ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Guinnane, Timothy

Working Paper

We Do Not Know the Population of Every Country in the World for the past Two Thousand Years

CESifo Working Paper, No. 9242

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Guinnane, Timothy (2021) : We Do Not Know the Population of Every Country in the World for the past Two Thousand Years, CESifo Working Paper, No. 9242, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/245423

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

CESifo WORKING PAPERS

9242
2021

August 2021

We Do Not Know the Population of Every Country in the World for the past Two Thousand Years

Timothy W. Guinnane

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9242*

We Do Not Know the Population of Every Country in the World for the past Two Thousand Years

Abstract

Economists have reported econometric results that rely on estimates of the population of every country in the world for the past two thousand or more years. The underlying source is usually McEvedy and Jones' *Atlas of World Population History*, published in 1978. The McEvedy and Jones data have important weaknesses. The reported populations for years before 1500 are, for most countries, little more than guesses, as are many estimates for more recent times. Research relying on McEvedy and Jones cannot take advantage of improved estimates reported since 1978. McEvedy and Jones often infer population sizes from their view of a particular economy, making their estimates poor proxies for economic growth. Although some economists treat the African data as pertaining to modern nation-states, in most cases it is not. With a few welcome exceptions, economists using this source do not take the measurement error issues seriously. Results that rest on McEvedy and Jones are unreliable. The willingness to rely on such data discourages effort to provide serious improvements.

JEL-Codes: N000, N100, O400, O470, J100, C180.

Keywords: long-run growth, historical populations, measurement error, McEvedy and Jones.

Timothy W. Guinnane
Department of Economics
Yale University
New Haven / CT / USA
timothy.guinnane@yale.edu

August 2, 2021
For comments and suggestions I thank Robert Allen, Cihan Artunç, Quamrul Ashraf, Gareth Austin, Howard Bodenhorn, Christophe Chamley, Tracy Dennison, Jeremy Edwards, Shari Eli, William English, James Fenske, Philip Hoffman, Maggie Jones, Thomas Mroz, Steven Nafziger, Tom Nicholas, Sheilagh Ogilvie, Paul Rhode, Larry Samuelson, Richard Tilly, Christopher Udry, and Qiyi Zhao. Benjamin English provided excellent research assistance.

"… we haven't just pulled the figures out of the sky. Well, not often."

McEvedy and Jones (1978, p.11)

In the past twenty years, economists have reported econometric exercises that rely in part on a published work that reports the population of every country in the world starting in the year 1 CE or even earlier. The existence of such data surprises those familiar with research on population history; we have only a rough idea of the population of most parts of the globe before 1500, and for many countries, the statistical lacuna extends closer to the present. Until the advent of modern censuses, which in most countries started during the nineteenth century, reckonings of the total population for even the best-studied cases remain subject to considerable error.

The source underlying these numbers is usually McEvedy and Jones's *Atlas of World Population History* (hereafter MJ). Published in 1978, this work draws on published literature and guesswork described below to report a population total for the countries of the world at intervals of a century or half-century. Many economics articles, include several highly-cited contributions in the leading journals, have used the MJ data in econometric exercises. The MJ estimates are also a source for historical population in the HYDE database.[1] The drawbacks to using such data are several. MJ's estimates (as they suggest themselves) lacked, in many cases, any firm foundation at the time they wrote. Often the estimates appear to reflect a judgment about the nature of the economy in question, rendering their use as economic proxies partially tautological. The MJ estimates are out-of-date for some countries; researchers have provided better figures in the past forty years. MJ note that they round their figures in ways that create

[1] The History Database of the Global Environment. See Goldewijk et al (2017). While this paper focuses on economics, other historically-minded social scientists have also relied on MJ. See, for example, Baumard et al (2015).

additional statistical problems. Economists have compounded these weaknesses with unwise disaggregation practices.

Reliance on MJ has two sets of undesirable implications. First, the population estimates suffer from considerable measurement error. Economists tend to dismiss measurement error issues by appealing to the implications of "classical" measurement error. For this context, classical measurement error is a strong and implausible assumption, and MJ's rounding practices are inconsistent with the assumption. The measurement error in MJ's population estimates is not classical. Second, the willingness to rely on data this rough discourages the additional work needed to provide better estimates going forward.

Table 1 demonstrates how important this source has become. Economists have used MJ for articles that appeared in the leading general-interest economics journals; in development and growth-oriented journals; and in the main field journals for economic history. Several of these papers have been cited many times. Economists routinely use population figures (or a derivative, such as urbanization rates) as a proxy for economic growth in periods prior to the availability of good GDP estimates. In those cases, MJ provide not just another control, but the key variable of interest. In their defense, MJ did not disguise the rough nature of their data, as their epigraph notes, and we should distinguish what they report from the way others used their work. Several economists point to a U.S. Census Bureau summary that appears to endorse MJ's estimates. The Bureau simply notes that MJ's estimates for *world* population are not too different from the others that appear in Table 2.[2] As MJ (pp. 353-4) state, that agreement is largely by construction.

The most natural way to assess the consequence of using MJ's estimates would be to re-estimate published models using correct data. This is obviously not possible. Instead, I first describe how MJ assemble their estimates and then discuss the way their approach likely biases econometric results. I then discuss more specifically the way some economists have used this population data.

[2] Acemoglu et all (2002, p. 1242) and Ashraf and Galor (2011, Note 14) each refer to a webpage that no longer exists. This page https://www.census.gov/data/tables/time-series/demo/international-programs/historical-est-worldpop.html lists only world population totals.

1. The source

MJ report their efforts as a series of graphs of total population in a country (or region – see below), with labels at centuries or half-centuries. Figure 1 reproduces the page for Germany as an example. For the twentieth century, and in some cases the nineteenth, MJ reproduce official census tabulations as discussed by earlier scholars, sometimes adjusted for changes in national boundaries.[3] Modern censuses did not start anywhere until the late eighteenth century were not widespread until the nineteenth century. Few countries have official national figures available before 1820. For the pre-census era, MJ rely on a variety of published sources as well as judgments as described below.

One would think from reading the economics literature that MJ report precise numbers. Graphs such as Figure 1, along with MJ's descriptions, suggest a different picture. "There are almost no data on which to base a population estimate for Germany until we reach the late Middle Ages" (p.70). "Estimates of Poland's population before the 14th century are based on nothing more than general ideas about likely [population – T.G.] densities" (p.76). For the Maghreb, "There is really nothing on which to base any calculations before the 19th century" (p.220). These comments are admirably frank, but MJ do, in fact, report population totals for Germany, Poland, and the Maghreb, and economists have used those observations to test hypotheses we view as important. MJ include a bibliography for each group of population estimates. While I do not doubt they consulted these sources, in most cases they do not say how they used the references they list. For Burma, to take an extreme example, they note that the quantitative record consists of a single publication based on a count of houses in 1783 as well as colonial censuses that began in 1871. Yet MJ report population sizes for that country as far back as 400BCE (pp.190-192).

[3] With exceptions discussed below, MJ rely on the definitions and boundaries of countries as of the year they wrote. Imputing constant-territory population for countries whose borders change over time raises additional issues I do not address.

How did MJ derive population estimates from before, as they say, there was anything on which to base such estimates? Reading their descriptions and examining the figures suggests five overlapping approaches. First, they start with the earliest official census and work backward. What Colin Clark (1968, p. 61) calls "jobbing back" can yield good population estimates given the right raw materials and technique. The population of a country in 1500 equals its 1600 population minus deaths and net emigrants, plus births in the period 1500-1600. Wrigley and Schofield (1981, Chapter 7) offer a good example of this approach. They start with the reliable census of 1841 and work back in time using careful estimates of births and deaths, along with more speculative estimates of net migration, to produce annual estimates of the total English population back to 1541. The challenge for these periods is that we rarely have anything like good counts of births and deaths, much less migrants, and the effort demands attention to complex sources. Creating the vital events series was the heart of Wrigley and Schofield's project. [4] Austin (2008, p.1002) stresses that momentous historical events such as the rise of the Atlantic slave trade greatly complicate such efforts. Few areas of the globe have been entirely spared such destabilizing episodes.

A second hint comes from the suspiciously round progression of population figures for single countries. In MJ's reckoning, England's population grew by 750 thousand between 1600 and 1650, and by another 750 thousand in the next half-century (p.43). Austria added 250 thousand people every 50 years between 1650 and 1800 (pp.88-92). Thailand added 250 thousand people in both the sixteenth and seventeenth centuries (p.193). Burma's population growth in the same period was 500 thousand per century.

[4] Wrigley and Schofield offered more than population totals: they estimated fertility and mortality rates, which require estimates of population age-structure not available from counts of births, deaths, and migrants. The technique of "back projection" they use has been criticized by Lee (1985) as being under-identified. The MJ estimates are for population totals alone. Scholars have questioned Wrigley and Schofield's estimates in many ways, but these criticisms rely on the fact that the original authors had substantive information and clear methods with which others could disagree. Deng (2004) argues that MJ ignored the possibility of using reasonable Chinese sources to do something similar for that country.

Third, MJ were apparently concerned that their estimates fit what they thought was the right model. In their view, population until the late medieval period grew at a roughly constant rate. In disagreeing with an earlier author on the right total world population for the year 1000, for example, MJ note that "our figure for AD 1, being 100m below the agreed figure for AD 1000, fits better on the sort of exponentially rising curve that everyone agrees best describes mankind's population growth" (p.354).[5] As the quotation implies, MJ also worried about consistency between theirs and earlier estimates.

Fourth, in the face of ignorance, MJ felt comfortable assigning identical growth rates to places they thought were similar. For 29 of the 82 observations that have populations reported for the years 1 and 1000, MJ's figures imply a percentage change of precisely 100 percent. For another seven countries, there is no change at all. In the interval from 1000 to 1500, their estimates imply that 12 of 96 places grew not at all, and another 19 increased by 100 percent. Similarly round numbers (25 percent, etc.) account for another 10 regions.[6]

Finally, MJ's background as geographers shows in several instances and may underlie other, more subtle judgments. Especially for population before 1500, they tended to reason from the nature of the economy and the population they thought it could support. They are rarely entirely explicit about this tactic, but it shows through in remarks such "likely population densities" quoted above. To the extent they estimate population in this way, MJ's figures reflect not the population of a particular country at a point in time, but their views about the population density consistent with the kind of economy MJ thought the country had.

Maddison

MJ's effort reflects a long interest in the world's population for distant times. MJ draw on these earlier efforts, which include Clark (1968) and Durand (1974). Table 2 summarizes the leading examples.

[5] Deng (2004, Note 3) stresses this feature, worrying that it implies an assumption that all population histories must resemble Europe's.

[6] Unless otherwise stated, calculations from MJ reported in this paper rely on a dataset provide by James Fenske. MJ report many areas as regions rather than countries, a practice Fenske retained.

Caldwell and Schindlmayr (2002) discuss the intellectual history of these research projects, expressing skepticism about the apparent consensus in the figures. MJ's effort differs from their predecessors in one important respect: not content to estimate the populations of continents or large regions, they usually report populations for the areas that correspond to modern nation-states.

Several of the articles discussed below rely in part on estimates reported by the late Angus Maddison. Maddison famously constructed, updated, and used a database that offered estimates of population and GDP/capita for most of the world's countries, again, in some versions, going as far back as the year 1 CE. Some of Maddison's critics suggest that for most of the historical record, his GDP estimates follow mechanically from population figures because he assumed constant output per person for long periods prior to industrialization. This criticism is not directly relevant to our concerns although it does imply, for some periods, another source of error in Maddison's figures. What does concern us is the relationship between MJ's and Maddison's population figures.

For the last major revision of his estimates, Maddison says of his population data: "The following detailed estimates for 1500 onwards rely heavily on monographic country studies for the major countries. To fill holes in my dataset I draw on McEvedy and Jones (1978). For the preceding millennium and a half, I use their work extensively" (Maddison 2001, p. 230). Maddison adds that he relies on MJ rather than earlier accounts because MJ are "the most detailed and best documented." Thus for many places before 1500, Maddison's database just reproduces MJ's figures. This is not always the case, however; Maddison was able to incorporate the fruits of research published between 1978 and his own publication. This led to some substantive revisions. He updated 28 percent of MJ's observations for the year 1000, for example, and 43 percent of the observations for 1500. The average absolute difference in his changes for those two years was 19 and 20 percent. Some individual changes are larger, however. Maddison added 50 percent to Mexico's population for the year 1000, and he doubled Peru's population in that same year (Maddison 2001, Table B-5, p. 235). He increased the population of the territory that would be the United States by 125 percent for the year 1500. For later periods, especially in the twentieth century, Maddison extends and revises the MJ estimates more comprehensively. In 1850, 84 percent of Maddison's 51

observations have different values, although the average absolute difference (3 percent) is smaller than for earlier years.[7]

The literature contains scattered mentions of disquiet with MJ's data, as well as Gregory Clark's assertion that for the years prior to 1820, MJ "is largely a work of imagination" (p.1160). Systematic discussion of MJ has been rare, but specialists tend not to be impressed. As Austin (2008, p.1102) puts it, "If you look up McEvedy and Jones expecting a treatise, detailing the original evidence and the reasoning behind the judgements by which it was converted into useable data, you will be disappointed." In discussing one particular study that relies heavily on MJ, Austin (2008, p.1002) says that "there is simply no epistemological basis for Nunn's use of the word `data' – literally, `things that are given' or granted – to refer to the guesses that have been made about the population of future African countries in 1400." [8]

2. Measurement error and rounding

Relative to "perfect" data for every country in the world, how far wrong will MJ take us? It is worth reviewing some general consequences of measurement error for the kinds of linear models that most researchers use.[9] Denote the true population of country *i* in year *t* as $\dot{P}_{it}$. P_{it} is the MJ estimate. The difference between MJ's estimate and the true population is measurement error ε_{it} such that $P_{it} = \dot{P}_{it} + \varepsilon_{it}$. Classical measurement error is the special cases where ε_{it} is uncorrelated with $\dot{P}_{it}$.We have two general implications. First, classical measurement error in the dependent variable alone does not bias estimates. The ε_{it} are swept into the regression error term, and the only consequence is some efficiency loss. Second, measurement error in any regressor implies bias in *all of the estimates*.

Consider the following regression:

[7] Maddison (2007) offers further revisions and extends some of the estimates back to ancient times. Most of these revisions pertain to the twentieth century.
[8] Austin is referring to Nunn (2008).
[9] Standard textbook treatments include Greene (2018, pp.102-3 and pp. 281-288) as well as Wooldridge (2010, pp.78-82).

$$Y_{it} = \alpha + \beta X_{it} + \gamma P_{it} + \mu_{it} \quad (1)$$

Where P_{it} is the mis-measured variable. While I write (1) for a panel framework that is not necessary to what follows. Classical measurement error in P_{it} implies that the estimate for γ will be smaller in absolute value than it would be if we could use $\dot{P}_{it}$ instead. The estimate is attenuated. The estimate for β will also be biased in ways we cannot ordinarily sign. The problem arises from the correlation between the measurement error ε_{it} and the regression error term μ_{it}. Measurement error is a particular kind of endogeneity, which is why some researchers employ instrumental-variable techniques in using the MJ data. Panel data does not offer a panacea. The fixed-effects estimator does not necessarily yield unbiased estimates in the presence of even classical measurement error. Panel data only "fixes" measurement error if the errors in P_{it} are, for each country i, the same for all years. In that case, the measurement error becomes part of the estimated country fixed effects (Deaton (1997, pp. 108-110)).

What does this mean for econometric studies that used mis-measured population estimates? If we maintain the assumption that the measurement error is classical, we can say two things. When population is the dependent variable the estimates may be less efficient, but there should be no bias due to measurement error alone. If population is a regressor, on the other hand, then the estimate for population will be attenuated. Additionally, the other estimates in this case will be biased and inconsistent. Thus using population as a "control" can lead to bias even for variables not thought to suffer from measurement error.

This sounds bad enough. We often want to control for population size. The situation is actually worse: while economists tend to assume that measurement error is always classical, in this case, we know that not to be the case. MJ state that they have rounded their estimates in ways that make the measurement error depend on the true value:

> All figures are rounded on the following system: below one million to the nearest .1 million, between one and 10 millions to the nearest .25 million, between 10 and 20 million to the nearest

> .5 million and between 20 and 100 millions to the nearest million. Above 100 million the rounding is to the nearest 5 million, above a billion… to the nearest 25 million. (p.9)

MJ's rounding procedure also creates a related, subtle problem when a country's population crosses one of the thresholds implied by their rounding rule. Portugal, they report, had a population of 900 thousand in 1400 and 1.25 million in 1500 (p.103). These figures imply that Portugal's population increased by 350 thousand people, or 39 percent, in those 100 years. Taking the rounding into account, however, implies upper and lower bounds for the population estimate in both 1400 and 1500. The true increase could be as small as 19 percent and as great as 58 percent.

The more general problem is that the rounding rules imply that the magnitudes of the measurement error are larger for larger populations. Thus, the measurement error MJ is not classical, and we cannot appeal to the implications of classical measurement error. Even if MJ had not adopted this rounding procedure, the assumption of classical measurement error in this case seems implausible: surely the error associated with a large country must be, in absolute value, larger than for a small country. We cannot, of course, know the extent of the problem without the true values, and if we had the true values this entire discussion would be moot. The appendix reports two exercises that attempt to get at the problem indirectly.

3. Circularity

Many economists who use MJ's figures think of population (or a derivative such as population density) as a proxy for an economic aggregate such as output. Critics such as Caldwell and Schindlmayr (2002) and Austen (2008) note that this is true in MJ's estimates in a way the economists may not intend, because MJ often use ideas about the economy to derive an estimate of population size. This is especially true in places and times for which the population data are thin; we learn that there was this kind of land and technology, and it could support this many people, and from that image, MJ infer a population size. In one explicit example, MJ discuss agricultural conditions in a region that comprises the modern states of

Columbia, Venezuela and the Guyanas to defend their assumption that until 1500, Colombia always accounted for 2/3 of the region's population (p. 302).

This procedure has obvious drawbacks in general. Austen stresses that this approach makes their estimates hostage to ideas about an economy and economic change. It is a particular problem for Africa. Maddison (2001, p.238), for example, adopts MJ's estimates for Africa in preference to earlier alternatives because MJ "assumed a more dynamic growth process." That is, Maddison preferred MJ's *population* estimates because he agreed with their assessment of the African *economy*. Neither Maddison nor MJ offer independent evidence about the African economy. To the extent MJ assigned population estimates based on their perceptions of economic performance, a regression using population as a proxy for growth tells us more about MJ than about economic growth.

4. Soft clones

Researchers who use MJ's data treat them as if they imply independent observations; put differently, that if there are N countries listed for a given year, this reflects N pieces of information. This is not always true, for two distinct reasons. MJ themselves create what I will call "soft clones." Frankly admitting that they lack meaningful data on country X, they assign to some countries the population dynamics of countries they think are similar. Sometimes they make this approach explicit. After concluding that Afghanistan has no useful population data before the 20th century, MJ say that "Perhaps the best approach is to compare Afghanistan with Iran" (p. 156). What they did, in fact, is to assume that Afghanistan had half the population of Iran in every year before 1900. The growth rate of the reported populations of Afghanistan and Iran are thus, for the period 1-1900 CE, identical. The population estimates for Afghanistan thus have two problems. They share any measurement error in the figures for Iran, and they also suffer from the error implied by any deviation of Afghanistan's true population history from Iran's.

MJ's data has many, less explicit examples of these soft clones. They often appeal to the idea that neighboring countries should have similar population growth rates: "… the fact that population doubled

in most European countries between A.D. 1000 and 1300 can be taken as strong evidence for it doing so in other European countries for which direct evidence is lacking" (p.11). Thus, in their reckoning, Poland, Hungary, and Czechoslovakia each grew 20 percent between 1000 and 1100. In the fifteenth century European Russia and China each grew by one-third. As late as 1600-1700, Romania and Austria each grew by 11.11 percent, and both Italy and Germany by 18.18 percent. We noted earlier that 35 percent of countries have identical population growth rates in the period 1-1000; those are all soft clones.

5. Hard clones

A final problem reflects both MJ's estimates and the way some economists have used them. MJ sometimes report estimates for regions rather than modern countries. Some economists create country-level populations out of the regions by using fixed weights to allocate the regional population among the modern nation-states. I will call the resulting countries "hard clones."

This procedure is most common for Africa. [10] Figure 2 reports the MJ page for a portion of western Africa. Some economists turn those estimates for a single region into populations for seven countries. The literature has two different ways to create countries out of regions. Ashraf and Galor (2011, 2013) disaggregate the regions by assigning weights that assume that each country within a region has the same population density in each year. Nunn (2008, p.170) uses weights that correspond to each country's share of the region's population in 1950. Nunn and Qian (2011) do not say explicitly how they disaggregated the regions, but for most countries their population figure are similar to Nunn's, so the approach is probably similar. [11]

As a general proposition, the population of the cloned countries has considerable measurement error. How much, once again, is something we cannot say precisely because we do not know the true

[10] Some authors use this same procedure to derive population sizes for the constituent parts of now-dissolved countries such as Yugoslavia. See, for example, the online appendix to Ashraf and Galor (2013).
[11] "For groups of some smaller countries, population data are only disaggregated to a regional level." (p.170). The aggregation also affects some of the largest countries, including Nigeria, Angola, and South Africa. The appendix for Nunn and Qian (2011) does not discuss this issue, nor does the replication code show how the hard clones were created.

populations of the clones in those years. The appendix reports an exercise that disaggregates the regions of the United States into the individual states using the two approaches discussed here. Comparing the resulting "states" to actual state populations demonstrates how wrong this approach can go.

Africa

MJ report only 12 regions for African; most of the African observations used in econometric estimates are hard clones. Africa in Nunn (2008) has 52 countries, while in Nunn and Qian (2011) it has 47.[12] Ashraf and Galor use different weights, so we should not be surprised that their country-level population estimates differ from Nunn or Nunn and Qian. The differences between their implied populations and Nunn and Qian's can be huge, however. Nunn and Qian estimate that Angola's population in 1500 was about 1.6 million, while Ashraf and Galor put it at 1.9 million. For Nigeria, Nunn and Qian estimate a population of 6.5 million, while Ashraf and Galor put it at 3.9 million. The percentage difference for smaller Liberia is even larger, with 157 thousand in the Nunn-Qian dataset and 417 thousand in Ashraf and Galor.[13]

Population in the cross-section thus differs considerably between the two datasets, not because of differences in the underlying source but because of the weights the authors use to break up the regions into countries. The differing results reflect one of two different sources of measurement error for hard clones. The first comes from MJ; MJ's regional estimates are themselves noisy. Creating clones assigns that noise to each of the country-level figures. The weights add a second source of measurement error, one that is negatively correlated for countries within a given MJ region. The cloned population estimates

[12] The difference in the number of observations reflects five small islands that Nunn (2008) treats as countries but Nunn and Qian (2011) exclude: Comoros, Cape Verde Islands, Mauritius, São Tomé and Principe, and Seychelles. Alternative specifications in Nunn (2008) show that these five observations do not drive the results. The Nunn (2008) replication dataset does not include population per se. Rather, I compute population using a variable that is the log of population density and another that is area. Nunn (2008) and Nunn and Qian (2011) sometimes do not agree on the population of a given country. Some differences are less than one percent of the population, and may reflect differences in rounding, but for one-fourth of the 47 countries, the absolute difference is more than 10 percent.

[13] By allocating historical populations according to population sizes in 1950, the Nunn-Qian approach makes the historical clone's size in earlier depend on later population growth. I thank Quamrul Ashraf for this observation.

cannot be "correct," not least because the true weights would not be the same for hundreds of years. (Put differently, to know the true weights we would have to know the actual populations in each year and thus would not need the weights.) The implications for change over time, however, are the same: in both datasets, 19 African countries grow by 100 percent between 1000 and 1500. These are just the growth rates MJ assign to the regions. Breaking these regions up into observations does not create more information. It just creates clones.

6. How economists use MJ

Table 1 shows that MJ has become a standard source for many economists. The citations alone do not tell us how it is actually used, however. To gather a clearer idea of how economists use these data, I examined every paper listed in Table 1 that was published in one of the first four journals listed until 2019. These four journals are the "Top 5" outlets that have published articles drawing on MJ. Some of these papers I set aside for the rest of this discussion. This list includes a few articles that cite MJ but do not use the data in econometric exercises. Shiue and Keller (2007, p.1194), for example, cite MJ and other authorities as implying that their two regions, China and Europe, had similar populations at the end of the eighteen century. Rogers (1994, p.467) cites MJ to defend an assumption that long-term population growth rates were nearly zero until relatively recent. This usage seems consistent with the spirit in which MJ offer their estimates. I also set aside papers that only use MJ's estimates for 1900 and later. By that date, the information MJ report comes almost entirely from reasonable census reports. This includes articles such as Acemoğlu et al (2001).[14] Table 3 summarizes several dimensions of the remaining papers. Table 3 does not necessarily do justice to some discussion in these articles, as we will see. The table entries stress the way the article uses MJ in the main analysis; the population data in some cases appears in different ways in robustness checks.

[14] I also set aside two intermediate cases. Putterman and Weil (2010) rely on MJ's year 1500 estimates for some of their analysis, but they first aggregate the data to 11 large regions "because population data for 1500 are very noisy, particularly at the country level." See their note 11. Voigtländer and Voth (2012) rely on MJ's figures for Europe in the period 1000-1700, but do not report econometric estimates.

The first question is the time period for which the article relies on MJ. The year 1500 does not form a magical dividing line, but it is the earliest year for which we have anything like reliable estimates for populations of even most European countries, which tend to have the best-founded estimates. Several of the articles listed in Table 3 depend in a serious way on MJ's population estimates from before 1500. Ashraf and Galor (2011)'s report econometric results that depend critically on population data from the years 1 and 1000. The same is true of Nunn and Qian (2011), which starts with the year 1000. Nunn (2008) relies on the 1400 estimates alone.[15] Most of the other papers listed also rely on data from 1500-1800.

The second issue is whether MJ's population figures form the dependent variable or a regressor. Here I include variables constructed from the MJ population data and something else. The most common such examples are population densities and the urbanization rate. As noted, even classical measurement error cannot be benign when the variable in question is a regressor. Many articles use as their dependent variable either population itself, or a quantity divided by population. The latter has the effect of controlling for population size but doing so in the dependent variable, where the measurement error does least harm. These include Acemoğlu et al (2002, 2005) as well as Ashraf and Galor (2011, 2013) and Nunn and Qian (2011).[16]

Some articles, however, create regressors from MJ's estimates. This list includes Iyigun (2008) as well as Gennaioli and Voth (2015). Iyigun (2008) studies whether military pressure from the Ottoman Empire helped reduce conflict among European states in the early-modern period. The econometric models rely on annual observations for the period 1450-1700. The dependent variables measure intra-European conflict. The controls include measures of Ottoman military pressure as well as the populations of Europe and, in some specifications, the Ottoman Empire's. Iyigun describes the population data as a proxy for economic "size and strength." (p.1476). The estimated effect for European population size is

[15] "Rely on" in the sense that these years are part of the sample used to estimate the primary regression results. I have not re-estimated the models to assess whether the results hold if some years are dropped. Nunn and Qian (2011) do not need data from the western hemisphere.
[16] Acemoğlu at el (2005) use MJ's population figures to weight their regressions.

imprecisely-estimated in most specifications, while Ottoman population is more precisely estimated but switches signs, depending on the dependent variable. The point-estimates for both population variables must be attenuated because of measurement error, so we cannot really say whether Europe became more peaceful simply because of economic growth, nor can we assess the implications of Ottoman economic conditions for European conflict. Moreover, the estimates for his main variable of interest, the extent of Ottoman military incursion into Europe, may be biased because of the measurement error in population.

Gennaioli and Voth (2015, Table 3) address a related question, and their population figures cause similar trouble. They study the determinants of battle success in early-modern European conflicts. The authors set this up as a horse-race between fiscal capacity on the one hand and population size on the other. Greater fiscal capacity allows a state to pay more mercenaries and support more allies. Population size could matter to early-modern war because larger populations make it easier to field larger armies. In most specifications, the fiscal variable has a positive and significant effect on battlefield success, while the relative populations of the two combatants has almost none. They interpret their regressions as saying "Differences in population size do not have a systematic effect on the chance of battlefield success." (p.1430) This result could reflect nothing more than the measurement error in MJ's estimates. The population variable suffers from the standard attenuation, and because of the measurement error in population, the estimate for fiscal capacity is biased in unknown ways.

The third issue pertains to how the authors discuss the possibility of measurement error in the population data. Acemoğlu et al (2002, 2008) explicitly discuss measurement error. Most of the other articles by Acemoğlu and co-authors avoid using the population estimates as a regressor, and when they do, they use 2SLS models that are motivated in part by a desire to overcome it (for example, Acemoğlu et al (2002, 2008)). Others take a different approach. Ashraf and Galor (2011, p. 2011) claim:

> The most comprehensive worldwide cross-country historical estimates of population and income per capita since the year 1 CE have been assembled by Colin McEvedy and Richard Jones (1978) and Angus Maddison (2003), respectively. Indeed, despite inherent problems of measurement associated with historical data, these sources remain unparalleled in providing comparable

> estimates across countries in the last 2,000 years and have, therefore, widely been regarded as standard sources for such data in the long-run growth literature.

They do not argue that MJ's data meet any particular standard. Rather, they know of nothing better (it is "unparalleled") and everyone else uses it (it is the "standard source in the long-run growth literature").[17]

Nunn and Qian (2011, p.616) address measurement error more explicitly, but their discussion consists of general statements that are not relevant to the MJ data:

> Accuracy is an obvious concern for historical data that span such a long time horizon and broad cross-section. However, classical measurement error in our outcome variables will not bias our regression estimates. Similarly, any systematic measurement error that varies by time-period or by country is captured by the country and year fixed effects, which are included in all specifications.

Population is their dependent variable so they are correct that if the measurement error is classical, it does not bias their results. They provide no reason to think this is true, and as noted, MJ say it is not true. The second statement about fixed effects is equally true but irrelevant to the case; neither of these extreme assumptions is likely. Nor can they be true simultaneously.

7. Conclusions

We know the population of the United States in the year 2020 to a high degree of accuracy. If we wanted to improve on that accuracy, we could draw on a variety of sources to reduce the error in the existing estimates. This is probably true of many other times and places although in some instances the improvements may be relatively small. Refining the estimates for Poland in 1400, for example, may just require consulting more published works, but might require original research using (for example) essentially archeological approaches. But it can be done. It would not be useful to assert that because we

[17] Ashraf and Galor (2011) treat the entire western hemisphere as a single observation with its own dummy variable "which is natural given the historical period examined." (Notes to Table 2, p. 2015.) This may be a reference to measurement error.

cannot know the population of Poland in 1400 to the same accuracy as we can in 2020 that there is no point in using historical population counts. The opposite extreme is more common and pernicious: with some welcome exceptions, many economists take the view that the accuracy of the data does not matter because they can never be as precise as modern reports.

Econometric estimates that rely on MJ appear as part of particular literatures that take "cross-country regression" approaches to economic growth and related questions. Economists differ on the wisdom of the general approach, and those who favor such studies may insist that some data are better than none. Even those who take this view, however, should be aware of the pitfalls in the source and the way some use it. As I have noted, Acemoğlu and his co-authors tend to use MJ as carefully as one can. Others have been less careful, in some cases compounding MJ's weaknesses by trying to create information that is not in the source.

Some of the research discussed here appeals to the idea that classical measurement error does not cause bias in linear models when the measurement error affects the dependent variable only. This observation is mathematically true but not relevant to the MJ data. MJ adopt rounding rules that make the measurement error depend on the true value in every case. Thus the error is not classical. More generally, the idea of classical measurement error in this case strains credulity: how would the errors for larger populations *not* be larger?

This paper documents a series of problems in a published source that underpins many articles published in the leading general-interest economics journals. Publication in these outlets has strong professional rewards and conveys signals. One signal is that if everyone does something inappropriate, then it is fine. The second signal discourages the original work necessary to improve the basis of our knowledge. The researchers who did the spadework on which MJ is based understood themselves as contributing to broader literatures in the social sciences. Their contributions were rewarded within their own niches. The same applies to all of the effort that went into constructing the considerable information on historical economies that Maddison summarizes. To the extent the profession signals lack of interest in

such work then it is unlikely we will ever learn more about, for example, the population of Poland in 1400.

We can, and in fact have, done better. Figure 3 summarizes population growth rates for England for the period 1600-1850 implied by both MJ and Wrigley and Schofield. The differences are substantial. MJ missed the population decline of the second half of the 17th century and significantly under-stated the population growth of the first half of the 18th century. Population figures underlie, directly, any statement about per-capita GDP or its growth rate and are thus central to understanding the Industrial Revolution. Wrigley and Schofield set a high standard, and the evidentiary basis for this project is less rich in many places. Yet we can do much better than MJ's guesses for many countries, especially in the period since 1500 (eg, Vos 2014, pp. 366-369).

This discussion also contains a simpler lesson. Many economists today download a dataset and merge it onto other datasets without consulting the original sources. Examining MJ's book is instructive. The introduction explains the problem of non-classical measurement error. A cursory look at the graphs (such as my Figure 1) would lead most to treat the data with considerable caution. And anyone look at the Figure 2 should immediately wonder how Africa can have so many countries in the dataset.

Appendix

Rounding

MJ's rounding rules create a possible dependence of the measurement error on the true value, which would imply that the measurement error is not classical. Since we do not know the true value we cannot assess this problem's extent directly. But we can consider two exercises that shed some light. The first uses the population of the fifty United States at the decadal census years 1900-1970.[18] I apply the same rounding criteria that MJ use and compare the implied measurement error to the true values. The results are mixed. The simplest way to assess the problem is to compute the zero-order correlation between the true value and the implied measurement error. For the dataset as a whole, the correlation is small. In some years, however, we cannot reject to null of no correlation. The differences reflect sensitivity to particular values, especially the large errors created by the gross rounding for larger states. Only a handful of states had more than twenty million people, for example, and that handful can drive the results.[19]

Appendix Table A1 reports the distribution of these errors for four census years. If the rounding produced symmetric errors within the size groups, and econometric estimates used only populations from a single population size group, then there would be no correlation between the true value and the measurement error. For the U.S. data, the first condition is never true: there appears to be no greater tendency to round up than to round down, but in every census year, there is more of one than the other. The second condition, not using countries from different rounding groups, could be true but that would make it hard to undertake the kind of research for which MJ is used. Note the difficulty with the largest group; here the errors can be largest, as indeed they are for the single state in 1920 that had more than 20 million people. That state, New York, had 20.77 million inhabitants in that year. If New York had

[18] I create a balance panel of 50 states by treating as states territories such as Hawaii that did not become states until after 190.

[19] The correlation for all years is -.06 (p=.20). For 1920, it is -.23 (p=.10); for 1940, -.49 (p=0); and for 1950, -.23 (p=.11).

reported one million fewer, it would have fallen into the 10-20 million category and its population would have been rounded down, instead of up, and its apparent error would be negative instead of positive.

A second approach sticks closer to MJ. I use a Monte Carlo experiment to "unround" the numbers MJ report. Define λ as the rounded-off portion of the population estimate. We do not know λ but we can assume its distribution and approximate it empirically with a Monte Carlo exercise. MJ's rounding rules imply an upper and lower bound for λ in each size group. For example, if the reported value is between one and ten million, λ must lie on an interval from minus to plus 250 thousand. I assume that the λs follow a uniform distribution on that interval, and generate random numbers to construct true values using the figures MH report. Applying this procedure to every country in the MJ dataset yields one possible pre-rounded set of populations. I then compute the correlation between the "true" estimates this experiment creates and the rounding. By running the Monte Carlo many times, I consider many possible values of the rounding. In a large majority of these cases, MJ's population and λ are correlated at significant levels. The magnitudes depend on the year, because in earlier years more countries fit into a single rounding group. In the year 1000, 68.65 of the replications imply a correlation that is significantly different from zero at a 5 percent or better confidence level. In 1500, that figure is 80.52 percent, and in 1800 it is 85.08. In 81 percent of these Monte Carlo experiments, the correlation between the true value and the measurement error is significant at a level of 5 percent or better.

This simulation relies on the assumption that λ has a uniform distribution. Appendix Table A1 shows this is not an obvious assumption, but it does not suggest a particular alternative, either. Re-running the Monte Carlos assuming λ follows a normal distribution implies even stronger correlations between the true value and the error. Asymmetric distributions might imply something different, but the basic message remains: the rounding procedure makes the measurement error depend on the true value. This is strong evidence that the measurement error is not classical.[20]

[20] The results differ slightly across years, which reflects the changing distribution of countries across MJ's rounding groups. (This is the same issue as in the U.S. state data). I considered the years 1000, 1500, 1700, and 1800, and in each case a Monte Carlo based on 10,000 experiments implies that in at least 80 percent of the cases, the p-value for the correlation is less than .05.

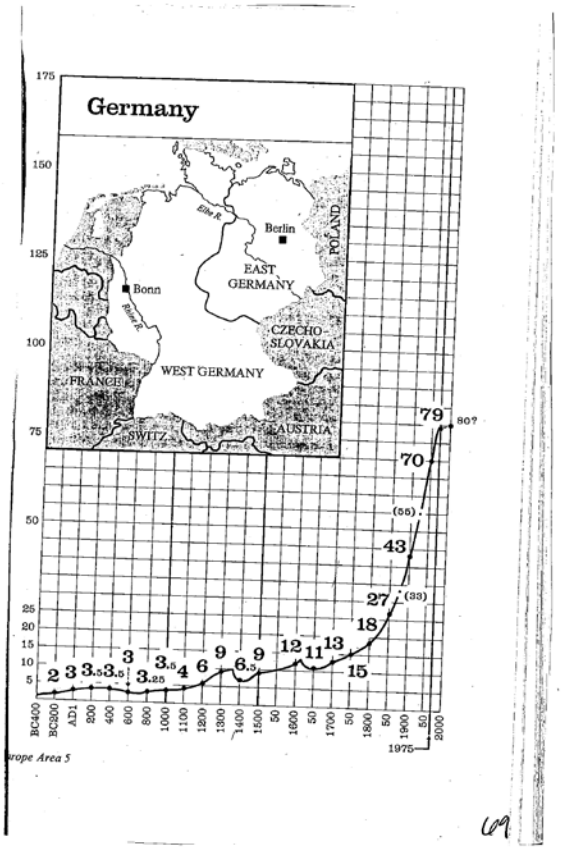
Hard clones

We do not know the true population of Angola in 1400 and thus cannot precisely state the consequences of creating hard clones the way some economists have done. We can, however, examine other situations where we know both the population of a region and the population of its sub-units. Using the ICPSR's standard regions, I aggregate the U.S. states into four regions and then disaggregate them into "states" that are clones of the sort researchers create for Africa. This allows me to compare the actual state population to the clone population. In what follows, P^T denotes the state population as reported by the census. I construct two sets of "clones" from the, regional-level population data. P^{NQ} assumes that each state's share of its regions's population is the same as in 1940. P^{NQ} approximates the Nunn-Qian approach with one difference: their weights come from 1950, long after any observation in their sample. Note that for 1940, $P^T = P^{NQ}$ by construction, which is not true for any year in their data. To construct P^{AG} I assume that the population density of each state within a region is equal for every census year. P^{AG} is what Ashraf and Galor do for Africa.

Next I computed E^{NQ} and E^{AG} by subtracting the "cloned" population using the two methods from the true value for that state and year, and examine the correlation between the true value (which we know) and the measurement error created by the cloning. The results show that cloning leads to severe problems in some cases. If we examine the years 1900-1930, relying on the Nunn-Qian approach and 1940 population weights, the correlation between the true value and the error is -.17 (p=.02). Using the 1940 weights seeks to mimic Nunn and Qian's use of 1950 weights. If we use those same years and weights taken from the 1970 census, the results are better: the correlation is -.05 (p=.49). This difference is not really good news for the Nunn-Qian approach. The weights for 1970 turn out to be closer to correct for the period 1900-1930 than those from 1940. This is unexpected and one cannot assert it would be true in any particular context. The results are even worse for Ashraf and Galor's approach. For the entire period 1900-1970, the correlation is .26 (p=0) and for 1900-1930 it is .30 (p=0).

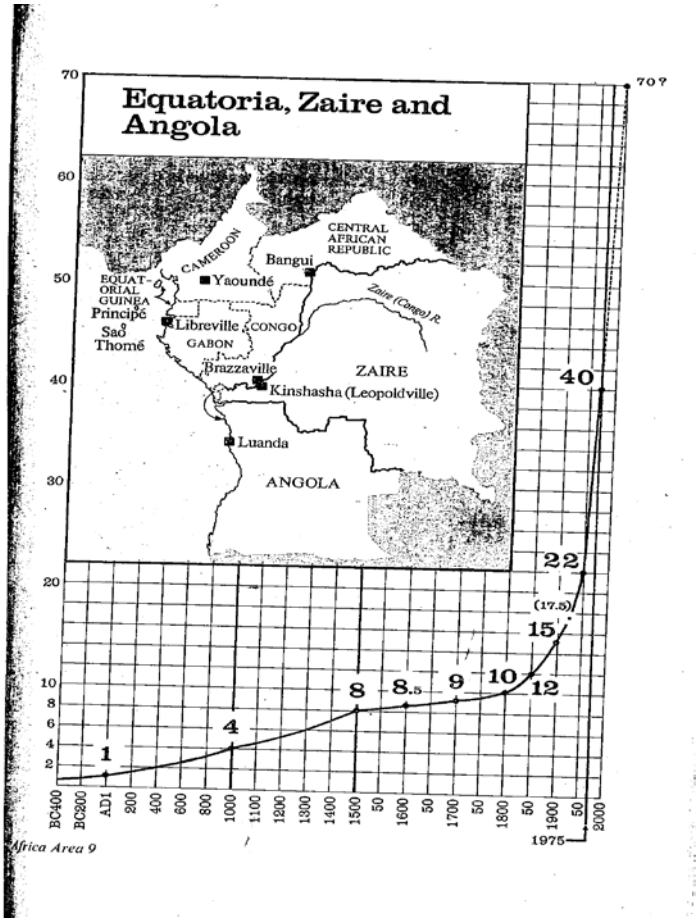
Is this a fair comparison? Nunn-Qian's approach assumes that countries within a region grew at the same rate. This assumption fails in the U.S. I compute the percentage change in population from 1900 to 1940 for each state. In an ANOVA framework, region only "explains" about 28 percent of the variation in this variable. Most heterogeneity in state-level population growth is within, not across, regions. For Ashraf-Galor the issue is whether the countries had the same population densities. Their assumption performs especially poorly in the U.S. because of the unequal population densities within some U.S. regions; the "Midwest" region, for example, includes states like Ohio (204 persons per square mile in 1900) as well as states like North Dakota (9 persons per square mile).

Figure 1: MJ's page for Germany



MJ p.69

Figure 2: MJ's page for western Africa



MJ p. 249

Figure 3: Average annual population growth for England, MJ versus Wrigley and Schofield, 1600-1850

Source: MJ p.43; Wrigley and Schofield 1981, Appendix Table A.3

Note: MJ's estimates are for England and Wales, while Wrigley and Schofield's are for England less Monmouth. MJ's population estimate for 1600 is 4.2 million; Wrigley and Schofield's estimate is about 4.1 million. Wrigley and Schofield (1981, Table A5.3) use an inflation factor of .93383 to convert "England and Wales" to "England less Monmouth." Using this factor, the MJ estimates are about 11 percent lower than Wrigley and Schofield in 1600 and about 2 percent higher in 1850.

Table 1: Citations to MJ in selected economics journals

Journal	Number of citations to MJ
American Economic Review	7
Journal of Political Economy	1
Quarterly Journal of Economics	7
Review of Economic Studies	3
Journal of Development Economics	8
Economic Development and Cultural Change	1
Journal of Economic Growth	9
The Economic Journal	4
Journal of Economic History	13
Explorations in Economic History	4
European Review of Economic History	3
Economic History Review	6

Source: Google scholar as of December 31, 2019

Notes: This list is restricted to the main general-interest, development, and economic history journals, along with two others that often publish articles on economic growth. There were no citations located for *Econometrica* or the *World Bank Research Observer*. The figures for the *American Economic Review* exclude the May issue.

Table 2: Alternative estimates of the population, by continents, 0-1700

Year	0	1000	1500	1700
		Europe (including former USSR)		
Clark	44500	44200	73800	111800
Durand	42500	45500	79000	n.a.
Biraben	43000	43000	84000	125000
McEvedy and Jones	32800	38800	85500	126150
Maddison	33350	39013	87718	126810
		Americas		
Clark	3000	13000	41000	13000
Durand	12000	37500	46500	n.a.
Biraben	12000	18000	42000	12000
McEvedy and Jones	4500	9000	14000	13000
Maddison	6320	12860	19750	13250
		Asia (including Australasia)		
Clark	185000	173000	227000	416000
Durand	207000	189500	304000	n.a.
Biraben	171000	152000	245000	436000
McEvedy and Jones	114200	183400	277330	411250
Maddison	174650	183400	284350	402350
		Africa		
Clark	23000	50000	85000	100000
Durand	35000	37500	54000	n.a.
Biraben	26000	38000	87000	107000
McEvedy and Jones	16500	33000	46000	61000
Maddison	16500	33000	46000	61000

Note: Figures in thousands.

Source: Maddison Appendix B, Table B-1.

Table 3: How the MJ data have been used in general-interest economics journals

	(1)	(2)	(3)	(4)
Article	Years for which MJ data used	How used	Is there discussion of measurement error?	Africa?
Acemoglu et al AER 2005	1300, 1400, 1500, 1600, 1700, 1750, 1800, 1820, 1850	Dependent variable	No	Yes
Acemoglu et al AER 2008	1500	Regressor	No	Yes
Acemoglu and Johnson JPE 2005	1500	Dependent variable	No	Yes
Acemoglu et al QJE 2002	1,1000,1500,1800	No	Yes	Yes
Ashraf and Galor AER 2011	1, 1000, 1500	Dependent variable	No	Yes
Ashraf and Galor AER 2013	1, 1000, 1500	Dependent variable	No	Yes
Gennaioli and Voth REStud 2015	1500-1800	Regressor	No	No
Iyigun QJE 2008	1401 to 1700	Regressor	No	Yes
Nunn QJE 2008	1400	Regressor	No	Yes
Nunn and Qian QJE 2011	1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, 1850, 1900	Dependent variable	No	Yes

Notes: See text for discussion and qualification of these simple classifications. In particular, the years reported in (1) include both the primary analysis and robustness checks. (2) is coded as “regressor” for any article that uses MJ as regressor in the central analysis, even if it also uses MJ to construct dependent variables for some purposes.

Appendix Table A1: Implications of rounding for U.S. state populations, selected years 1900-1960

Rounding category	25th	Median	75th	Number of states
1900				
<1 million	-14,938	-3,338	23,544	17
1-10 million	-69,368	13,250	65,090	31
10-20 million	37,788	71,009	104,230	2
1920				
<1 million	-31,676	-1,208	11,793	11
1-10 million	-36,758	43,744	86,824	35
10-20 million	-59,966	-29,440	18,788	3
20-100 million	229,546	229,546	229,546	1
1940				
<1 million	3	19,478	39,029	8
1-10 million	-73,308	33,870	76,536	35
10-20 million	199,640	185,001	170,352	6
20-100 million	-41,716	-41,716	-41,716	1
1960				
<1 million	-39,868	-29,444	-20,238	5
1-10 million	-55,046	15,131	75,654	36
10-20 million	-87,206	133,564	146,388	5
20-100 million	398,330	-99,476	298,362	4

Source: computed from the populations of the United States as described in the text.

Note: the table reports the measurement error implied by following MJ's rounding rules for the states. The error is defined as true population – rounded population, so a positive value implies rounding down.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson, 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review*, 91(5), pp. 1369-1401.

Acemoglu, Daron, Simon Johnson and James A. Robinson, 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4), pp. 1231-1294.

Acemoglu, Daron, Simon Johnson, and James Robinson, 2005. The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *The American Economic Review*, 95(3), pp 546-579.

Acemoglu, Daron and Simon Johnson, 2005. Unbundling Institutions. *Journal of Political Economy*, 113(5), pp. 949-995.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared, 2008. Income and Democracy. *The American Economic Review*, 98(3), pp. 808-842.

Ashraf, Quamrul and Oded Galor, 2011. Dynamics and Stagnation in the Malthusian Epoch. *The American Economic Review*, 101(5), pp 2003-2041.

Ashraf, Quamrul and Oded Galor, 2013. The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *The American Economic Review*, 103(1), pp. 1-46.

Austin, G., 2008. The 'reversal of fortune' thesis and the compression of history: Perspectives from African and comparative economic history. *Journal of International Development: The Journal of the Development Studies Association*, *20*(8), pp.996-1027.

Baumard, N., Hyafil, A., Morris, I. and Boyer, P., 2015. Increased affluence explains the emergence of ascetic wisdoms and moralizing religions. *Current Biology*, *25*(1), pp.10-15.

Caldwell, J.C. and Schindlmayr, T., 2002. Historical population estimates: Unraveling the consensus. *Population and development review*, *28*(2), pp.183-204.

Cantoni, David and Noam Yuchtman, 2014. Medieval Universities, Legal Institutions, and the Commercial Revolution. *The Quarterly Journal of Economics*, 129(2), pp 823-888.

Clark, Colin., 1968. *Population growth and land use*. MacMillan.

Clark, Gregory, 2009. Review of Angus Maddison's Contours of the World Economy. *Journal of Economic History* pp. 1156-1161.

Deaton, Angus, 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: Johns Hopkins.

Deng, Kent G., 2004. "Unveiling China's True Population Statistics for the Pre-Modern Era with Official Census Data." *Population Review* 43(2).

Durand, John D.,1974. *Historical Estimates of World Population: An Evaluation*. Philadelphia.

Galor, Oded and Andrew Mountford, 2008. Trading Population for Productivity: Theory and Evidence. *The Review of Economic Studies*, 75(4), pp. 1143-1179.

Gennaioli, Nicola and Hans-Joachim Voth, 2015. State Capacity and Military Conflict. *The Review of Economic Studies*, 82(4), pp. 1409-1448.

Goldewijk, K Klein., Beusen, A., Doelman, J. and Stehfest, E., 2017. Anthropogenic land use estimates for the Holocene–HYDE 3.2. *Earth System Science Data*, *9*(2), pp.927-953.

Greene, William H., 2018. *Econometric Analysis*. (8th edition).

Iyigun, Murat, 2008. Luther and Suleyman. *The Quarterly Journal of Economics*, 123(4), pp. 1465-1494.

Maddison, Angus, 2001. *The World Economy: A Millenial Perspective*. OECD

Maddison, Angus, 2007. *Contours of the World Economy, 1-2030: Essays in Macro-Economic History*. Oxford.

McEvedy, C. and Jones, R., 1978. *Atlas of world population history*. Penguin Books Ltd, Harmondsworth, Middlesex, England.

Nunn, Nathan, 2008. The Long-Term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics*, 123(1), pp. 139-176.

Nunn, Nathan and Nancy Qian, 2011. The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment. *The Quarterly Journal of Economics*, 126(2), pp. 593-650.

Putterman, Louis and David N. Weil, 2010. Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality. *The Quarterly Journal of Economics*, 125(4), pp. 1627-1682.

Rogers, Alan R., 1994. Evolution of Time Preference by Natural Selection. *The American Economic Review*, 84(3), pp. 460-481.

Shiue, Carol H. and Wolfgang Keller, 2007. Markets in China and Europe on the Eve of the Industrial Revolution. *The American Economic Review*, 97(4), pp. 1189-1216.

Voigtländer, Nico and Hans-Joachim Voth, 2013. The Three Horsemen of Riches: Plague, War, and Urbanization n Early Modern Europe. *The Review of Economic Studies*, 80(2), pp. 774-811.

Vos, Jelmer, 2014. "Work in Times of Slavery, Colonialism, and Civil War: Labor Relations in Angola from 1800 to 2000." *History in Africa* 41: 363-385

Wooldridge, Jeffrey M., 2010. *Econometric Analysis of Cross Section and Panel Data*. (2nd edition). Cambridge, MA: MIT Press.

Wrigley, E. A. and Roger Schofield, 1981. *The Population History of England 1541-1871*. Cambridge, MA: Harvard University Press.