French Soudan" mortality rates (Reynaud, 1898) confirms this.
- "French Soudan" is the name for Mali from 1890 to 1899 and from 1920 to 1959. "Haut-Senégal-Niger" is the name for Mali and Burkina Faso from 1904 to 1920 and for Niger from 1904 to 1911. It appears that AJR were unaware that both terms apply to Mali. The text and map in Curtin's chapter containing the "Haut-Senégal-Niger" table (pp. 74-89) indicate that these rates refer to campaigns in western Mali. Footnotes of the table include the word "Soudan" three times, with one mention of "le Soudan français."
- Although "Haut-Senegal" and "Niger" are not synonyms, AJR state (Data Appendix, p. 3) "In Haut-Senegal (Niger), in 1880-83 there was a death rate of 400 per 1000 mean strength (Curtin 1998, p. 85)," to justify assigning the earliest mortality rate of 400 (1880-83) from the "Haut-Senégal-Niger" table to Niger only.
- Incorrectly, AJR state (Data Appendix, p.3) that "Burkina Fasu [sic], Central African Federation [sic], Chad, French Congo and Mauritania were part of French Soudan." and assign to these countries the first rate in the "French Soudan" table

of 280 from 1884.[i] This rate is then assigned to Angola, Cameroon, Gabon, and Uganda under the unexplained premise that their disease environments are similar.

- The mortality rate of 2940 used for Mali comes from the two-month Logo expedition from 1874 (Curtin, 1998, p. 81), discussed in Section 1.A of the main text.

The net result is that AJR assign three different rates, all from western Mali, to three different sets of countries: a rate of 400 from 1880-3 to Niger, a rate of 280 from 1884 to the five dispersed countries of Burkina Faso, Cameroon, Gabon, Angola and Uganda, and an epidemic-based rate of 2940 from two months in 1874 to Mali.

The Mali-based rates are all classified as campaign rates. As Curtin uses the terms "campaign" and "expedition" interchangeably, the rate from the Logo expedition is a campaign rate. The other rates are primarily campaign rates according to Curtin (1998, p. 84), "The annual deaths reported for most years in the 1880s are therefore a combination of garrison rates for four to six months, plus campaign rates for six to eight months during the dry season." Furthermore, the rates of 280 and 400 are higher than the average rate of 200.24 that Curtin (1998) singles out as the "Annual Mortality of French Troops on Campaign in the French Sudan, 1883-88" in his Table 4.3, titled "Senegal and French Sudan – Barracks and Campaign." Separate mortality rates for the other six countries are not evident in Curtin's texts.

The rates applied to Congo and Kenya are clearly maximum rates and taken from African laborers. The rate of 240 for Congo, also used for Zaire, is in Curtin et al. (1995, p. 463):

> …workers had to be recruited elsewhere; and they were recruited by force from all parts of French Equatorial Africa… Many of these men came from the savanna country in Chad and Ubangi-Shari. They were therefore

[i] AJR's confusion over the "French Soudan" may be due to how the general term "Soudan" in French – as seen in older editions of *Le Petit Larousse* – refers to a large swath of land south of the Sahara from Mali (French Soudan), to Darfour in modern (Anglo-Egyptian) Sudan. However this area should have still included Mali, and excluded French Congo and most of the Central African Republic. The latter two countries were part of French Equatorial Africa, while Mali, Burkina Faso, and Niger were part of French Occidental Africa. It is also not clear why AJR assign the rate of 280 from Curtin (1998) to Gabon because it is a neighbor of Congo, which they claim was part of French Soudan, while Congo itself receives the rate of 240 from Curtin et al. (1995).

> unprepared for the diseases they encountered in the forested Mayombe highlands… They were also underfed and ill-housed. As a result, the overall death rate reached 100 per thousand per annum, and as high as 240 per annum at the peak of mortality…

The laborer rate of 145 for Kenya, also used for Tanzania, comes from this passage (p. 491):

> some of East Africa's highland peoples who lived in malaria-free areas knew that if they stayed overnight in lowland regions where mosquitos bred, they would be likely to suffer from malaria… The Europeans had the armed force to require that African men move to the lowlands… The result, in Kenya, was that workers from the highlands died at annual rates as high as one hundred and forty-five per thousand in Mombasa and nearby coastal regions on the eve of the First World War.

AJR (Data Appendix, p. 3) are not specific about which examples in Curtin (1968) they refer to when justifying the use of African coerced labor as a lower bound for the mortality of Europeans. The validity of AJR's lower bound argument seems questionable, as all of the mortality rates for blacks refer to troops, not laborers, who may have faced quite different conditions. The mortality of the Congolese laborers may have been abnormally high as they were "underfed and ill-housed," (Curtin, 1995, p. 463). In the Ethiopian campaign of 1867-68 reported in Curtin (1998, pp. 45-6), badly treated laborers died at a much higher rate than the European soldiers they accompanied. It also appears that the origin of the troops may be different than that of the laborers. Moreover, these examples compare average rates to other average rates, and thus do not justify AJR's practice of taking maximum rates from the laborers. That is not to say that Congo was a healthy place for Europeans, only that AJR lack comparable data for this country.[ii]

For Madagascar, AJR did not use the first available peacetime rate mentioned in the following passage (Curtin, 1998, p. 181) "In 1880, the peacetime garrison at the

---

[ii] They may have also overestimated mortality in Kenya, which is next to Ethiopia, with much lower mortality, and since there appear to be healthier highlands in Kenya. AJR (2005, p. 33) state that when rates from multiple regions from within a country, they prefer to take the lowest rate.

French post at Nossi-Bé had an annual death rate of about 75 per thousand." AJR (2006, p. 7) argue that the Nossi-Be rate should not be used since it comes from an island just off the coast, even though it has been formally part of Madagascar since 1896. Yet in the next sentence, Curtin (1998, p. 181) mentions a similar mortality rate on land: from 1884 to 1885, "those who went ashore for the occupation of Tamatave died at 95 per thousand." Later in the paragraph, Curtin mentions another mortality rate of "59 per thousand for the Malagasay expedition of 1884-85." Instead, AJR use a rate of 536.04 from 1895 during the Madagascar Expedition (p. 188), consisting of about 15,000 soldiers. Curtin (pp. 184-188) refers to this expedition repeatedly as a "campaign," making its coding clear. Yet, Curtin (1998, p. 191) writes that mortality rates diminished quickly after the 1895 campaign:

> From 1896 to 1903, the French were no longer fighting a single campaign, but many small campaigns. The death rate of French troops over these later years is partly a campaign death rate – not just a barracks rate – but by 1897, the *annual* mortality [of 30] had dropped to less than the monthly mortality during the campaign of 1895.

In 1897 the mortality rate was 30 per year, and by 1904, it was 7.69.

AJR choose a later campaign rate – the highest rate available – for Madagascar. AJR's reasoning in their Appendix (p. 2) appears to conflict with their stated preference for choosing the earliest peacetime rate available:

> The very high mortality rate from disease in Madagascar in 1895 is a reasonable estimate of what settlers would have expected because this was not a particularly well-organized campaign.

The mortality rate that AJR take from the Gold Coast (Ghana, which they also apply to Togo and Cote d'Ivoire) of 668 comes from 1823 to 1826 (Curtin, 1998, p. 18) and appears to be a campaign rate from the First Anglo-Asante War. According to Curtin (p.18), "these troops were not on campaign during the whole four years, but their death rate from disease can be taken to be a campaign, rather than a barracks, death rate." The yearly data (Army Medical Department, 1840, p.7) show an increase in mortality from 1823 to 1824 when the war began, as this coincided with a yellow fever epidemic, which subsided a few years later. The role of poor living conditions is clear from a passage in the *United Service Journal* (1840, p. 516): [iii]

> Here diseases of the bowels proved nearly as great a source of mortality as fever; but considering that the soldiers were fed constantly on salt provisions, and that the water was of the very worst quality this is not to be wondered at. None of the officers died from that class of diseases, though thirty-seven perished from fever; which shows that they must have been in great measure attributable to some other cause than climate.

The war also led to substantial troop buildups in the Sierra Leone Company, which included the Gambia. Referring to the mortality rate of 483 for Sierra Leone, Curtin (1998, p. 10) states that "European mortality represented in the chart was not the normal West African experience, but only typical of what could happen during a yellow fever epidemic." Because the mortality rate is averaged using troop strengths as weights, and as there was a military build-up during the epidemic period – the number of troops rose from 6 in 1824 to 571 in 1825, dropping to 9 by 1830 – the mortality rate for the entire period of 1817 to 1836 reflects the epidemic period (Feinberg, 1974). According to Curtin (1989, p. 18),

> The Sierra Leone rate of more than 400 per thousand was somewhat higher than usual, but peacetime rates of 100 to 200 per thousand had been common enough in the past and were to persist for several decades to come.

Although evidence of actual conflict is not given, the living conditions up until 1826 at the Sierra Leone camp were abysmal, with "diseases of the stomach and bowels" killing

[iii] In their Data Appendix, AJR consider using the Sierra Leone rate for Ghana, Togo, Nigeria, and Cote d'Ivoire, although this rate was impacted by the epidemic as well (see below).

many. [iv]

The deaths from Gambia come from a particularly hard hit sub-population of the Sierra Leone Command from 1825 to 1826, precisely during the worst time of the yellow fever epidemic. AJR use this period to infer a mortality rate of 1470, which they state (Data Appendix, p. 3), "is high but not extraordinary for that time and place." Since Gambia was part of the Sierra Leone Command, the Sierra Leone rate could be revised to avoid double-counting mortality in Gambia: Tulloch (Army Medical Department, 1840, p. 7) suggests a rate of 350. It seems reasonable to classify rates for both countries as campaign rates given the nearby conflict, the poor conditions, the yellow fever epidemic, and Curtin writing that the rates are higher than typical peacetime rates,. Classifying them as barracks rates has little effect on the results.[v]

For an expedition of 159 Europeans in 1841 up the Niger, which AJR use for Nigeria, Curtin (1998, p. 21) explains:

> The longest time any of the steamers spent in the river that year was just over two months, but even for this brief period the death rate was 350 per thousand. Eighty-two percent of the crews came down with malaria, and the case-fatality rate was 30 percent. The standard statistical measure of annual deaths per thousand at risk is obviously impossible to determine… The death rate per month gives a measure more nearly comparable with similar expeditions, and the overall monthly death rate in this case, based on the number of days each man was at risk on a steamer in the Niger, comes to 162 per thousand, per month. [vi].

[iv] The Army Medical Department (1840, pp. 6-9) describes these conditions, including the brackish water soldiers drank, "liable to cause slight affections in the bowels in persons unaccustomed to it." While soldiers had rations of one pound of meat and one pound of bread per day, the meat given to these soldiers was of poor quality, with little else being added to it unlike in other stations. The author states that some portion of the sickness and mortality was due to the privation of diet. All of the troops were also "commuted punishment men" of "the most degraded class of soldiers" and may have had higher mortality rates than typical soldiers.

[v] If the rates for Gambia and Sierra Leone are classified as barracks rates then the first stage estimates of $\beta$ using the campaign dummies in Table 2 become slightly more significant without controls and slightly less significant in other specifications.

[vi] AJR appear to have misread this passage as they write (Appendix, p. 3) that "In a period of several months, 82 percent of the Europeans died from malaria," when in fact only 30 percent of this 82 percent died. Case-fatality rates from malaria are high, but far from 100 percent.

In another passage, which AJR (2006) cite, Curtin (1996, p. 101) writes

> It would not be statistically legitimate to annualize a death rate of this magnitude, but if one did, it would have been more than 2,000 per thousand annual strength.

Despite this passage and the comparability issues which Curtin raises, AJR take the monthly rate of 162 and multiply it by 12 to infer an annual rate of 2004 for Nigeria. Whatever the case, this rate is from an "expedition" and is thus properly labeled a campaign rate.

*North and Northeast Africa:* As mentioned earlier, North Africa has a fairly healthy Mediterranean climate suitable for European settlement. In Sudan and Ethiopia, the hot and arid summers may not have been particularly comfortable for European soldiers, a number of whom died from sunstroke when they had inadequate shelter, but they were not decimated by tropical disease. Nevertheless, for most of these countries, AJR choose mortality rates abnormally high because of waterborne diseases that stem from campaign conditions.[vii] For these countries, AJR do not follow a consistent rule of picking either the earliest available rate or the earliest peacetime rate, but instead tend to choose the highest mortality rate available, as with Madagascar.

The rate AJR used for Algeria from 1831-38 is clearly a campaign rate. The relevant passage in Curtin (1989, p. 17) is worth repeating in its entirety:

> Climatically, the south shore of the Mediterranean was much like the north shore in Italy or southern France. On grounds of general climate, one might expect even smaller mortality differences than those reported on Table 1.1. The British posts at Gibraltar, Malta, and the Ionian islands, however, show a range of relocation costs from 7 to 65 percent [greater mortality than Britain, proportionally], which seems to indicate that the disease environment was not much different from that of northern Europe.

---

[vii] To explain the role of typhoid in affecting these mortality rates: Curtin names chapters in his (1998) book "The Typhoid Campaigns: Northeastern Africa in the 1880s," and "Typhoid and the Egyptian Garrison."

> The high Algerian figure of 288 percent [greater mortality than France, proportionally] in the 1830s was certainly the result of campaigning in the conquest period. Within a decade or so, the Algerian death rate was closer to the rates of the Mediterranean islands. Indeed, by 1909-13, Gibraltar, Malta, and Cyprus had all begun to show small relocation benefits for British troops (Table 1.2).

AJR (2005) claim they do not use rates from wartime. Yet, the Algerian campaign was not a minor expedition: Curtin calls it "a major military operation" (1989, p. 28), "with a total force of 100,000 in round numbers" (p. 5). In their Response, AJR (2005) cite a rate mentioned in passing in Tulloch (1847) that early Algerian settlers experienced a mortality of 70, which does not mention which diseases were responsible. The relatively high mortality of these settlers is not altogether surprising given that this was during the initial settlement of Algeria and during a war – the conditions facing the settlers were far from ideal, and yet they still fared better than the initial settlers of Jamestown, who died at four times this rate (Earle, 1979). Later experience would show Algeria to be quite healthy once sanitary conditions were established: the hundreds of thousands of European settlers who came to Algeria certainly attest to this (see Curtin et al., 1995, pp. 434-5).

Furthermore, AJR overlook mortality rates from more peaceful years, mentioned several times in Curtin's tables and texts, including (1989, p. 30)

> The mortality rate for the French in Algeria dropped steeply over the midcentury… from 78 per thousand in the first five years of conquest (1831-35) to 15 per thousand in the late 1860s (1863-69).

Curtin (1998, p. 152) singles out a particular Algerian barracks rate of 12.64 from 1872-76 to compare with the campaign rate of 63.3 from Tunisia in 1881. AJR use the latter for Tunisia.

AJR state that they use a peacetime number of 67.8 for Egypt although it is not truly a peacetime rate. Three rates are reported in Curtin's Table 6.2 (1998, p. 158):

- The first rate of 24.7 is from a short campaign from July to October of 1882. The death rate from enemy action, listed in the same table, was 31.1

- The second rate of 67.8, used by AJR, is labeled "Post-Campaign." It is from the period following the initial invasion, from October to December of 1882. The death rate from enemy action from during this time was 18.2, only 41 percent lower, and thus it was not a time of actual peace. The disease mortality rate is also listed in Table 6.4 (p.169) titled "Mortality in Northeast African Campaigns," which suggests Curtin saw it as a kind of campaign rate.
- The third rate of 32.64 is from the following year, 1883. There were no deaths from enemy action over this year, so this appears to be the first peacetime rate.

It appears that mortality rose temporarily at the end of 1882 because of the after-effects of dirty water drunk while campaigning and to deplorable conditions in Cairo after the British captured it:

> The outbreak [of typhoid] seems to have been caused principally by the condition of the Cairene barracks and water sources at the moment of the British occupation. Sanitary officers described their shock in finding human wastes on the stairways and in the corridors of the barracks they took over from the Egyptian army (p. 161).

This, along with the deaths from enemy action, is likely why Curtin calls 67.8 a "post-campaign" rate rather than a "barracks" rate. From 1895 to 1904 Curtin (p. 198) reports a further decline of mortality to 10.56. Of the rates available for Egypt, AJR choose the highest rate, not the first peacetime rate. Given the conditions of the soldiers, it makes sense to follow Curtin (p. 169) and label the rate from 67.8 applied to Egypt a campaign rate.

For Sudan, AJR choose the fourth available campaign rate of 88.2. It is much higher than the other rates and occurred while British forces were besieged by their Sudanese enemies. In Chapter 6, Curtin (1998) reports five mortality rates that could be used for the Soudan, all of which occurred during the Anglo-Sudan (or Mahdi) War:

- The first rate of zero comes from a campaign launched from Suakin, a coastal Sudanese city, from February to April of 1884 with 4,500 troops. According to Curtin (1998 p. 173) there were no deaths from disease: "Medically, this brief period… appears as one of unaccountable success, unless there were errors in the reporting. The force had only 127 hospital admissions and no deaths at all from disease…"

- The second rate of 30.36 (p.169) is from the Nile Expedition, which also occurred in southern Egypt, although primarily in Sudan.[viii] It spanned 15 months from March 1884 to July 1885 and consisted of 10,771 men.
- The third rate of 10.9 (p.169) comes from a second Suakin expedition from March into May of 1885, with a troop strength of 7,235.
- The fourth rate of 88.2 (p.173), used by AJR, comes from May 15, 1885, to the end of the year, when the British left a garrison in Suakin, which had an average annual strength of 465. According to Royle (1900, pp. 459-60) this garrison was "besieged" by Sudanese enemies throughout the year.[ix] 41 soldiers died: 20 of typhoid, and 15 of sunstroke (Army Medical Department, 1885). The deaths from sunstroke were blamed partly on the tents the soldiers had for accommodations, which were inadequate shelter against the summer sun and heat (p. 145).
- The fifth rate of 27.24 (p.169) is from a force of 5,873 men on the Sudan Frontier from November 1885 to January 1886.

Given the intense fighting and adverse conditions the soldiers faced during this colonial war, these rates are all effectively campaign rates. AJR's choice of the fourth rate is odd, not only because it is unusually high, but because it is based on the smallest sample. Curtin never calls this rate a peacetime rate, although he does not explicitly call it a campaign rate. It does seem reasonable to classify this rate as a campaign rate as the soldiers were under constant attack, and as the death toll was higher than in all of the other campaigns, with half of the deaths from typhoid, and most of the others from sunstroke due to inadequate shelter,[x]

---

[viii] This rate was not mentioned in previous drafts, because it is not clear in Curtin's text how much of the expedition occurred in Egypt relative to Sudan.

[ix] According to Royle (1900, pp. 459-60), "In May, 1885, when Graham's force withdrew from Souakim, General Hudson took over the command. The troops left to protect the town consisted of 930 Europeans, 2,405 Indians, and the Egyptians forming the regular garrison…. Day after day the Dervish scouts approached the forts and cavalry patrols went out and fired upon them; night after night parties of the enemy took up positions from which they fired on the town… Thus Souakim continued to be besieged, the enemy refraining from any serious attack, and devoting themselves principally to raiding the friendly tribes in adjoining territory. On 11th May, 1886, the remainder of the British and Indian troops left." It is not clear why Royle mentions 930 Europeans, when an average annual strength of 465 men over seven and-a-half months would imply 744 Europeans. Perhaps some were lost to combat or others counted as officers.

[x] In their Response (2005, p. 22), AJR state that the first available mortality rate of zero for the Sudan is one which "Curtin clearly suggests is mistaken." This seems like a strong interpretation of Curtin's text above, and it does not explain why they overlook the second and third rates.

Ethiopia's rate of 26 is unquestionably a campaign rate as it is taken from Table 2.1 in Curtin (1998, p. 44) labeled, "British Military Mortality on the Magdala Campaign, 1867-68."

*Asia and Oceania:* Curtin (1989, p. 8) reports mortality rates for three presidencies in India: 36.99 for Bombay, 71.41 for Bengal, and 48.63 for Madras for years 1829 or 1830 to 1838. AJR assign these rates to Pakistan, Bangladesh and India respectively. Yet all three presidencies overlapped with India, but for reasons not explained, AJR apply the Madras rate for India, although this does not correspond to their rule of taking the lowest available mortality rate for a given country.

Discussing the mortality rates he reports for India, Curtin (1989, p. 25) writes

> it reflects the high mortality from wartime years…. Campaigning was frequent especially in the two northern presidencies [Bengal and Bombay]. The comparative peace for Madras accounts for its lower mortality rates. After 1852, death rates fell sharply in all three presidencies, again the probable result of relative peace in these years.

Although it is not obvious what to do here, given the passage above, it makes sense to deem Bengal and Bombay campaign rates and Madras a barracks rate.

The mortality rate of 140 for Vietnam comes from a campaign in Cochin China, which the original source, Reynaud (1898, pp. 92-101, 471), describes as being deadly due to the deplorable, dirty conditions the soldiers endured and multiple cholera epidemics.[xi]

The mortality rate of 14.9 AJR use for Hong Kong belongs to the British China Field Force who fought in 1860 during Arrow's War. For at least three reasons this rate cannot be considered a campaign rate for Hong Kong: it was not annualized, most of the campaign was fought in Beijing, and as it is 19 times lower than an actual barracks rate for Hong Kong from 1842-5 of 285 (Tulloch, 1847, p. 254).

The rate of 17.7 used by AJR for Malaysia and Singapore is from a small sample

[xi] AJR chose this rate over a lower rate of 60 for Tonkin from 1864 (Curtin, 1998, p. 238), although it is not clear why. The Cochin China happens only 3 years before the Tonkin rate, despite AJR's stated preference for the lowest mortality rate reported within a country.

in Penang, Malaysia. Curtin (1989, p. 17) writes:

> the Straights Settlements of the 1830s enjoyed a brief reputation for superior healthfulness – for European troops, curiously enough, but not for those recruited in British India. The death rate of European troops in Penang during 1929-38 was only 17.7 per thousand, although the mean strength of this force was too low to be significant.

In their Response, AJR (2005) find that this mortality rate is repeated in the Statistical Society of London (1841), which also gives a combined mortality rate of 20.0 for a much larger group of soldiers in Penang, Malacca, and Singapore combined (p. 146).

For Indonesia, AJR use a rate of 170 from 1819-1829 in the Dutch East Indies, during a period of conflict which included the Java War. Curtin (1989, p. 18) states:

> The very high mortality rate for the Dutch East Indies (Indonesia) was unrepresentative for the same reason the Algerian figure was. These years included those of the Java War, with tough campaigns, high casualties from combat and high disease rates.

AJR do not use a rate from a more peaceful era of 39.15, from 1859 to 1914 (p.82), which is much closer to the rate for Malaysia. The rate AJR use for Indonesia is clearly a campaign rate.

In their response, AJR (2005) report a new barracks mortality rate of 14.1 for Australia (Balfour, 1845, p. 195). It is higher than the rate of 8.55 previously assigned from New Zealand.

*The Caribbean and Latin America:* AJR's mortality rates of 130 for Jamaica and 85 for the Windward Leeward Command (including Barbados, Guiana, and Trinidad and Tobago) from 1817 to 1836 appear to be barracks rates, although soldiers were suffering from the residual effects of war (Curtin 1989, pp. 25-8):

> As expected from campaign conditions, the death rates of all the islands had been high in wartime, but these levels continued into the immediate

> peacetime (Figure 1.2). The then-current military policy of keeping troops in the islands for a dozen years or more may be at fault to the extent that these men were veterans of the heavy fighting and the disease that went with it. By the 1830s and 1840s, however, death rates were more stable at a lower level.

AJR apply the Jamaican rate of 130 to Haiti and the Dominican Republic, and the Windward Leeward Command rate of 85 to the Bahamas. This latter choice is questionable given the Bahamas closer proximity to Jamaica, although the Bahamas does lie lower than Jamaica, which presumably makes it unhealthier. AJR claim (Data Appendix, p. 4) that "information from Gutierrez 1986 indicates that these were similar disease environments," although nowhere in Gutierrez (1986) do I see corroboration for this claim. The inaccuracy of AJR's extrapolation made plain by AJR's (2005, p. 33) own finding of a rate from the Bahamas of 189 (Tulloch, 1838b, p. 229), over twice the extrapolated rate of 85.

AJR's (2005) Response also mentions that Tulloch (1838a, p. 32) reports disaggregated mortality rates for Trinidad and Tobago (106.3) and Guiana (84). Trinidad has a lower rate (106.3) than Tobago (152.8). The rate AJR has applied to Guyana of 32.18 is based on French Guyana. Tulloch (1838b, p. 231) reports that a small contingent in Honduras suffered a mortality rate of 95.2; a case where AJR's benchmarked rate of 78.1 was close relative to other cases.

Moving forward, it is worth considering how AJR construct their mortality rates using Latin American bishops. Understanding this construction makes it possible to understand how many possible benchmarking systems similar to AJR's could be used to infer mortality rates; these systems may produce widely differing rates, undermining the practice of benchmarking.

Gutierrez defines low, medium, and high temperature regions as areas with mean temperatures of less than 20°C, 20°C to 24.9°C, and 25°C or greater, respectively, and assigns cities to these regions according to temperatures in Showers (1979). In the text of his article, Gutierrez lists a number of cities in low and high temperature regions, although not in medium temperature regions. These cities are shown in Figure A1, along with AJR's assignment of mortality rates to different countries based on the bishops. Note that several countries, including Uruguay, Paraguay, Venezuela, Costa Rica, Honduras, and El Salvador do not contain cities that Gutierrez mentions. Using additional cities listed in Showers (1979) the following countries have cities in multiple regions:

- Low and medium: Bolivia, Ecuador
- Medium and high: El Salvador, Honduras
- Low, medium, and high: Brazil, Colombia, Mexico, Venezuela

Basing the classification on capital cities would produce a similar classification to AJR's, except for Brazil which would change from low to medium. [xii]

No actual evidence is given that this classification by temperature creates an accurate classification of disease environments, suitable for benchmarking. There appear to be cases where the classification breaks down. For example, the one city listed in the Bahamas, Nassau, has a mean temperature of 24.8°C, placing the Bahamas in the medium mortality region. Despite being in a medium mortality region, the Bahamas has a higher observed mortality, 189, than the Windwards and Leewards or Jamaica, countries in a high mortality region, with rates of 85 and 130. This casts doubt on AJR's classification system and is inconsistent with AJR's assertion that the Bahamas and the Windward Leeward Command share "similar disease environments."

The bishop mortality rates are benchmarked using a rate from Mexico. However, there are two mortality rates listed for Mexico on facing tables in Curtin (1998, pp. 238-9): a rate of 53 from Reynaud (1898) and a rate of 71 from an anonymous article in the *British Medical Journal* (1898). AJR take the latter, despite their stated preference for lower rates. Although there is little explanation of these rates or why they differ, it is clear from the table that they are from a campaign: in fact it was a major campaign with over 30,000 troops. The French faced heavy resistance and lost the Battle of Puebla, which is celebrated to this day by Mexicans on Cinco de Mayo, a national holiday. Naturally, all of the mortality rates benchmarked to this rate should be labeled as from a campaign.

Furthermore, the rate of 71 used by AJR is not annualized. The original source (Reynaud 1898, 113-121 and pp. 471-2) shows that the lower rate of 53 applies to the

---

[xii] More discrepancies would occur if one were to redo their classification based on the mean temperature variable from Parker (1997) that AJR use as a control variable in their regressions: Brazil would change from low to high, and Peru and Uruguay from low to medium.

army, while the rate of 71 applies to the army, navy, and marines combined.[xiii] The mortality rate for the navy and marines, 82 was higher as they were staying on the unhealthy Mexican coast, while the army went into the highlands. Using the mortality of the navy and marines to benchmark the high mortality region results in mortality rates 50 percent lower for Latin America; using the army rate to benchmark the low mortality region results in rates 25 percent lower.[xiv]

AJR claim their results would be "virtually the same" (Data Appendix, p. 4) if they used the rate of 130 from Jamaica as their benchmark for the high mortality region, although this would lower mortality rates of Latin American countries by 20 percent.[xv] Other benchmarking systems from the Caribbean produce widely varying results. For instance, the city of Cayenne, French Guiana, is mentioned by Gutierrez as being in a high mortality zone, giving it a bishop rate of 32.8. Curtin reports a very similar soldier mortality rate of 32.18 for French Guiana – which AJR assigned to (British) Guyana. As these rates are so close, AJR's benchmarking assumptions could be used to justify using the bishop mortality rates directly, instead of multiplying them by a ratio of 4.25. Similar methods would also lower the rates for Latin America: benchmarking high mortality areas with the rate of 85 from the Windward Leeward Command (Curtin, 1989, Table 1.5), or the rate of 84 from Guiana lowers Latin American mortality by almost 50 percent. On the other hand, using the Bahamas rate of 189 as a medium mortality benchmark results in mortality rates 141 percent higher.

To cross-validate the accuracy of their Latin American numbers, AJR benchmark their rates using naval station data for 1825-45 from Curtin (1964. p. 486) and the mortality rate of 483 from Sierra Leone. From this data they take the ratio of the mortality rate of "South American Stations," which is in fact quite low at 7.7, to the mortality rate from the anti-slavery blockade off of the African coast, 54.4, which may or

---

[xiii] Using the text I was able to estimate the number of troop-years (34,319) that French troops were at risk and the number of deaths in this population (2,095) to calculate an overall annual mortality rate of 61.

[xiv] American troops during the Mexican War experienced a mortality rate of 100 (Adams, 1952, p. 194). Scott's 1847 campaign during this war took a similar route from Veracruz to Mexico City, although much campaigning occurred in what is now the United States and just south of the Mexican border.

[xv] According to AJR (Data Appendix, p. 4) "Running our regressions using data that takes Jamaica as the base for applying the Gutierrez's ratios actually strengthens our results." This is only true for the IV estimates, the opposite is true for the first stage. Lowering the mortality rates for Latin American lowers the first-stage estimate but increases the second-stage IV estimate. This makes sense using the indirect least squares formula ($\alpha = \pi/\beta$) if the reduced form coefficient $\pi$ changes by less than the first stage $\beta$.

may not apply to Sierra Leone, to get 0.142. Multiplying this ratio by Sierra Leone's rate of 483, AJR impute a solider mortality rate of 68.9 for South America, which they apply directly to Argentina and Chile in their data instead of 71, without mentioning their justification for why this only should apply to these two countries. Still, these rates imply that Argentina and Chile, largely temperate countries, are over four times more deadly than the United States – an incredible claim. Then, AJR claim that their benchmarking is robust since 68.9 is close to the imputation of 71 for the low mortality region (including Argentina and Chile) using bishop data.

AJR's cross-validation method appears to have been chosen quite arbitrarily: using the same methodology with the other naval station data in Curtin (1968), it is possible to calculate a number of other possible benchmarked rates for the low mortality region. Seven alternative examples are shown in Table A2. Another indication that the benchmarking system is flawed is apparent when comparing the naval station rates to the soldier rates in the benchmarked countries. If benchmarking is appropriate, then the ratios between naval station rates should be similar to ratios between corresponding soldier rates (e.g. comparing Malta and Jamaica 16.3/130 should be close to 9.3/18.1). As these ratios are often very different, often not obeying the same ordinal rankings, benchmarking can produce widely varying and inconsistent imputed mortality rates.

Despite its obvious relevance, AJR do not consider Gutierrez's own comparison of European and Latin American bishop mortality, based on remaining life-expectancy at age 40. In the low, medium, and high temperature regions, Gutierrez finds life-expectancies of 22.1, 21.2, and 17.6 years for Latin-American bishops. In the seventeenth century, the Latin American bishops had an overall life expectancy of 19.8 years, French bishops 27.4 years; in the eighteenth century, the corresponding figures are 20.3 and 29.0. These rates reveal that mortality was approximately 40 percent higher in Latin America than it was in Europe. This is nowhere near the over 300 percent difference suggested by AJR's rates. Furthermore, bishops in Latin America originally born in Europe had a life-expectancy of 20.5 years, while those born in Latin America had an expectancy of 19.4 years, implying that these differences do not appear to be due to acquired immunities.

Overall, it appears that the benchmarking systems proposed by AJR fail to provide stable or reliable mortality estimates for Latin America. AJR's choices result in imputed Latin American mortality rates that appear to be above average among possible imputations, and most comparable to campaign rates.

*North America:* The mortality rates AJR use for the United States and Canada of 15 and 16.1 are clearly barracks rates. These peacetime rates are much lower than rates seen during the Civil War, where white soldiers in the Union army had mortality rates from disease of 53.4 and black soldiers had mortality rates of 143.4 (Adams, 1952. p. 239). Much like the campaigns in Mexico and North Africa, typhoid ("continued fever") and diarrhea-dysentery accounted for about half of the Civil War deaths, with malaria playing a minor role, as the soldiers faced campaign-type conditions. Although the Civil War was a major war, it is not clear that the actual campaigns suffered from worse conditions than European soldiers on campaign in colonial countries, particularly as Union soldiers typically had access to medical services, hospitals, and fresh food (see Adams, 1952). The Civil War rates are remarkably close to the mortality rate of 61 for French soldiers campaigning in Mexico in overlapping years, 1862 to 1863, mentioned above.[xvi]

## A.II. Additional Checks

Table A3 presents estimates showing the effect of dropping all countries with conjectured rates separately from the effect of changing the mortality rate for Mali from 2940 to 280. Panel A presents the results using the original data, while Panel B drops the 36 conjectured mortality rates, but retains the rate of 2940 for Mali. Without the conjectured rates, log mortality is not significant at the 5 percent level in all of the specifications with controls. Meanwhile, the controls become more significant, as in Panel B of Table 1,

[xvi] In the Revolutionary War British soldiers had a mortality rate in the neighborhood of 26 (Cantlie, 1974, p. 156), although the British soldiers spent much of their time in barracks in New York, Boston, and Halifax. Hessian mercenaries died at a rate of 34. American soldiers probably died at a rate of approximately 52 (close to the Civil War rate), as they were apparently twice as sickly as the British, with sickness rates of 200 as opposed to 100 for the British, probably due to differences in campaign conditions.

It is worth noting that even within the United States, AJR choose a low mortality rate. Instead of choosing 15 for the northern United States, they could have chosen a mortality rate of 34 from the southern United States (both rates are shown in Curtin, 1989, p. 7). Mitchener and McLean (2003) collected soldier mortality rates for states from 1829 to 1838. The directly measured rates (some are inferred) reveal that mortality was not only higher in the South, but also in what were then less settled frontier states such as California (31) and Missouri (41).

repeated in Panel E. In Panel C, the original data is retained, and the only change made is for Mali, whose rate is dropped to 280 - changing this rate alone causes the significance of the instrument to fall noticeably. The other countries with rates based on Mali – Angola, Burkina Faso, Cameroon, Gabon, Niger, and Uganda – are also dropped in Panel D, causing a further drop in significance. Even if all of AJR's other conjectures correct, adjusting only for the misuse of rates from Mali is enough to make $\beta$ insignificant in most specifications.

Table A4 investigates the robustness of the results to how the bishop mortality are used. Panels B and C examine how the first stage results are affected if countries in Latin America are simply dropped – along with the other conjectured rates – or assigned un-benchmarked bishop mortality rates directly. Given that AJR's benchmarking system is unreliable and produces rates in the upper range of possible benchmarking systems, it is worth considering how results would be affected if the mortality rates of bishops were simply applied directly without benchmarking, which may be justified using the benchmark from soldiers in French Guyana, who had mortality similar to bishops in the high-temperature region. When the Latin American data are simply dropped in Panel B, the results are only mildly affected. If instead the mortality rates are simply lowered to the actual bishop mortality rates, as in Panel C, the estimates of $\beta$ are noticeably lowered. Even the specification without Africa and no controls remains on the cusp of significance at the 5 percent level. Dropping the Latin American data completely, as in Panel B of Table 1, is more favorable to AJR's hypothesis than using the original bishop mortality rates. Thus dropping conjectured Latin American rates is a conservative way of dealing with the data problems they raise.

An alternative strategy of dealing with bishops is to use a dummy variable to indicate whether a mortality rate is derived from benchmarking, as shown in Panel D. This dummy does not have a large effect on the results. However, when used in conjunction with campaign rate and laborer dummies, mortality becomes insignificant in most specifications.

Appendix Tables A5 and A6 check the robustness of the first stage using alternate measures of institutions – Constraint on Executive in 1990 and Law and Order Tradition in 1995 – applying the two checks in the main text simultaneously. These results show that mortality is still not a robust predictor of either of these measures, although it does a slightly better job with the Law and Order Tradition measure, than with the Constraint on Executive measure, where the correlation changes direction in most of the specifications.

Appendix Table A7 presents possible revisions for the mortality rates. The first set, shown in Panel A, uses sources from AJR's (2005) Response to create geographically more precise rates where previous rates were lacking, used in Panel E of Tables 1 and 2. In six cases (Australia, Bahamas, Guyana, Honduras, Hong Kong, and Singapore), the sources provide rates for countries with previously conjectured rates. In two cases (Trinidad and Sierra Leone), the additional data are more geographically disaggregated than the original data, allowing for better geographically targeted rates.

These revised data are used in Appendix Table A8, which applies the same checks as in Table 1. None of the basic conclusions change: with either check mortality is insignificant in almost all of the specifications with controls; with both checks, it is insignificant without controls.

The second panel in Table A7 presents alternative mortality rates to replace the rates for countries in Africa where AJR's choices are clearly inconsistent with their own stated rules. While the alternative rates may be imperfect, in the case of Egypt and Madagascar, they are based on what appear to be the earliest available peacetime rates. In the case of Sudan, where no peacetime rates are available, the idiosyncratically high rate of 88.2 from the besieged garrison is replaced by that of the Nile Expedition rate of 30.8, which is from an expedition with over 10,000 men and which lasted over a year.

To demonstrate the assertion that AJR's inconsistent use of campaign rates is favorable to their hypothesis, Table A9 presents results correcting for these inconsistencies. Simply dropping the countries with inconsistencies in Panel B does not greatly affect the results: AJR's inconsistent rates were chosen in a manner which produces results similar to the other data. However, when the rates for the countries are instead revised in Panel C, the significance of log mortality falls below 5 percent in a number of cases. Thus, had AJR chosen their data consistently, these countries would have provided damaging counterexamples to their hypothesis. Correcting for the inconsistencies lowers the significance of log mortality further in Panel D, where the conjectured rates are dropped and Mali is corrected; in Panel E, which adds controls for campaigning soldiers and African laborers; or Panel F, which does all of these things. The results are still weak in Panel G which, on top of revising the inconsistencies and applying the two robustness checks, incorporates the revisions using new data.

**Additional References for the Appendix**

Adams, George Worthington. *Doctors in Blue: The Medical History of the Union Army in the Civil War*. New York: Henry Schuman, 1952.

*Army Medical Department. *Statistical Report on the Sickness, Mortality and Invaliding among the Troops in Western Africa, St. Helena, the Cape of Good Hope, and the Mauritius*. London, 1840.

Army Medical Department. *Statistical, Sanitary and Medical Reports of the Army Medical Department*. London, 1885.

*Balfour, T. Graham. "Statistical Data for Forming Troops and Maintaining them in Health in Different Climates and Localities." *Journal of the Statistical Society of London*, September, 1845, *8*(3), pp. 193-209.

Feinberg, H. M. "New Data on European Mortality in West Africa: The Dutch on the Gold Coast, 1719-1760." *The Journal of African History*, 1974, *15*(3), pp. 357-371.

Mitchener, Kris James and Ian W. McLean. "The Productivity of US States since 1880." *Journal of Economic Growth*, *8*, pp. 73-114.

*Statistical Society of London. "Second Report of a Committee of the Statistical Society of London, appointed to collect and enquire into Vital Statistics, upon the Sickness and Mortality among the European and Native Troops serving in the Madras Presidency," *Journal of the Statistical Society of London*, July 1841, *4*(2), pp. 137-155.

*Tulloch, A.M. "On the Sickness and Mortality Among the Troops in the West Indies [part I]." *Journal of the Statistical Society of London*, July 1838(a), *1*(3), pp. 129-142.

*______. "On the Sickness and Mortality Among the Troops in the West Indies [part II]." *Journal of the Statistical Society of London*, August 1838(b), *1*(3), pp. 216-230.

"Western Africa and its effects on the Health of the Troops." *The United Service Journal and Naval and Military Magazine*, 1840, Part II , pp. 509-519.

* Source brought to attention in AJR's (2005) Response, along with Barfour (1845), Cantlie (1974), and Tulloch (1847) mentioned in the main text.

APPENDIX TABLE A2: POSSIBLE RATES USING AJR'S BENCHMARKING SYSTEM WITH NAVAL STATION DATA

| Naval Station | Naval Station Rate | Latin American Station Ratio | Benchmarked Soldier Country | Benchmarked Soldier Rate | Inferred Rate for Low Mortality Region | Scaling Factor for Bishop Rates |
|---|---|---|---|---|---|---|
| Formula | Rate 1 | = 7.7/ Rate 1 | | Rate 2 | = Ratio x Rate 2 | = Inferred Rate / 16.7 |
| | (1) | (2) | | (3) | (4) | (5) |
| African Station | 54.4 | 0.142 | Sierra Leone | 483 | 68.9 | 4.12 |
| African Station | 54.4 | 0.142 | Gold Coast | 668 | 94.9 | 5.68 |
| Mediterranean Station | 9.3 | 0.828 | Gibraltar | 21.4 | 17.7 | 1.06 |
| Mediterranean Station | 9.3 | 0.828 | Malta | 16.3 | 13.5 | 0.81 |
| Home Station | 9.8 | 0.786 | England | 15.3 | 12 | 0.72 |
| East Indian Station | 15.1 | 0.51 | India | 48.63 | 24.8 | 1.49 |
| West Indian Station | 18.1 | 0.425 | Trinidad & Tobago | 85 | 36.2 | 2.17 |
| West Indian Station | 18.1 | 0.425 | Jamaica | 130 | 55.3 | 3.31 |

Naval station rates are taken from Curtin (1964, p. 486). Benchmarked soldier rates from AJR (2001), except for Gibraltar and England, which are from Curtin (1989, pp. 7-8), Table 1.1., used by AJR in the first step of their data construction. The first line, combining the "African Station" rate with the "Sierra Leone" rate gives the calculation used by AJR. Other lines use the same formula using different stations and countries. "Latin American Station Ratio" gives the ratio of the mortality rate in the "Latin American Station" of 7.7 to the "Naval Station Rate" listed in column 1. Since the "Latin American Station" presumably refers to a low mortality region in Gutierrez (1986), the "Inferred Rate for Low Mortality Region" is computed by taking this ratio and multiplying it by the "Benchmarked Soldier Rate," of which there are often several possible alternatives. The "Scaling Factor for Bishop Rates" is the ratio of inferred rate to the actual bishop mortality rate in the low-temperature region of 16.7.

APPENDIX TABLE A3: REMOVING CONJECTURED RATES AND CORRECTING MALI RATE SEPARATELY
(Dependent Variable: Expropriation Risk)

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents Dummies & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) |
|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (64 countries and 36 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.61 | -0.52 | -0.40 | -0.44 | -0.35 | -0.29 | -0.42 | -0.44 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.19) | (0.19) | (0.25) |
| *p* -value of log mortality | 0.001 | 0.01 | 0.03 | 0.04 | 0.11 | 0.13 | 0.03 | 0.08 |
| *p* -value of controls | - | 0.17 | - | 0.40 | 0.34 | 0.001 | 0.02 | 0.20 |
| *Panel B: Removing conjectured mortality rates only (28 countries and mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.59 | -0.42 | -0.32 | -0.31 | -0.22 | -0.21 | -0.29 | -0.28 |
| (heteroscedastic-clustered s.e.) | (0.19) | (0.22) | (0.19) | (0.20) | (0.23) | (0.23) | (0.21) | (0.28) |
| *p* -value of log mortality | 0.01 | 0.07 | 0.10 | 0.13 | 0.35 | 0.36 | 0.19 | 0.31 |
| *p* -value of controls | - | 0.05 | - | 0.01 | 0.002 | 0.001 | 0.02 | 0.06 |
| *Panel C: Original data, correcting Mali only (64 countries, 36 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.62 | -0.51 | -0.37 | -0.40 | -0.29 | -0.26 | -0.40 | -0.40 |
| (heteroscedastic-clustered s.e.) | (0.20) | (0.22) | (0.19) | (0.23) | (0.24) | (0.21) | (0.22) | (0.28) |
| *p* -value of log mortality | 0.004 | 0.03 | 0.06 | 0.09 | 0.23 | 0.22 | 0.07 | 0.16 |
| *p* -value of controls | - | 0.20 | - | 0.34 | 0.29 | 0.001 | 0.02 | 0.13 |
| *Panel D: Original data, correcting Mali and removing countries with Mali -based conjectured rates (58 countries, 34 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.59 | -0.47 | -0.33 | -0.39 | -0.27 | -0.22 | -0.38 | -0.39 |
| (heteroscedastic-clustered s.e.) | (0.21) | (0.24) | (0.20) | (0.24) | (0.25) | (0.21) | (0.23) | (0.28) |
| *p* -value of log mortality | 0.01 | 0.05 | 0.11 | 0.11 | 0.29 | 0.31 | 0.10 | 0.17 |
| *p* -value of controls | - | 0.15 | - | 0.36 | 0.27 | 0.000 | 0.022 | 0.18 |
| *Panel E: Removing conjectured mortality rates and correcting Mali, same as Panel B in Table 1 (28 countries and mortality rates)* | | | | | | | | |
| 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) |
| *p* -value of log mortality | 0.02 | 0.18 | 0.22 | 0.28 | 0.65 | 0.57 | 0.39 | 0.59 |
| *p* -value of controls | - | 0.05 | - | 0.01 | 0.001 | 0.002 | 0.010 | 0.02 |

See text and Table 1 for more details.

APPENDIX TABLE A4: DATA ROBUSTNESS CHECKS WITH ALTERNATIVE TREATMENTS OF LATIN AMERICAN RATES

(Dependent Variable: Expropriation Risk)

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) | Without Africa (9) |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (64 countries, 36 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.61 | -0.52 | -0.40 | -0.44 | -0.35 | -0.29 | -0.42 | -0.44 | -1.21 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.19) | (0.19) | (0.25) | (0.18) |
| *p*-value of log mortality | 0.00 | 0.01 | 0.03 | 0.04 | 0.11 | 0.13 | 0.03 | 0.08 | 0.000 |
| *p*-value of controls | - | 0.17 | - | 0.40 | 0.34 | 0.001 | 0.02 | 0.20 | - |
| *Panel B: Eliminating countries with rates "benchmarked" to bishops (48 countries, 33 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.62 | -0.51 | -0.39 | -0.40 | -0.32 | -0.23 | -0.37 | -0.23 | -1.13 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.20) | (0.17) | (0.20) | (0.23) | (0.19) | (0.20) | (0.23) | (0.23) |
| *p*-value of log mortality | 0.001 | 0.02 | 0.03 | 0.06 | 0.17 | 0.25 | 0.07 | 0.33 | 0.000 |
| *p*-value of controls | - | 0.23 | - | 0.27 | 0.36 | 0.001 | 0.01 | 0.02 | - |
| *Panel C: Using bishop data directly (64 countries, 35 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.44 | -0.34 | -0.28 | -0.30 | -0.23 | -0.11 | -0.21 | -0.13 | -0.66 |
| (heteroscedastic-clustered s.e.) | (0.15) | (0.15) | (0.12) | (0.16) | (0.18) | (0.15) | (0.17) | (0.21) | (0.32) |
| *p*-value of log mortality | 0.01 | 0.03 | 0.03 | 0.08 | 0.20 | 0.45 | 0.22 | 0.54 | 0.050 |
| *p*-value of controls | - | 0.07 | - | 0.08 | 0.04 | 0.000 | 0.03 | 0.02 | - |
| *Panel D: Using a dummy variable for rates "benchmarked" to bishops (64 countries, 36 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.63 | -0.54 | -0.41 | -0.42 | -0.35 | -0.28 | -0.40 | -0.27 | -1.14 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.18) | (0.18) | (0.23) | (0.21) |
| *p*-value of log mortality | 0.001 | 0.01 | 0.03 | 0.04 | 0.12 | 0.13 | 0.04 | 0.24 | 0.000 |
| *p*-value of bishop dummy | 0.11 | 0.16 | 0.64 | 0.25 | 0.31 | 0.01 | 0.002 | 0.003 | 0.37 |
| *p*-value of controls | - | 0.21 | - | 0.29 | 0.36 | 0.001 | 0.003 | 0.01 | - |
| *Panel E: Using separate dummy variables for campaign, laborer, and bishop rates (64 countries, 36 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.48 | -0.41 | -0.31 | -0.39 | -0.32 | -0.17 | -0.33 | -0.22 | -1.11 |
| (heteroscedastic-clustered s.e.) | (0.20) | (0.23) | (0.19) | (0.22) | (0.24) | (0.22) | (0.19) | (0.25) | (0.23) |
| *p*-value of log mortality | 0.03 | 0.08 | 0.12 | 0.09 | 0.20 | 0.44 | 0.09 | 0.37 | 0.000 |
| *p*-value of all dummies | 0.17 | 0.26 | 0.49 | 0.45 | 0.56 | 0.014 | 0.01 | 0.02 | 0.62 |
| *p*-value of controls | - | 0.32 | - | 0.63 | 0.69 | 0.001 | 0.01 | 0.02 | - |
| *Panel F: Eliminating countries with rates "benchmarked" to bishops or conjectured from Mali, and correcting Mali (42 countries, 31 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.59 | -0.46 | -0.33 | -0.34 | -0.22 | -0.14 | -0.31 | -0.14 | -1.13 |
| (heteroscedastic-clustered s.e.) | (0.21) | (0.25) | (0.21) | (0.23) | (0.26) | (0.22) | (0.23) | (0.26) | (0.23) |
| *p*-value of log mortality | 0.008 | 0.07 | 0.12 | 0.15 | 0.42 | 0.54 | 0.19 | 0.59 | 0.000 |
| *p*-value of controls | - | 0.21 | - | 0.25 | 0.30 | 0.000 | 0.01 | 0.00 | - |

Countries "benchmarked" to bishops given in Appendix Table A1. Results in Panel C derived by replacing mortality rates of 71 and 68.9 with 16.7, 78.1 with 17.5, and 163.3 with 32.8. See text and Table 1 for more detail.

APPENDIX TABLE A5: FIRST STAGE ESTIMATES USING CONSTRAINT ON EXECUTIVE IN 1990 AS AN INSTITUTIONS MEASURE

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) | Without Africa (9) |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (34 countries, 60 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.97 | -0.93 | -0.86 | -0.47 | -0.46 | -0.62 | -0.57 | -0.25 | -0.76 |
| (heteroscedastic-clustered s.e.) | (0.21) | (0.26) | (0.27) | (0.29) | (0.32) | (0.29) | (0.25) | (0.28) | (0.25) |
| $p$-value of log mortality | 0.000 | 0.001 | 0.003 | 0.12 | 0.16 | 0.04 | 0.03 | 0.38 | 0.01 |
| $p$-value of controls | - | 0.70 | - | 0.000 | 0.000 | 0.01 | 0.000 | 0.000 | - |
| *Panel B: Removing conjectured mortality, correcting Mali, campaign and laborer dummies (28 countries and mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.06 | 0.09 | 0.02 | 0.02 | 0.14 | 0.21 | 0.05 | 0.30 | -0.22 |
| (heteroscedastic s.e.) | (0.38) | (0.44) | (0.43) | (0.41) | (0.47) | (0.45) | (0.41) | (0.39) | (0.24) |
| $p$-value of log mortality | 0.87 | 0.84 | 0.96 | 0.95 | 0.76 | 0.65 | 0.90 | 0.45 | 0.38 |
| $p$-value of dummies | 0.000 | 0.000 | 0.000 | 0.002 | 0.001 | 0.002 | 0.001 | 0.000 | 0.000 |
| $p$-value of controls | - | 0.13 | - | 0.39 | 0.47 | 0.204 | 0.16 | 0.06 | - |

Constraint on Executive in 1990 is on a scale from 1 to 7 with a higher score indicating more constraints, taken from the Polity III data set. Sample does not include the Bahamas, Hong Kong, Malta and Sierra Leone. See text and other tables for additional information.

APPENDIX TABLE A6: FIRST STAGE ESTIMATES USING LAW AND ORDER TRADITION IN 1995 ASAN INSTITUTIONS MEASURE

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) | Without Africa (9) |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (63 countries, 36 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.54 | -0.42 | -0.41 | -0.39 | -0.29 | -0.37 | -0.41 | -0.41 | -1.01 |
| (heteroscedastic-clustered s.e.) | (0.13) | (0.15) | (0.14) | (0.14) | (0.15) | (0.16) | (0.14) | (0.18) | (0.15) |
| $p$ -value of log mortality | 0.000 | 0.01 | 0.01 | 0.01 | 0.06 | 0.03 | 0.01 | 0.03 | 0.000 |
| $p$ -value of controls | - | 0.04 | - | 0.01 | 0.002 | 0.03 | 0.10 | 0.41 | - |
| *Panel B: Removing conjectured mortality, correcting Mali, campaign and laborer dummies (28 countries and mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.42 | -0.25 | -0.27 | -0.28 | -0.11 | -0.19 | -0.20 | -0.31 | -1.04 |
| (heteroscedastic s.e.) | (0.22) | (0.24) | (0.21) | (0.22) | (0.27) | (0.24) | (0.20) | (0.28) | (0.20) |
| $p$ -value of log mortality | 0.07 | 0.32 | 0.21 | 0.22 | 0.69 | 0.43 | 0.31 | 0.28 | 0.00 |
| $p$ -value of dummies | 0.83 | 0.42 | 0.96 | 0.84 | 0.38 | 0.93 | 0.81 | 0.84 | 0.73 |
| $p$ -value of controls | - | 0.04 | - | 0.01 | 0.01 | 0.03 | 0.000 | 0.63 | - |

Law and Order Tradition in 1995 is measured on a scale from 0 to 6, with a higher score meaning more law and order, from Political Risk Services. Sample does not include El Salvador. See text and other tables for additional information.

APPENDIX TABLE A7: POSSIBLE DATA REVISIONS

| Country | Old Rate | New Rate | From Campaign | Reason |
|---|---|---|---|---|
| *Panel A: Revisions using added data from AJR (2005)* | | | | |
| Hong Kong | 14.9 | 285 | | Actual rate |
| Bahamas | 85 | 189 | | Actual rate |
| Australia | 8.55 | 14.1 | | Actual rate |
| Guyana | 32.18 | 84 | | Actual rate |
| Honduras | 78.1 | 95.2 | | Actual rate |
| Singapore | 17.7 | 20 | | Actual (shared) rate |
| Trinidad | 85 | 106.3 | | Lowest rate within Trinidad & Tobago |
| Sierra Leone (& Guinea) | 483 | 350 | ✓ | Eliminates mortality from Gambia |
| *Panel B: Revisions to correct for AJR's inconsistencies in Africa* | | | | |
| Algeria (& Morocco) | 78.2 | 15 | | Original not earliest peacetime rate. |
| Sudan | 88.2 | 30.36 | ✓ | Original not a peacetime rate. No peacetime rate available. Earliest long campaign rate. |
| Egypt | 67.8 | 32.6 | | Original not a peacetime rate |
| Madagascar | 536.04 | 75 | | Original not earliest peacetime rate. |

Data sources from Acemoglu et al. (2005). See Appendix for more information.

TABLE A8: FIRST STAGE ESTIMATES WITH NEW DATA FROM AJR (2005)
(Dependent Variable: Expropriation Risk)

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) | Without Africa (9) |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (64 countries, 36 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.61 | -0.52 | -0.40 | -0.44 | -0.35 | -0.29 | -0.42 | -0.44 | -1.21 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.19) | (0.19) | (0.25) | (0.18) |
| $p$-value of log mortality | 0.001 | 0.01 | 0.03 | 0.04 | 0.11 | 0.13 | 0.03 | 0.08 | 0.000 |
| $p$-value of controls | - | 0.17 | - | 0.40 | 0.34 | 0.001 | 0.02 | 0.20 | - |
| *Panel B: Original data revised with new data (64 countries, 40 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.60 | -0.48 | -0.35 | -0.38 | -0.29 | -0.24 | -0.37 | -0.38 | -1.06 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.21) | (0.21) | (0.19) | (0.19) | (0.24) | (0.27) |
| $p$-value of log mortality | 0.001 | 0.02 | 0.05 | 0.07 | 0.17 | 0.22 | 0.06 | 0.12 | 0.00 |
| $p$-value of controls | - | 0.11 | - | 0.18 | 0.14 | 0.001 | 0.01 | 0.10 | - |
| *Panel C: Revising with new data, removing conjectured mortality rates, and correcting Mali (34 countries and mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.62 | -0.48 | -0.30 | -0.32 | -0.20 | -0.13 | -0.29 | -0.25 | -0.86 |
| (heteroscedastic s.e.) | (0.22) | (0.25) | (0.22) | (0.23) | (0.25) | (0.25) | (0.25) | (0.27) | (0.26) |
| $p$-value of log mortality | 0.007 | 0.07 | 0.19 | 0.18 | 0.43 | 0.61 | 0.26 | 0.38 | 0.004 |
| $p$-value of controls | - | 0.10 | - | 0.01 | 0.005 | 0.000 | 0.013 | 0.01 | - |
| *Panel D: Revising with new data, adding campaign and laborer dummies (64 countries, 40 mortality rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.45 | -0.38 | -0.27 | -0.35 | -0.28 | -0.11 | -0.25 | -0.28 | -0.94 |
| (heteroscedastic-clustered s.e.) | (0.18) | (0.19) | (0.17) | (0.20) | (0.21) | (0.20) | (0.19) | (0.22) | (0.30) |
| $p$-value of log mortality | 0.02 | 0.05 | 0.12 | 0.09 | 0.19 | 0.56 | 0.19 | 0.21 | 0.01 |
| $p$-value of dummies | 0.12 | 0.17 | 0.26 | 0.25 | 0.36 | 0.10 | 0.16 | 0.23 | 0.17 |
| $p$-value of controls | - | 0.21 | - | 0.59 | 0.46 | 0.001 | 0.01 | 0.12 | - |
| *Panel E: Revising with new data, removing conjectured rates, correcting Mali, and adding campaign and laborer dummies (34 countries and rates)* | | | | | | | | | |
| Log mortality ($\beta$) | -0.36 | -0.22 | -0.10 | -0.25 | -0.10 | 0.02 | -0.15 | -0.14 | -0.78 |
| (heteroscedastic s.e.) | (0.22) | (0.24) | (0.21) | (0.25) | (0.24) | (0.24) | (0.23) | (0.27) | (0.29) |
| $p$-value of log mortality | 0.11 | 0.35 | 0.66 | 0.32 | 0.69 | 0.93 | 0.53 | 0.61 | 0.01 |
| $p$-value of dummies | 0.01 | 0.02 | 0.02 | 0.28 | 0.30 | 0.00 | 0.06 | 0.10 | 0.47 |
| $p$-value of controls | - | 0.11 | - | 0.14 | 0.15 | 0.001 | 0.04 | 0.03 | - |

See Panel A of Table A7 for added data and text for further detail.

TABLE A9: FIRST STAGE ESTIMATES WITH INCONSISTENCIES IN AFRICA CORRECTED
(Dependent Variable: Expropriation Risk)

| Control Variables | No Controls (1) | Latitude Control (2) | Without Neo-Europes (3) | Continent Dummies (4) | Continents & Latitude (5) | Mean Temp and Min Rain (6) | Percent European, 1975 (7) | Malaria in 1994 (8) |
|---|---|---|---|---|---|---|---|---|
| *Panel A: Original data (64 countries, 36 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.61 | -0.52 | -0.40 | -0.44 | -0.35 | -0.29 | -0.42 | -0.44 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.19) | (0.19) | (0.25) |
| *p* -value of log mortality | 0.001 | 0.01 | 0.03 | 0.04 | 0.11 | 0.13 | 0.03 | 0.08 |
| *p* -value of controls | - | 0.17 | - | 0.40 | 0.34 | 0.001 | 0.02 | 0.20 |
| *Panel B: Dropping countries - Egypt, Madagascar, and Sudan - with inconsistent observations (61 countries, 33 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.60 | -0.49 | -0.38 | -0.47 | -0.37 | -0.26 | -0.42 | -0.50 |
| (heteroscedastic-clustered s.e.) | (0.17) | (0.19) | (0.17) | (0.20) | (0.21) | (0.19) | (0.19) | (0.24) |
| *p* -value of log mortality | 0.002 | 0.02 | 0.03 | 0.03 | 0.09 | 0.18 | 0.04 | 0.05 |
| *p* -value of controls | - | 0.15 | - | 0.54 | 0.38 | 0.002 | 0.03 | 0.43 |
| *Panel C: Original data revising inconsistencies in Egypt, Madagascar, and Sudan (64 countries, 36 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.55 | -0.43 | -0.32 | -0.33 | -0.22 | -0.19 | -0.34 | -0.31 |
| (heteroscedastic-clustered s.e.) | (0.18) | (0.20) | (0.17) | (0.21) | (0.22) | (0.18) | (0.19) | (0.25) |
| *p* -value of log mortality | 0.005 | 0.04 | 0.08 | 0.13 | 0.32 | 0.30 | 0.09 | 0.23 |
| *p* -value of controls | - | 0.15 | - | 0.20 | 0.19 | 0.001 | 0.01 | 0.07 |
| *Panel D: Revising inconsistencies, removing conjectured mortality rates and correcting Mali (28 countries and mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.46 | -0.17 | -0.11 | -0.13 | 0.03 | 0.02 | -0.05 | 0.02 |
| (heteroscedastic s.e.) | (0.24) | (0.27) | (0.21) | (0.24) | (0.27) | (0.23) | (0.23) | (0.29) |
| *p* -value of log mortality | 0.07 | 0.53 | 0.59 | 0.59 | 0.90 | 0.94 | 0.83 | 0.94 |
| *p* -value of controls | 0.006 0.001 | 0.03 | - | 0.004 | 0.001 | 0.000 | 0.003 | 0.001 |
| *Panel E: Revising inconsistencies, adding campaign and laborer dummies (64 countries, 36 mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.40 | -0.33 | -0.24 | -0.25 | -0.17 | -0.04 | -0.20 | -0.14 |
| (heteroscedastic s.e.) | (0.19) | (0.21) | (0.18) | (0.23) | (0.24) | (0.21) | (0.20) | (0.25) |
| *p* -value of log mortality | 0.05 | 0.13 | 0.19 | 0.28 | 0.48 | 0.83 | 0.34 | 0.58 |
| *p* -value of dummies | 0.23 | 0.34 | 0.38 | 0.40 | 0.55 | 0.20 | 0.24 | 0.30 |
| *p* -value of controls | - | 0.26 | - | 0.389 | 0.367 | 0.000 | 0.008 | 0.045 |
| *Panel F: Revising inconsistencies, removing conjectured mortality rates, correcting Mali, and adding dummies (28 countries and mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.21 | 0.03 | 0.06 | -0.04 | 0.15 | 0.19 | 0.06 | 0.16 |
| (heteroscedastic s.e.) | (0.28) | (0.29) | (0.23) | (0.27) | (0.29) | (0.26) | (0.24) | (0.29) |
| *p* -value of log mortality | 0.45 | 0.92 | 0.81 | 0.90 | 0.62 | 0.46 | 0.81 | 0.57 |
| *p* -value of dummies | 0.09 | 0.25 | 0.14 | 0.53 | 0.56 | 0.01 | 0.25 | 0.24 |
| *p* -value of controls | - | 0.04 | - | 0.021 | 0.009 | 0.000 | 0.012 | 0.010 |
| *Panel G: Same as in Panel F revised with new data (34 countries and mortality rates)* | | | | | | | | |
| Log mortality ($\beta$) | -0.28 | -0.13 | 0.03 | -0.14 | 0.02 | 0.15 | -0.01 | -0.01 |
| (heteroscedastic s.e.) | (0.24) | (0.26) | (0.22) | (0.26) | (0.25) | (0.21) | (0.24) | (0.26) |
| *p* -value of log mortality | 0.25 | 0.62 | 0.90 | 0.60 | 0.95 | 0.46 | 0.98 | 0.97 |
| *p* -value of dummies | 0.04 | 0.12 | 0.05 | 0.51 | 0.56 | 0.00 | 0.12 | 0.40 |
| *p* -value of controls | - | 0.12 | - | 0.04 | 0.05 | 0.000 | 0.01 | 0.01 |

See Panel B of Table A7 for corrected inconsistencies and text for further detail.


FIGURE A1: ASSIGNMENT OF MORTALITY RATES TO LATIN AMERICA USING "BENCHMARKING"