The European Origins of Economic Development

William Easterly and Ross Levine*

August 2015

Abstract: Although a large literature argues that European settlement outside of Europe during colonization had an enduring effect on economic development, researchers have been unable to assess these predictions directly because of an absence of data on colonial European settlement. We construct a new database on the European share of the population during colonization and examine its association with economic development today. We find a strong, positive relation between current income per capita and colonial European settlement that is robust to controlling for the current proportion of the population of European descent, as well as many other country characteristics. The results suggest that any adverse effects of extractive institutions associated with small European settlements were, even at low levels of colonial European settlement, more than offset by other things that Europeans brought, such as human capital and technology.

Keywords: Institutions; Human Capital; Political Economy; Natural Resources

JEL Classification Codes: 043; 01; P48, N5

Countries have followed divergent paths of economic development since European colonization. Some former colonies, such as the Congo, Guinea-Bissau, Malawi, and Tanzania, have experienced little economic development over the last few centuries, with current per capita Gross Domestic Product (GDP) of about $2 per day. Others are among the richest countries in the world today, including Australia, Canada, and the United States, with per capita GDP levels of about $140 per day. Others fall along the spectrum between these extremes.

To explain these divergent paths, many researchers emphasize that the European share of the population during colonization shaped national rates of economic growth through several mechanisms. Engerman and Sokoloff (1997) (ES) and Acemoglu, Johnson, and Robinson (2001, 2002) (AJR) stress that European colonization had enduring effects on political institutions. They argue that when Europeans encountered natural resources with lucrative international markets and did not find the lands, climate, and disease environment suitable for large-scale settlement, only a few Europeans settled and created authoritarian political institutions to extract resources. The institutions created by Europeans in these "extractive colonies" impeded long-run development. But, when Europeans found land, climate, and disease environments that were suitable for smaller-scale agriculture, they settled, forming "settler colonies" with political institutions that fostered development. This perspective has two testable implications: (1) former settler colonies with a large proportion of Europeans during colonization will create more inclusive political institutions that foster greater economic development than former extractive colonies with a small proportion of Europeans, and (2) colonial European settlement will have a stronger association with development today than current European settlement (the proportion of the population that is of European descent today) because of the enduring effect of institutions created during the colonial period.

As an additional, potentially complementary mechanism, ES and Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) (GLLS) note that the European share of the population during colonization influenced the rate of human capital accumulation. They argue that Europeans brought human capital and human capital creating institutions that shape long-run economic growth, as emphasized by Galor (2011). According to this human capital view, European settlers directly and

immediately added human capital skills to the colonies and also had long-run effects on human capital accumulation. These long-run effects emerge because human capital disseminates throughout the population over generations and it takes time to create, expand, and improve schools. Furthermore, this human capital view suggests that a larger share of Europeans during colonization could facilitate human capital accumulation across the entire population both because it would increase interactions among people of European and non-European descent and because it might accelerate expanded access to schools, as emphasized in ES. This human capital view also yields two testable implications: (1) the proportion of Europeans during colonization will be positively related to human capital development and hence economic development today, and (2) the proportion of Europeans during colonization will matter more for economic development than the proportion of the population of European descent today because of the slow dissemination of human capital and creation of well-functioning schools. Although the political institutions and human capital views emphasize different mechanisms, they provide closely aligned predictions about the impact of colonial European settlement on current economic development.

Other researchers, either explicitly or implicitly, highlight additional mechanisms through which European migration had positive or negative effects on development. North (1990) argues that the British brought comparatively strong political and legal institutions that were more conducive to economic development than the institutions brought by other European nations. This view stresses the need for a sufficiently strong European presence to instill those institutions, but does not necessarily suggest that the proportion of Europeans during colonization will affect economic development today beyond some initial threshold level. More recently, Spolaore and Wacziarg (2009) stress that the degree to which the genetic heritage of a colonial population was similar to that of the economies at the technological frontier positively affected the diffusion of technology and thus economic development, where European migration materially affected the genetic composition of economies. Putterman and Weil (2010) and Chanda, Cook, and Putterman (2014) emphasize that the experiences with statehood and agriculture of the ancestors of people currently living within countries help explain cross-country differences in economic success. And,

Comin, Easterly, and Gong (2011) likewise find that the ancient technologies of the ancestors of populations today help predict per capita income of those populations. In all of these papers, the ancestral nature of a population helps account for cross-country differences in economic development, where European colonization materially shaped the composition of national populations.[1] As in the human capital view, the emphasis is on things that Europeans brought with them, such as technology.

Although this considerable body of research emphasizes the role of European settlement during colonization on subsequent rates of economic development, what has been missing in the empirical literature is the key intermediating variable: colonial European settlement. While researchers, including AJR, have examined the European share of the population in 1900, this is well after the colonial period in several countries, including virtually all of the Western Hemisphere. To the best of our knowledge, researchers have not directly measured colonial European settlement and examined its association with current economic development.

In this paper, we construct a new database on the European share of the population during colonization and use it to examine the historical determinants of colonial European settlement and the relation between colonial European settlement and current economic development. Although we do not isolate the specific mechanisms linking colonial European settlement with current levels of economic development as emphasized in each of the individual theories discussed above, we do assess the core empirical predictions emerging from the literature on the relationship between European settlement and economic development. In particular, we assess whether the proportion of

---

[1] An extensive and growing body of research explores the historical determinants of economic development, which has been insightfully reviewed by Spolaore and Wacziarg (2013). For example, Michalopoulos and Papaioannou (2013) show that pre-colonial political institutions had enduring effects on regional economic development, while Michalopoulos and Papaioannou (2014) show that variation between African ethnic groups is more important than variations between nations in Africa in explaining comparative economic development, advertising the broader notion that different peoples carry growth-shaping features with them across borders. Furthermore, as suggested by the work of Bisin and Verdier (2000), Fernandez and Fogli (2009), and Tabellini (2008), culture also extend beyond national borders with prominent effects on economic development. And, other scholars address the deep historical roots of modern-day levels of social capital, civic capital, or democracy, including Stephen Haber (2014), Luigi Guiso, Paola Sapienza, Luigi Zingales (2013), Torsten Persson and Guido Tabellini (2010), and Tabellini (2010).

Europeans during colonization is positively related to economic development today and whether the proportion of Europeans during colonization is more important in accounting for cross-country differences in current economic development than the proportion of the population of European descent today.

We begin by compiling a new database on the European share of the population during colonization. For each country, we gather data from an assortment of primary and secondary sources for as many years as possible going back the 16th century. From these data, we construct several measures of European settlement that differ with respect to the date used to measure colonial European settlement. We first construct a country-specific measure based on the colonial history of each country. To do this, we use information on when Europeans first arrived in the country and when the colony became independent to select a date on which to measure colonial European settlement. For this country-specific measure, we seek a date, subject to data limitations, that is early in its colonial period but sufficiently after Europeans first arrived to allow for the formation of colonial institutions. We next construct measures that use a common date. For each country, we average the annual observations on the European share of the population over 1500-1800, 1801-1900, and 1500-1900. We obtain consistent results using these different methods for dating and measuring colonial European settlement.

We then examine the historical determinants of colonial European settlement both to check the credibility of our new data and to examine differing views about the factors shaping European colonization. As a guide, we employ a very simple model of the costs and benefits of European settlement. Some determinants have already been discussed in the literature, such as pre-colonial population density, latitude, and the disease environment facing Europeans. Pre-colonial population density raises the costs to Europeans of obtaining and securing land for new settlers, and might also raise the benefits since the European often exploited and enslaved the indigenous population. Latitude raises the benefits of simply transferring European technologies (such as for agriculture) to the newly settled areas. A harsh disease environment facing Europeans raises the expected costs of settlement.

To this list of common determinants of European settlement, we add one very important new variable: indigenous mortality from European diseases. Indigenous mortality from European diseases is a tragic natural experiment that is a very good predictor of European settlement, since it removed or weakened indigenous resistance to Europeans invading new lands, and made plenty of fertile land available to settlers. The phenomenon is limited to lands that had essentially zero contact with Eurasia for thousands of years, since even a small amount of previous contact was enough to share diseases and develop some resistance to them. For example, trans-Sahara and trans-Indian Ocean contacts were enough to make Africa part of the Eurasian disease pool (McNeil 1976, Karlen 1995, Oldstone 1998). Historical studies and population figures show that only the New World (the Americas and Caribbean) and Oceania (including Australia and New Zealand) suffered large-scale indigenous mortality due to a lack of resistance to European diseases (McEvedy and Jones 1978).

Our examination of the historical determinants of colonial European settlement yields three findings. First, we find that colonial European settlement tends to be smaller (as a share of total population) in areas where there was a highly concentrated population of indigenous people and where the indigenous population did not die in large numbers from diseases brought by Europeans. This finding provides the first direct empirical support for AJR's (2002) hypothesis that in areas with a high concentration of indigenous people, Europeans did not settle in large numbers and instead established extractive regimes. This finding is a key building block in AJR's (2002) theory of a "reversal of fortunes," in which formerly successful areas, i.e., areas with a high concentration of indigenous people, became comparatively poorer due to the enduring effects of extractive political regimes. Second, Europeans tended to settle in large concentrations in lands further from the equator. Third, although biogeography—a measure of the degree to which an area is conducive to the domestication of animals and plants—explains human population density before the era of European colonization (Ashraf and Galor 2011), it does not account for colonial European settlement after accounting for indigenous population density.

We next assess the two key predictions emerging from the political institutions and human capital views concerning colonial European settlement and current economic development and discover the following. First, colonial European settlement is strongly, positively associated with economic development today. This relationship holds after controlling for British legal heritage, the percentage of years the country has been independent since 1776, and the ethnic diversity of the current population. The strong, positive association between European settlement and economic development today is also robust to controlling for the mortality of the indigenous population during colonization, latitude, availability of precious metals, distance from London, ability to cultivate storable plants and domesticate animals, malaria ecology, and European mortality during colonization, as well as soil quality, access to navigable waterways, and continent dummy variables. However, the relationship between economic development today and the proportion of Europeans during colonization weakens markedly when controlling for either current educational attainment or government quality, which is consistent with the views that human capital and political institutions are intermediating mechanisms through which European settlement shaped current economic development.

Second, the European share of the population during colonization is more strongly associated with economic development today than the percentage of the population today that is of European descent. Europeans during the colonization era seem to matter more for economic development today than Europeans today. This finding is consistent with the view that Europeans brought growth-promoting characteristics—such as institutions, human capital, technology, connections with international markets, and cultural norms—that had enduring effects on economic development. This result de-emphasizes the importance of Europeans per se and instead emphasizes the impact of what Europeans brought to economies during colonization.

The estimated positive relation between colonial European settlement and current development is economically large. Based our parameter estimates, we compute for each country the projected level of income in 2000 if colonial European settlement had been zero. We then compare this counterfactual level of current income to actual current income and compute the share

of current income attributable to colonial European settlement. The data and estimates indicate that 40% of current development *outside of Europe* is associated with the share of Europeans during the colonial era. Though such projections must be treated cautiously, they suggest that the European origins of economic development deserve considerable attention.

We check the robustness of the positive relation between colonial European settlement and current economic development in several ways. First, we were concerned that the relation between colonial European settlement and current development might breakdown when eliminating "neo-Europes," such as Australia, Canada, New Zealand, and the United States, or other countries that had a comparatively high proportion of Europeans during colonization. Thus, we redid the analyses while omitting countries with colonial European settlement greater than 12.5%, which represents a natural break in the data that omits the neo-Europes, Argentina, and a few, small Latin American countries.

Rather than the results breaking down when the sample is restricted to countries with a small proportion of Europeans during colonization, the estimates become larger. When examining only those former colonies with colonial European settlement less than 12.5%, we find that the estimated positive relation between current income and colonial European settlement is more than double the estimate from examining the full sample of non-European economies.

We were also concerned about that the inclusion of many countries for which our data indicate zero colonial European settlement might affect the parameter estimates. Besides statistical robustness, the previous literature has not explicitly addressed whether colonial institutions with small European settlement during colonization are better or worse than those with no settlement.[2] Even if small colonial European settlements created worse institutions than those created in areas with no Europeans during colonization, the positive things that Europeans brought with them, such as human capital or technology, could offset the negative development effects of worse institutions. Our data allow us to provide a first evaluation of these issues.

---

[2] Below, we discuss the few cases of non-European countries that escaped colonial rule altogether.

We address this in two ways. First, we omit all of these "zero" countries from the analyses and confirm the results. Second, we include a dummy variable for countries in which our data indicate that colonial European settlement equals zero. We continue to find that colonial European settlement enters positively and significantly in the economic development regressions. We also find that the dummy variable for zero European settlement often enters with a positive coefficient, which provides some empirical support for the political institutions view that small European settlements that create extractive institutions are worse for current economic development than countries with essentially no European settlers during the colonial era. Combining the coefficient estimates on colonial European settlement and the dummy variable, the findings indicate that once colonial European settlement is above 4.8%, any adverse effects from extractive institutions associated with small colonial European settlements were more than offset by other things that Europeans brought during colonization, such as human capital, technology, familiarity with global markets, and institutions, that had enduring, positive effects on economic development.

Ample qualifications temper our conclusions. First, we do not assess the welfare implications of European colonization. Europeans often cruelly oppressed, enslaved, murdered, and even committed genocide against indigenous populations, as well as the people that they brought as slaves (see Acemoglu and Robinson 2012 for compelling examples). Thus, GDP per capita today does not measure the welfare effects of European colonization; it only provides a measure of economic activity today within a particular geographical area. Although there is no question about European oppression and cruelty, there are questions about the net effect of European colonization on economic development today. Second, we do not separately identify each potential channel through which the European share of the population during colonization shaped long-run economic development. Rather, we provide the first assessment of the relationship between colonial European settlement and comparative economic development and thereby inform debates about the sources of the divergent paths of economic development taken by countries around the world since the colonial period.

The remainder of the paper is organized as follows. Section 1 defines and discusses the data, while Section 2 provides preliminary evidence on the determinants of human settlement prior to European colonization and the factors shaping European settlement. Section 3 presents the paper's core results on the relationship between colonial European settlement and current economic development. Section 4 reports an exercise in development accounting to calculate what share of global development can be attributed to Europeans. Section 5 concludes.

1. Data

This section describes the two data series that we construct: (1) the European share of the population during colonization and (2) the degree to which a region experienced large scale indigenous mortality due to the diseases brought by European explorers in the 15th and 16th centuries. The other data that we employ are taken from readily available sources, and we define those variables when we present the analyses below.

1.1 Euro share

We compile data on the European share of the population during colonization (*Euro share*) from several sources. Since colonial administrators were concerned about documenting the size and composition of colonial populations, there are abundant—albeit disparate—sources of data. Of course, there was hardly anything like a modern statistical service in colonial times, so that different administrators across different colonies in different time periods used different and often undocumented methods for assembling population statistics. Thus, we use a large variety of primary and secondary sources on colonial history to piece together data on the European share of the population.

Although the Data Appendix provides detailed information on our sources, the years for which we compiled data on each country, and discussions about the quality of the data, it is worth emphasizing a few points here. First, we face the challenge of choosing a date to measure European share. We would like a date as early as possible after initial European contact to use European

settlement as an initial condition affecting subsequent developments. At the same time, we do not want to pick a date that is too early after European contact since it is only after some process of conquest, disease control, and building of a rudimentary colonial infrastructure that it became possible to speak of a European settlement. Given these considerations, we try to choose a date at least a century after initial European contact, but at least 50 years before independence. This means that for conceptual reasons we do not seek to use a uniform date across all colonies. For example, Europeans were colonizing and settling Latin America long before colonizing Africa. We also lack a continuous time series for each country; rather, the data reflect dates when colonial administrators in particular locales happened to measure or estimate populations.[3] Given these data limitations, we cannot always adhere to our own guidelines for choosing the date on which to measure *Euro share*. In sensitivity analyses discussed below, we show that the results are robust to measuring *Euro share* as the average value over three different uniform periods: (i) 1500-1800, (ii) 1801-1900, or (iii) 1500-1900.

Second, we adopt a "dog did not bark" strategy for recording zero European settlement. If we find no historical sources documenting any European settlement in a particular colony, we assume that there were no such settlers. This procedure runs the risk of biasing downward European settlement. However, we believe colonial histories (which are virtually all written by European historians) are extremely unlikely to fail to mention significant European settlements. We checked and confirmed the validity of this procedure of setting colonial European settlement to zero using the Acemoglu et al. (2001) data appendix, which gives the share of Europeans in the population in 1900. Furthermore, as presented below, the results hold when eliminating all countries with zero colonial European settlement or when including a fixed effect for these countries in the regression analyses.

---

[3] When we have several observations near our "ideal" date for measuring colonial European settlement, we take the average. The online dataset provides the date of each observation.

1.2 Indigenous mortality

We examine several predetermined factors that potentially influenced European settlement—including the degree to which Europeans brought diseases that wiped out the indigenous population. Others have carefully documented this tragic experience, but we believe that we are the first to use it to explain the nature of colonization and its effect on subsequent economic development.

Although Europeans established at least a minimal level of contact with virtually all populations in the world during the colonial period, this contact had truly devastating effects on indigenous populations in some regions of the world but not in others. Some regions had been completely isolated from Eurasia for thousands of years, and thus had no previous exposure or resistance to Eurasian diseases. When Europeans then made contact with these populations—which typically occurred during the initial stages of global European exploration and hence long before anything resembling "European settlements," European diseases such as smallpox and measles spread quickly through and decimated the indigenous population. For example, when the Pilgrims arrived in New England in 1620, they found the indigenous population already very sparse because European fisherman had occasionally landed along the coast of New England in the previous decades. Similarly, De Soto's expedition through the American South in 1542 spread smallpox and wiped out large numbers of indigenous people long before British settlers arrived.

Thus, we construct a dummy variable, *Indigenous mortality*, which equals one when a region experienced large-scale indigenous mortality due to the spread of European diseases during the initial stages of European exploration. To identify where Europeans brought diseases that caused widespread fatalities, we use the population data of McEvedy and Jones (1978) and three epidemiological world histories (McNeil 1976, Karlen 1995, Oldstone 1998). Diseases had circulated enough across Eurasia, Africa and the sub-continent, so that indigenous mortality did not shoot up with increased exposure to European explorers, traders, and slavers during European colonization. The New World (Americas and Caribbean) and Oceania (the Pacific Islands, Australia, and New Zealand) were different. When European explorers and traders arrived, the

microbes that they brought triggered extremely high mortality rates, which accords with their previous isolation from European diseases. The evidence suggests that mortality rates of 90 percent of the indigenous population after European contact were not unusual. Although we compiled a country-by-country variable for large-scale indigenous mortality, our review of the evidence and the historical narrative indicates little measurable variation within the New World and Oceania. As a result, *Indigenous mortality* wound up being a simple dummy for countries in the New World and Oceania. (As discussed below, however, this paper's findings on the associations of current development and colonial European settlement are robust to including continent fixed effects.) This dummy variable measure suggests caution in interpreting the results on *Indigenous mortality*. Although the data indicate that large-scale indigenous mortality occurred in the New World and Oceania but not elsewhere (McEvedy and Jones, 1978, McNeil, 1976, Karlen, 1995, Oldstone, 1998), *Indigenous mortality* is ultimately a dummy variable for these regions of the world and might proxy for other features of these regions, such as geographic isolation, rather than European-induced mortality. These same areas that were isolated from Europeans prior to colonization—and hence more susceptible to European-borne diseases—also had lower population densities in 1500 AD. This may be related to Spolaore and Wacziarg's (2009) result on diffusion of technology as a function of when different branches of humanity became separated. Populations in Oceania and the Western Hemisphere had been isolated from the rest for a very long time, and hence they did not get either (1) the more advanced technology originating in the Old World that would have helped support a larger population or (2) the exposure to European diseases before colonization that would have helped them become more resistant to European diseases and hence to European settlement. We will see that this combination of low indigenous population density and vulnerability to European diseases plays a large role in accounting for where Europeans settled.

2. Preliminaries: Where Did Europeans Settle?

Table 2 provides regression results concerning which factors shaped European settlement during colonization. The dependent variable is the proportion of Europeans in the colonial population (*Euro share*).

The regressors in Table 2 are as follows. First, we include *Population density 1500*. Note that this is a measure of the pre-Columbian population for the New World (i.e. before 1492) even though the date is conventionally rounded off to 1500. Hence, this number does NOT include any initial population decrease due to indigenous mortality to European diseases.

Since the regressions also include other variables to control for the attractiveness of the land for settlement, we examine the relationship between *Euro share* and *Population density 1500* conditional on the generalized attractiveness of the land for human settlement. A plausible interpretation of the conditional impact of *Population density 1500* on *Euro share* is that it gauges the ability of the indigenous population to resist European settlement.[4] Second, *Indigenous mortality* provides additional information on the inability of the indigenous population to resist European settlers. If European diseases eliminate much of the indigenous population, this would clearly reduce their ability to oppose European settlement.

Third, *Latitude* might have special relevance for European settlers to the extent that they are attracted to lands with the same temperate climate as in Europe. *Latitude* measures the absolute value of the distance of the colony from the equator.

Fourth, *Precious Metals* is an indicator of whether the region has valuable minerals since this might have affected European settlement. Fifth, one cost of settling in a particular country

---

[4] Although there could also be a mechanical negative relation between indigenous *Population density 1500* and *Euro share* because the denominator of *Euro share* is the sum of the indigenous and settler populations, we normalize European settlers by total colonial population because the political institutions and human capital views frame their predictions about the enduring effects of the colonial period on economic development in terms of the *proportion* of Europeans in the colonial population. The indigenous population could potentially attract European settlers to the extent that the indigenous peoples represent a readily available labor supply to be exploited by the Europeans. Thus, the net effect of the indigenous population on European settlement is an empirical question. Our result of *Euro share* responding negatively to log indigenous population density is consistent with some positive response of *absolute numbers* of European settlement to indigenous population as long as the elasticity of that response is less than one.

might be its distance from Europe, so we use the distance from London to assess this view (*London*). Finally, we examine other possible determinants of the attractiveness of the land for settlement, including *Biogeography, Malaria ecology*, and *Settler mortality. Biogeography* is an index of the prehistoric (about 12,000 years ago) availability of storable crops and domesticable animals, where large values signify more mammalian herbivores and omnivores weighing greater than 45 kilograms and more storable annual or perennial wilds grasses, which are the ancestors of staple cereals (e.g., wheat, rice, corn, and barley).[5] *Malaria ecology* is an ecologically-based spatial index of the stability of malaria transmission in a region, where larger values signify a greater propensity for malaria transmission.[6] *Settler mortality* equals historical deaths per annum per 1,000 European settlers (generally soldiers, or bishops in Latin America) and is taken from AJR (2001).

The results show that three factors account for the bulk of cross-country variation in European settlement. First, the density of the indigenous population matters. In regions with a high concentration of indigenous people who could resist European occupation, Europeans comprised a much smaller fraction of the colonial population than in other lands. Second, indigenous mortality matters. Where the indigenous population fell drastically because of European diseases, European settlers were more likely to settle. Third, there is a positive relationship between *Euro share* and *Latitude*, even when conditioning on *Population density 1500* and *Indigenous mortality*. Europeans were a larger proportion of the colonial population in higher (more temperate) latitudes, plausibly because of the similarity with the climate conditions in their home region.[7][8]

---

[5] Taken from Hibbs and Olsson (originally 2004, later expanded to a larger sample), *Biogeography* equals the first principal component of (a) the number of annual perennial wild grasses known to exist in the region in prehistoric times with mean kernel weight of greater than ten milligrams and (b) the number of domesticable large mammals known to exist in the region in prehistoric times with a mean weight of more than 45 kilos. Ashraf and Galor 2011 found another version of this measure to be a good predictor of the timing of transition to agriculture and through that channel a good predictor of 1500 AD population density.

[6] The Malaria ecology index is from Kiszewski et al (2004) and measures the biological characteristics of mosquitoes that influence malaria transmission, such as the proportion of blood meals taken from human hosts, daily survival of the mosquito, and duration of the transmission season and of extrinsic incubation.

[7] Ashraf and Galor 2011 find that latitude had the opposite effect on where pre-1500 population was dense – there was less settlement in temperate regions and more in tropical regions.

[8] We explored robustness of these results to adding other geographic variables that will also be considered in the exercises below: soil fertility and distance to waterways from Ashraf and Galor (2011, 2013), and continent dummies. The results are robust to these additions.

These three characteristics, *Population density 1500*, *Indigenous mortality*, and *Latitude* help explain in a simple way the big picture associated with European settlements, or the lack thereof, in regions around the world. Where all three factors were favorable for European settlement, such as Australia, Canada, New Zealand, and the United States, the European share of the colonial population was very high. When only some of the three factors were favorable, there tended to be a small share of European settlers. Latin America suffered large-scale indigenous mortality, but only some regions were temperate, and most regions had relatively high pre-Columbian population density (which is why more people of indigenous origin survived in Latin America compared to North America, even though both regions experience high indigenous mortality rates when exposed to European diseases). Southern Africa was temperate and had low population density, but did not experience large-scale indigenous mortality. These factors can also explain where Europeans did not settle. The rest of sub-Saharan Africa was tropical and again did not experience much indigenous mortality from exposure to the microbes brought by Europeans during colonization. And, most of Asia had high population density, did not suffer much indigenous mortality from European borne diseases, and is in or near the tropics, all of which combine to explain the low values of *Euro share* across much of Asia.

None of the other possible determinants that we consider are significant after controlling for these three determinants. Indeed, European colonial settlement, unlike pre-Columbian population (the latter as verified by Ashraf and Galor (2011, 2013)), was NOT driven by the intrinsic, long-run potential of the land—as measured especially by *Biogeography*.

One of the most famous variables in the literature on explaining European settlement is *Settler mortality*. Our data on colonial settlement allows for the first assessment of the ability of this variable to explain European settlement during colonization. The results are mixed. *Settler mortality* has a negative and significant simple correlation with colonial European settlement (not shown), confirming the prediction in AJR. It becomes insignificant when including the three variables that we found most robust in accounting for colonial European settlement, and does not materially alter the statistical significance of the other variables. Yet, when we include all RHS variables

simultaneously (in column 8 of Table 2), which yields a much smaller sample of countries, *Settler mortality* returns to significance. In sum, the relation between *Euro share* and *Settler mortality* is highly sensitive to changes in the sample and the control variables.

.

3. Results: Europeans during Colonization and Current Economic Development

3.1 Simple graphical analyses

To assess the relationship between the European share of the population during colonization and the current level of economic development, we begin with simple graphs. We measure the current level of economic development as the average of the log of real per capita GDP over the decade from 1995 to 2005 (*Current income*). Using data averaged over a decade reduces the influences of business cycle fluctuations on the measure of current economic development.

Figure 1 shows (1) the number of countries with values of *Euro share* within particular ranges, (2) the actual countries with these values of *Euro share*, and (3) the corresponding median level of *Current income* for countries with values of *Euro share* within the particular ranges. Two key patterns emerge. First, median *Current income* is positively associated with *Euro share*. Second, very few countries have Euro share greater than 0.125. While ES and AJR do not provide an empirical definition of a "settler colony," we use 12.5% as a useful benchmark.

Figures 2a and 2b illustrate the relationship between *Current income* and *Euro share* using *Lowess*, which is a nonparametric regression method that fits simple models to localized subsets of the data and then smooths these localized estimates into the curves provided in Figures 2a and 2b. Figure 2a illustrates the relationship for the full sample of non-European countries. Figure 2b provides the curve for the sub-sample of countries with measured values of *Euro share* less than 12.5%. Figure 2c omits zero observations from Figure 2b. As shown, the relationship between *Euro share* and *Current income* is positive throughout. There is no apparent region in which an increase in *Euro share* is associated with a reduction in *Current income,* although 2c shows a steeper relationship at very low values of Euro share than 2b. We examine this more formally below.

3.2 Euro share and economic development today

In this section, we use regressions to condition on a range of national characteristics and assess the independent relationship between *Current income* and *Euro share*.

We consider the following cross-country regression:

$$\text{Current income} = \alpha * \text{Euro share} + \beta' X + u, \tag{1}$$

where X is a matrix of national characteristics that we define below, and u is an error term, potentially reflecting economic growth factors that are idiosyncratic to particular countries, as well as omitted variables, and mis-specification of the functional form. Different theories provide distinct predictions about (a) the coefficient on *Euro share* (α), (b) whether α changes when conditioning on particular national characteristics, and (c) how α changes across sub-samples of countries.

We get some insight into the channels connecting *Euro share* and *Current income* by examining how α changes when controlling for the different potential channels discussed above: political institutions and human capital. If *Euro share* is related to current levels of economic development through the formation of enduring political institutions, then *Euro share* may not enjoy an association with economic development today conditioning on political institutions. And, if *Euro share* is related to economic development today through the spread of human capital, then *Euro share* may not have an association with development today conditioning on educational attainment today. Of course, both current political institutions and educational attainment are endogenous to current economic development, so these findings must be interpreted cautiously.

We begin by evaluating equation (1) while conditioning on an array of national characteristics (X). *Legal origin* is a dummy variable that equals one if the country has a common law (British) legal tradition. This dummy variable both captures the argument by North (1990) that the United Kingdom instilled better growth-promoting institutions than other European powers and

the view advanced by La Porta et al (2008) that the British legal tradition was more conducive to the development of growth-enhancing financial systems than other legal origins, such as the Napoleonic Code passed on by French and other European colonizers. Further differentiating across different civil law traditions, as in La Porta et al. (2008), does not alter the results. *Education* equals the average gross rate of secondary school enrollment from 1995 to 2005 and is taken from the World Development Indicators. *Independence* equals the fraction of years since 1776 that a country has been independent. As in Beck, Demirguc-Kunt, and Levine (2003) and Easterly and Levine (2003), we use this to measure the degree to which a country has had the time to develop its own economic institutions. It could also be interpreted as a measure of the duration (and hence, perhaps, intensity) of recent colonialism across countries. *Government quality* is an index of current level of government accountability and effectiveness and is taken from Kaufman et al. (2002). *Ethnicity* is from Easterly and Levine (1997) and measures each country's degree of ethnic diversity. In particular, it measures the probability that two randomly selected individuals from a country are from different ethnolinguistic groups. Since the purpose of our research is to examine the impact of European settlement outside of Europe, all of the regressions exclude European countries.

Using ordinary least squares (OLS), Table 3a shows that there is—with two notable exceptions—a positive and statistically significant relation between *Current income* and *Euro share*. For example, regression (1) indicates that an increase in *Euro share* of 0.1 (where the mean value of *Euro share* is 0.07 and the standard deviation is 0.17) is associated with an increase in *Current income* of 0.36 (where the mean value of *Current income* is 8.2 and the standard deviation is 1.3). Table 1 lists descriptive statistics for the main variables used in the analyses. Below, we provide more detailed illustrations of the magnitude of the relation between the European share of the population during colonization and the current level of economic development. The strong positive link between the European share of the population during colonization and current economic development holds when conditioning on different national characteristics. Indeed, when simultaneously conditioning on *Legal origin, Independence*, and *Ethnicity*, the results hold.

The exceptions are that the coefficient on *Euro share* falls materially and becomes insignificant when conditioning on either *Education* or *Government quality*. These findings are consistent with—though by no means a definitive demonstration of—the view that the share of Europeans in the population during colonization shaped long-run economic development by affecting political institutions and human capital accumulation.

These results could be driven by a few former colonies in which Europeans were a large fraction of the population during economic development and that just happen to be well-developed former colonies today. As indicated in the introduction, an open question in the existing literature is whether small colonial European settlement (associated with extractive institutions) is better or worse for later development than colonial rule with no settlers. Thus, we conduct the analyses for a sample of countries in which *Euro share* was less than 12.5%. The goal of restricting the sample to only those countries where Europeans account for a small proportion of the population is to assess whether the relation between *Euro share* and *Current income* holds when there is only a small minority of Europeans. While there is no formal definition of what constitutes a "small European colony" that is conducive to extractive institutions, we use less than 12.5% European as a conservative benchmark of a small colonial European settlement and because there is a natural break in the distribution of *Euro share* across countries at this level.

As shown in Table 3b, however, the coefficient on *Euro share* actually becomes larger when restricting the sample to those countries in which *Euro share* is less than 12.5% (in the regressions that do not condition on either *Education* or *Government quality*). The increase in the coefficient on *Euro share* when restricting the sample to former colonies with *Euro share* less than 12.5%, suggests that the relationship between the European share of the population during colonization and the level of economic development does not simply represent the economic success of "settler colonies." Rather, a marginal increase in *Euro share* is associated with a bigger increase in subsequent economic development in colonies with only a few Europeans—one might characterize this as the diminishing marginal long-run development product of *Euro share*. Table 3b also shows that the relationship between *Current income* and *Euro share* remains sensitive to controlling for

political institutions and human capital accumulation. The association between *Current income* and *Euro share* shrinks and becomes insignificant when conditioning on *Education* or *Government quality*.

The coefficient on the *British legal origin* dummy variable is never significant (nor will it be in the rest of the paper). It is also of interest that many of the colonies with *Euro share* < 0.125 were Spanish colonies. Hence we find no evidence for the popular view that British colonization or legal origin led to more development than Spanish colonization or legal origin.

As robustness tests, we next expand the conditioning information set (*X*) and report these sensitivity tests in Table 3c. In particular, we first repeat the analyses in Tables 3a and 3b except that in all of the regressions we consider two additional sets of conditioning information: control variables from Table 2, and continent fixed effects. The Table 2 variables were associated with *Euro share*; the association in Tables 3a and 3b of today's development with *Euro share* may reflect an underlying association with the determinants of *Euro share* instead. Hence, in Table 3c, we simultaneously control for *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology, and Settler mortality*. As in the Table 3a and 3b analyses, we introduce *British legal origin, Education, Independence, Government quality,* and *Ethnicity* one at a time to assess the influence of controlling for these characteristics on the estimated coefficient on *Euro share*.

Although conditioning on the regressors from Table 2 reduces the sample by about 50%, the results hold. In no case does expanding the conditioning information set in Table 3c of Tables 3a and 3b cause the coefficient on *Euro share* to become statistically insignificant compared to the results reported and already discussed in the Table 3a and 3b analyses.[9]

More generally, there might be concerns that other features about the continents of the former colonies might account for the findings. Thus, in Table 3c, we include continent dummy variables to assess the robustness of the results. We use the UN coding and definitions of populated

---

[9] The results in Tables 3a – 4c are also robust to including a dummy variable for whether the country is a former colony, i.e., *Ex-colony* as defined in the Data Appendix and constructed by AJR (2001).

continents—Europe, Africa, Asia, the Americas, and Oceania. The UN coding of continents seems to us the least susceptible to later, possibly endogenous, splits of regions such as North Africa and Sub-Saharan Africa, or North and South America.[10] (Using the continent fixed effects as conditioning variables has the advantage that they do not reduce the sample. Since analyses exclude European countries, we do not include a dummy variable for the continent of Europe. When re-doing Tables 3a and 3b with continent dummy variables, we confirm the paper's findings.[11]

3.3 Is it Europeans during Colonization or Europeans today?

If *Euro share* proxies for the proportion of the population today that is of European descent, then it would be inappropriate to interpret the results in Tables 3a and 3b on *Euro share* as reflecting the enduring impact of Europeans during the colonization period on economic development today. Indeed, Figure 3 shows that there is a positive association between colonial *Euro share* and European share in 2000, which we call *Euro 2000 P-W* and is take from Putterman and Weil (2010). To assess the strength of the independent relationship between the level of economic development today and the European share of the population during the colonial era, we therefore control for the proportion of the population today that is of European descent.

In Tables 4a and 4b, we find that all of the results on the positive relationship between *Current income* and *Euro share* hold even when controlling for the current proportion of the population of European descent. Tables 4a and 4b are the same as Tables 3a and 3b except that the regressions also condition on the proportion of the population of European descent today. As shown, across the different regression specifications, samples, and control variables, the earlier results hold.

---

[10] Ashraf and Galor (2011, 2013) likewise say "a single continent dummy is used to represent the Americas, which is natural given the historical period examined."

[11] Tables 1 and 2 are also robust to including an Africa dummy. When controlling for *Indigenous Mortality*, we do not include fixed effects for the continents of the Americas or Oceania because *Indigenous Mortality* is equivalent to the sum of those two continent fixed effects.

Figure 3 helps in understanding that the proportion of Europeans during the colonization period is more strongly associated with current economic development than the proportion of the population today that is of European descent. Examining the scatter plot in Figure 3, consider three groups of countries: (1) countries in which *Euro share* was high both in colonial times and today (e.g. North America), (2) countries in which *Euro share* was low both in colonial times and today (e.g. South Africa), and (3) countries in which *Euro share* today is much higher than it was in colonial times (e.g. some Central and South American countries). If colonial *Euro share* did not have an independent link with incomes today, then we would expect group (3)'s income to be more like group (1)'s income. But, this is not what we find. In contrast, if colonial *Euro share* does matter independently for income today, then we would expect group (3)'s income to have lower income than group (1) and to have similar income to group (2). This is what we observe. The proportion of Europeans during the colonization period is independently associated with economic development today.

In Table 4c, we add the same conditioning variables to the Table 4a and 4b regressions that were used in Table 3c of Tables 3a and 3b. That is, in Table 4c, we not only control for the proportion of the population of European descent in 2000, we also include two different sets of conditioning information: continent dummy variables and the control variables *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology, and Settler mortality*. We then introduce *British legal origin, Education, Independence, Government quality, and Ethnicity* one at a time to assess the sensitivity of the estimated coefficient on *Euro share*. As shown, all of the results hold when conditioning on the continent dummy variables for both the full sample of countries and for the sample that eliminates settler colonies, i.e., former colonies with *Euro share* greater than 12.5%. However, for the full sample of countries, *Euro Share* is not robust to adding the full set of control variables from Table 2. But, when eliminating the few countries with *Euro share* greater than 12.5%, the results hold when including the full set of control variables, as shown Panel B of Table 4b. In particular, when eliminating "Neo-Europes" and when controlling for *Euro 2000 P-W* and the large set of control variables from Tables 1 and 2, the core results on the

relationship between *Current income* and the European share of the population during the colonization period hold.

3.4 Explaining the Reversal of Fortune

In a widely cited article, Acemoglu, Johnson, and Robinson (2002, p. 1231) document a reversal of fortune: "Among countries colonized by European powers during the past 500 years, those that were relatively rich in 1500 are now relatively poor." They proxy for the degree to which a country was relatively rich in 1500 with two indicators -- urbanization rates and population density. Using both indicators, they find a strong negative correlation between the economic success of a region before colonization and current levels of per capita income.

Acemoglu et al. (2002) argued that densely populated areas in 1500 were more likely to induce Europeans to adopt extractive institutions, and these extractive institutions stymied economic development, leading to the reversal of fortune. In particular, they hypothesized that successful areas before European colonization, as measured by population density, would attract only a few European settlers, who would establish extractive political institutions that would retard long-run growth. Acemoglu et al. (2002) also suggested a direct positive effect of indigenous population density on the productivity of extractive institutions: there was more prosperity for Europeans to tax away for themselves, and there was a large labor force to exploit in European-owned plantations and mines.

Our new data on colonial European settlement contribute to the study of the reversal of fortune in two ways. First, using actual data on colonial European settlement, we provide the first confirmation of AJR's prediction that indigenous population density was inversely associated with the share of European settlers during colonization (Table 2). Consistent with Acemoglu et al. (2002), we find that Europeans settled where there were not many other people. Second, as demonstrated in Tables 3a, 3b, 4a, and 4b, *Euro share* has a strong positive association with per capita income today. This finding is consistent with the explanation that colonial European

settlement is a key intermediating variable in explaining the reversal of fortunes:[12] lower pre-colonial population density facilitated more European settlers (as a share of total population) and these settlers brought human capital, political institutions, and other factors that fostered economic development. This explanation does not require that political institutions are the principal channel through which European settlement shaped the reversal of fortunes. It simply requires, as documented above, that *Euro share* is positively associated with subsequent economic development.

3.5 Additional Robustness tests

These results are robust to several sensitivity analyses. First, there might be concerns about the dating of *Euro share*. As discussed above, we attempt to use a date in each country's history that is at least a century after initial European contact, but at least 50 years before independence. This is both a bit arbitrary and often infeasible due to data availability. Thus, we re-do all of the analyses using three alternative ways of computing *Euro share*. We compute the average value of the share or European settlement in our data over three uniform periods: (1) 1500 – 1800, (2) 1801 – 1900, and (3) 1500 – 1900.[13] Tables 5a and 5b provide the regression results using these alternative methods for dating the share of colonial European settlement. As shown, all of the earlier results hold.

Second, there might also be concerns that (1) we assign a value of zero to *Euro share* if we find no historical sources documenting European settlement in a particular colony and (2) countries in which *Euro share* equals zero are special cases. Thus, we conducted two sensitivity tests. We simply re-did the analyses while eliminating all countries with *Euro share* equal to zero. All of the

---

[12] Chanda, Cook, and Putterman (2014) also explain the Reversal of Fortune as due to the movement of Europeans, but they use Euro 2000 P-W—the proportion of Europeans in the population in 2000, which we discuss and use above. Our story is similar to theirs except that we pinpoint the importance of the proportion of Europeans during the colonial period in accounting for the reversal.

[13] Note that the sample will include only observations with positive Euro Share for these periods. A "dog did not bark" method might be acceptable for the entire historical period to equate lack of mention of colonial settlement with zero colonial settlement. But this method is more problematic with finer breakdowns of historical sub-periods.

results hold. Next, we re-did the analyses while including a dummy variable called *Dummy for euro share equal to zero* that equals one if *Euro share* equals zero and zero otherwise. These results are reported in Table 7. As shown, all of the results on *Euro share* hold when including this dummy variable. The finding that *Dummy for euro share equal to zero* generally enters with a positive and significant coefficient in the Table 3b and 4b specifications that restrict the sample to *Euro Share* <.125) provides some suggestive evidence for the view that small European settlements had harmful effects on long-run economic development compared to no European settlement. The estimates, however, indicate that the positive effect of *Euro share* on *Current income* quickly surpasses the positive effect of the *Dummy for euro share equal to zero* on *Current income*. For instance, the estimates from column (1) of Table 7, Panel 2 indicate that a *Euro share* of greater than 0.048 is better for *Current income* than a *Euro share* of zero.

Third, we also assess whether other geographic endowments account for the findings on the relation between *Euro share* and *Current income*. In particular, we use variables from Ashraf and Galor (2013), the impact of soil quality and access to navigable waterways. For soil quality, they gauge the suitability of the soil for agriculture though measures of soil carbon density and soil pH. For navigable waterways, they use the average distance from grid cells throughout a country to the nearest ice-free coastline or sea-navigable river. When we include these measures of geographic endowments in our analyses, all of the results hold.

Fourth, there might be concerns that we do not measure the effect of European colonization compared to no colonization, and hence might not sufficiently distinguish European colonization from European settlement. This is a valid issue but difficult to study empirically, since few non-European countries completely escaped colonization or some intermediate form of European control, and there is ambiguity and disagreement as to which ones they are. For example, AJR identify the following countries as *not* former colonies: China, Iran, Japan, Korea, Liberia, Mongolia, Thailand, Turkey, and Uzbekistan. There are, however, disagreements because, for example, China had semi-colonial enclaves, Iran was a "sphere of interest" for Russia and Britain, Uzbekistan partially belonged to the Russian empire, Korea was a Japanese colony, and Liberia was

governed by descendants of American slaves. In the other direction, Ethiopia is often considered a non-colony but is classified as a colony by AJR. Nevertheless, when including a dummy variable that equals one if the country is a former colony as classified by AJR and zero otherwise, all of the paper's results hold.

4. How much development is attributable to Europeans?

In this section, we do some global development accounting to illustrate how much of development might be associated with European settlers. This exercise uses the estimated equation for *Euro share* with no controls

(2) $\quad \ln(CurrentIncome_i) = \alpha + \beta Euroshare_i + \varepsilon_i$

Next, define the counterfactual *CurrentIncome*^{CF} for every country outside of Europe by removing the European effect:

(3) $\quad CurrentIncome_i^{CF} = CurentIncome_i \cdot e^{-\beta EuroShare_i}$

Of course, $CurentIncome_i = CurrentIncome_i^{CF}$ for any country i where *Euroshare*$_i$=0.

The counterfactual population-weighted global mean is then simply the weighted mean across all non-European countries of $CurrentIncome_i^{CF}$, where P_i is population in country i, and P is total global population:

(4) $\quad \tilde{y}^{CF} = \Sigma_i \left(\frac{P_i}{P}\right) CurentIncome_i^{CF}.$

The global population-weighted per capita income $\tilde{y}$ is

(5) $\quad \tilde{y} = \Sigma_i \left(\frac{P_i}{P}\right) CurentIncome_i.$

The share of development attributed to European settlement is then $s_e = \left(\frac{\tilde{y}-\tilde{y}^{CF}}{\tilde{y}}\right).$

As an illustrative exercise, we use the sample and the coefficient from regression (1) of Table 3a, which is the simplest regression for the full sample of all countries outside of Europe. The coefficient estimate is $\beta = 3.623$.

Using the 2000 population weights, the data and estimated coefficients indicate that 40% of the development outside of Europe is associated with the share of European settlers during

colonization $\left(\frac{\tilde{y}-\tilde{y}^{CF}}{\tilde{y}}\right)$. We repeat our frequent caveat that global per capita income is not a welfare measure, especially in light of the history of European exploitation of non-Europeans.

As an exercise in positive analysis, however, it is striking how much of global development today is associated with the migration and settlement of Europeans during the colonial era (not even considering the development of Europe itself). It is even more striking that this large average income outcome in a non-European world today of over five billion is associated with the migration of only six million European settlers in colonial times.

5. Conclusions

The previous literature was correct to focus on colonial settlement by Europeans as one of the pivotal events in the history of economic development. In this paper, we provide the first direct evidence that the proportion of Europeans during colonization is strongly and positively associated with the level of economic development today. These findings are robust to using different subsamples of countries, controlling for an array of country characteristics, and conditioning on the current proportion of the population of European descent.

These results relate to theories of the origins of the divergent paths of economic development followed since Europeans colonization. ES and AJR stress that when endowments lead to the formation of settler colonies, this produced more egalitarian, enduring political institutions that fostered long-run economic development. And, ES and GLLS emphasize that Europeans brought human capital that slowly disseminated to the population at large and boosted economic development. The results presented in this paper are consistent with both of these effects: former colonies with larger colonial European settlements have much higher levels of economic development today than former colonies that had a smaller proportion of Europeans during the colonial period. Our results also paint a positive picture of minority colonial European settlements about which the previous literature was ambiguous. Specifically, the estimates indicate that once European settlement is above 4.8%, the small colonial European settlements have a positive effect

on development today compared to no colonial European settlement. This is suggestive that any adverse effects arising from the extractive institutions created by small colonial European settlements were more than offset by other things that Europeans brought during colonization, such as human capital, technology, familiarity with global markets, and institutions, which had lasting, positive effects on economic development.

Bibliography

Acemoglu, D., S. Johnson and J. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review, 91 (5): 1369-1401.

Acemoglu, D., S. Johnson and J. Robinson (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution", Quarterly Journal of Economics 117(4): 1231-1294.

Acemoglu, D., S. Johnson and J. Robinson (2005), "Institutions as the Fundamental Cause of Long-Run Growth", in Aghion and Durlauf, editors, Handbook of Economic Growth, Amsterdam: North Holland.

Acemoglu, D. and J. Robinson, (2008), "Persistence of Elites, Power and Institutions", American Economic Review, 98 (1): 267-293.

Acemoglu, D. and J. Robinson (2012), Why Nations Fail: The Origins of Power, Prosperity, and Poverty. Crown Publishers: New York,

Albouy, D. (2012), "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment", American Economic Review 102(6): 3059-3076.

Ashraf, Q. and O. Galor (2011), "Dynamics and Stagnation in the Malthusian Epoch." American Economic Review, 101(5): 2003-41.

Ashraf, Q. and O. Galor (2013), "The 'Out of Africa' Hypotheses, Human Genetic Diversity, and Comparative Economic Development", American Economic Review 103(1): 1-46.

Banerjee, D. and L. Iyer (2005), "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India", American Economic Review, 95(4): 1190-1213.

Beck, T., A. Demirguc-Kunt, and R. Levine (2003), "Law, Endowments, and Finance." Journal of Financial Economics, 70: 137-181.

Bisin, A. G. Topa and T. Verdier (2004), "An Empirical Analysis of Religious Homogamy and Socialization in the U.S, Journal of Political Economy, 112(3): 615-64.

Bisin, A. and T. Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits", Quarterly Journal of Economics, 115(3): 955-988.

Bockstette, V., A. Chanda, and L. Putterman (2002), "States and Markets: the Advantage of an Early Start," Journal of Economic Growth, 7: 347-369.

Castro, Baron (1942) "La poblacion de El Salvador: estudio acerca de su desenvolvimiento desde la epoca prehispanica hasta nuestros dias" (The population of El Salvador: study of their

development from pre-Hispanic era to the present day). <u>Spain: Consejo Superior de Investigaciones Cientificas.</u>

Chanda, Areendam, C. Justin Cook, and Louis Putterman. 2014. "Persistence of Fortune: Accounting for Population Movements, There Was No Post-Columbian Reversal." American Economic Journal: Macroeconomics, 6(3): 1-28.

Comin, D., W. Easterly, and E. Gong (2010), "Was the Wealth of Nations Determined in 1000 B.C.?" American Economic Journal: Macroeconomics, 2(3): 65–97.

Diamond, J. (1997), <u>Guns, Germs, and Steel: The Fates of Human Societies</u>, New York: Norton.

Easterly, W. and R. Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions." Quarterly Journal of Economics, 112 (4): 1203-1250.

Easterly, W. and R. Levine (2003), "Tropics, germs, and crops: the role of endowments in economic development." Journal of Monetary Economics, 50 (1): 3-40.

Easterly, W. (2007), "Inequality does cause underdevelopment." Journal of Development Economics, 84 (2): 755-776.

Engerman, S. and K. Sokoloff (1997), "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States" in How Latin America Fell Behind, in Haber, editor, Stanford: Stanford University Press: 260–304

Fernandez, R. and A. Fogli (2009), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility", American Economic Journal: Macroeconomics 1 (1): 146–77.

Galor, G. (2011), Unified Growth Theory, Princeton NJ: Princeton University Press.

Glaeser, E., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2004), "Do Institutions Cause Growth?", Journal of Economic Growth, 9 (3): 271-303.

Guiso, L. P. Sapienza, L. Zingales (2010), "Civic Capital as the Missing Link", in Jess Benhabib, Alberto Bisin and Matthew Jackson (editors) Handbook of Social Economics, Amsterdam: North Holland.

Guiso, L. P. Sapienza, L. Zingales (2013), Long-term Persistence, Einaudi Institute for Economics and Finance Working Paper 23/13, September 2013

Haber, S., (2014) "Where does democracy thrive? Climate, technology, and the evolution of economic and political institutions," Stanford University mimeo.

Hibbs, D.A. Jr. and O. Olsson , (2004) "Geography, biogeography, and why some countries are rich and others are poor" Proceedings of the National Academy of Sciences 101 (10): 3715-372.

Hofstede, G. (2001), Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations, 2nd Edition Thousand Oaks CA: Sage Publications.

Hofstede, G and G. Hofstede (2005), "Cultures and Organizations: Software of the Mind" 2nd Edition, New York: McGraw-Hill.

Karlen, A. (1995), Man and Microbes: Disease and Plagues in History and Modern Times New York: Simon & Schuster.

Kaufman, D. A. Kraay, and P. Zoido-Lobaton (2002) "Governance Matters II: Updated Indicators for 2000-01", World Bank Policy Research Working Paper 2772.

Kleibergen, F., and R. Paap (2006), "Generalized reduced rank tests using the singular value decomposition". Journal of Econometrics 133: 97-126.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S.E. Sachs and J. Sachs (2004), "A Global Index Representing the Stability of Malaria Transmission", American Journal of Tropical Medicine and Hygiene 70 (5): 486-498.

La Porta, R., F. Lopez-de-Silanes, and A. Shleifer (2008), "The Economic Consequences of Legal Origins," Journal of Economic Literature, 46: 285–332.

Licht, A., C. Goldschmidt, and S. Schwartz (2007), "Culture rules: The foundations of the rule of law and other norms of governance", Journal of Comparative Economics, 35: 659–688.

Madison A. (2006), "The World Economy", Organization for Economic Co-operation and Development. Development Centre, Vol. 1-2.

McAlister L. (1984), "Spain and Portugal in the New World, 1492-1700", USA: U of Minnesota Press.

McNeil W. (1976), Plagues and Peoples New York: Anchor Press.

McEvedy C. and R. Jones (1978), Atlas of World Population History, New York: Facts on File.

Michalopoulos, S, and E. Papaioannou. "Pre-Colonial Ethnic Institutions and Contemporary African Development." Econometrica 81.1 (2013): 113-152.

Michalopoulos, S., and E. Papaioannou. "National Institutions and Subnational Development in Africa." The Quarterly Journal of Economics 129.1 (2014): 151-213.

Miguel, E, and R. Fisman (2007), "Corruption, Norms and Legal Enforcement: Evidence from Diplomatic Parking Tickets," Journal of Political Economy, 115(6): 1020–1048.

Miguel, E., S. Saiegh, and S. Satyanath (2011), "Civil War Exposure and Violence," Economics and Politics 23 (1), March 2011.

North D. (1990), "Institutions, Institutional Change, and Economic Performance" UK: Cambridge University Press.

Nunn, N. (2009), "The Importance of History for Economic Development," Annual Review of Economics, 1(1): 65-92.

Nunn, N. (2008), "The Long-Term Effects of Africa's Slave Trades," Quarterly Journal of Economics, 123 (1): 139–176.

Oldstone M. (1998), Viruses, Plagues, and History, USA: Oxford University Press.

Olsson, Ola. "On the democratic legacy of colonialism." Journal of Comparative Economics 37, no. 4 (2009): 534-551.

Persson, T. and G. Tabellini (2010), "Democratic Capital: The Nexus of Political and Economic Change", American Economic Journal: Macroeconomics, 1(2): 88–126.

Putterman, L. and D. Weil (2010), "Post-1500 Population Flows and the Long Run Determinants of Economic Growth and Inequality", Quarterly Journal of Economics, 125(4): 1627-1682

Rosentblat A. (1954), "La población indígena y el mestizaje en América" (The indigenous population and racial mixing in America), Buenos Aires: Editorial Nova

Sokoloff, K. and S. Engerman, History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World, Journal of Economic Perspectives 14 (3), Summer 2000, 217–232

Spolaore E. and R. Wacziarg, R. (2009), "The Diffusion of Development", Quarterly Journal of Economics, 124 (2): 469-529.

Spolaore E. and R. Wacziarg, R. "How Deep Are the Roots of Economic Development?" Journal of Economic Literature 2013, 51(2): 325–369.

Tabellini, G. 2010. "Culture and Institutions: Economic Development in the Regions of Europe." Journal of the European Economic Association 8 (4): 677–716.

Tabellini, G. (2008), "The Scope of Cooperation: Values and Incentives," Quarterly Journal of Economics, 123: 905–950.

<u>Data Appendix</u>

This appendix describes the construction of the dataset on the European share of the population in countries around the world during colonization. The primary goal is to define what we did, so that the numbers are transparent and replicable. At the end of the appendix, we list some of the problems that we encountered in constructing the database, and the difficulties that we faced in choosing which years to use in defining a country's "European share of the population during colonization." The dataset and other key information are contained in the excel workbook titled "Appendix Europeans," which is available on request. In this Appendix when discussing details of the dataset, we refer to specific worksheets within this workbook. We also have created Stata do files that replicate the results in the tables; these files and the full dataset is also available on request.

Data sources and definitions

We primarily rely on 46 sources, which are listed in the worksheet titled "bibliography" and the worksheet titled "web." Many of these are scholarly books about particular regions or countries and some are atlases. As a few examples, Robert Wells wrote *The Population of the British Colonies in America before 1776*; Simeon Ominde wrote *The Population of Kenya, Tanzania, and Uganda*; and, McEvedy and Jones assembled *Atlas of World Population History*. We also use primary sources (such as national and colonial censuses) to both check these sources and to expand the number of countries and data points. Besides the books and official documents listed in the worksheet "bibliography," some datasets are provided online. We list these in the worksheet "web."

In terms of defining "European settlers," we strive in collecting the data to identify Europeans as people from the geographic region of Europe; it is NOT a racial or ethnic description. So, some observers might consider the populations of some countries outside of Europe as racially or ethnically equivalent to Europeans, but that is irrelevant to us. We are only concerned with resettlement from Europe to outside Europe, defined geographically. Furthermore, in assembling

the data on settlers, we strive to exclude colonial officials or business people that are temporarily stationed abroad; we strive to only include Europeans who permanently resettle outside of Europe.

Data: By country, year, and source

For each country, we provide an entry for each year for which we found information on the European share of the population. For each data point, we provide the source of the information (including the page number) and brief notes about the data, whenever relevant. Some of these notes are important. For example, the 1744 and 1778 values for Argentina are based only on the population around Buenos Aires, for the country as a whole. Similarly, one of the values for Ecuador in 1781 measures only the population around Quito. These notes, the data, and the sources of each data point are listed in the worksheet, "country_year_source."

For example, Lyle N. McAlister, in Spain and Portugal in the New World, 1492 -1700, (University of Minnesota Press, 1984) provides data on page 131 on the population of Argentina in 1570. He notes that there are 2000 whites, 4000 blacks, 300,000 "others" in Argentina. Since whites are typically used to describe people of European descent, we calculate European share in Argentina in 1570 as 0.0065.

In some cases, the data sources provide a range of years (rather than a single year) corresponding to data on the share of the population that is European. For example, one of the observations on Mexico is listed as 1568-1570 in the underlying data. In these cases, we choose the average of the range and enter this as the year for the observation. Thus, for the Mexico example, we choose the year 1569 when entering this data. All of these cases are separately identified in the worksheet "periods." This has no bearing on our analyses, but might be relevant for others that use these data.

In a few cases, we found two data sources that provide information on the same year (or range of years) for the same country. In some cases, these two data sources agree, in which case we simply report both observations within the worksheet "country_year_sources." In a few cases, the data sources give different numbers for the share of Europeans in a country in a given year. In this

case, we report both numbers in the worksheet "country_year_sources" and use the average of the two observations when constructing the data on which we conduct our analyses.

Data: Share of Europeans used in the analyses

From these data scattered over many years since the 16th century, we construct several measures of the share of Europeans during colonization for each country, where we have one measure per country. To do this, we arrange the data in a manner that is amenable for the construction of a single measure of the share of Europeans during colonization for each country. In the worksheet "euro share," each row is a country. The columns provide possible data entries for many years running from 1540 (which is our first observation, for Chile) through to the late 20th century. We do not include all years as column headings; rather, we only include years for which we have at least one non-missing value for one country.

First, we construct simple, objective measures that average the value of each country over particular periods. For example, we average each country's entries over the period from 1500 to 1801; and, we average values over the period from 1700 to 1950. These measures are provided in the worksheet titled "euro share." Other researchers can obviously take these data and use whichever periods they find most appropriate.

These simple, objective measures for computing the share of Europeans during colonization, however, have some limitations. Specifically, averaging over uniform time periods for all countries might not create accurate measures for each particular country of the proportion of the population that is European during a colony's formative period—the period when a colony was creating an initial set of (potentially enduring) political, educational, and cultural institutions. We fully recognize that there is not a precise definition of "the" formative period of colonization. Nevertheless, influential studies of comparative economic development emphasize the potential role of Europeans during a colony's history when it establishes major institutional norms. This motivates our efforts to give empirical substance to this amorphous notion. From this perspective, using the

European share of the population of Mexico in 1650 might be more appropriate than using the share in 1850, but using the European share of the population in 1650 in some parts of Sub-Saharan Africa (or other parts of the world) might be inappropriate because European colonization evolved differently there. Thus, we face a challenging goal: account for these historical differences in the timing and process of colonization to construct a more conceptually useful measure of "euro share" for each country. We face this challenge while operating under severe data constraints.

Thus, the second method for constructing a measure of each country's European share of the population during the formative years of colonization attempts to select the best year, or range of years, given the particulars of the country and data availability. We would like a date as early as possible after initial European colonization to use European settlement as an initial historical condition affecting subsequent developments. At the same time, we can't pick a date that is too early after the start of European colonization. It was only after some process of conquest, disease control, and building rudimentary colonial infrastructure that it becomes possible to speak of a European community that might influence economic, political, and cultural conditions. Given these considerations, it would not make sense to use a uniform date across all colonies.

Thus, we formulated the following "guidelines." Subject to data limitations, we tried to constrain the timing of the European measure to be at least a century after initial European contact. Furthermore, we tried to choose a date that was at least 50 years before independence to measure the colonial period. Finally, if there were a few measures close together, we took the average.

In the worksheet "euro share," we provide a measure of each country's European share of the population – *Euro share*—that represents our assessment of the best year, or range of years, for measuring European share during the formative years of colonization for each particular country, where this assessment is almost always done subject to extreme data constraints. When using a range of years, we average to compute *Euro share*. The worksheet also provides the year, or range of years, used to compute *Euro share*. This second method is neither simple nor fully objective. Though we do our best to follow the guidelines outlined above, data limitations and the idiosyncrasies of colonial histories make things complex and subjective. Nevertheless, we believe

Euro share is a more accurate representation of the share of Europeans during the formative years of colonization that simply averaging over a uniform time period for all former colonies.

Though the excel spreadsheet "euro share" provides the details for each country, it is valuable to illustrate some of the constraints that we face and the choices that we made. For much of Latin America, Angel Rosentblat (1954) provides detailed estimates of the composition of the population in 1650. We have used these estimates when available. In many countries, the next available observation is not until a century (or more) later. For example, after 1650, the next observation is not until 1798 in Brazil, 1940 in Bolivia, 1777 in Mexico, and 1744 in Argentina. For some of these countries, we have earlier population estimates that are reported in the excel file. For example, we have observations in 1570 for Argentina, Bolivia, Brazil, and Mexico. But, following the guidelines sketched above, we determined that this was too early after Europeans first arrived.

There are other problems, some of which force us to break with the "guidelines" sketched above. For example, the first number that we have for the United States is the 1790 census, which is obviously not fifty years before the country became independent. Similarly, we do not have data on the composition of the population before 1950 for several countries in Africa, including the Democratic Republic of Congo, Djibouti, Gabon, Sao Tome and Principe, Senegal, Tunisia, and Rwanda. Jamaica and El Salvador provide some particular challenges. For Jamaica, the numbers on European share of the population show considerable variability over the period from 1570 to 1673; but then, the numbers (provided by various sources) are quite consistent from 1700 through 1943. So, to compute *Euro share*, we take the average over the period from 1700-1750. For El Salvador, there is one extremely large observation in 1796 (0.48) that deviates from other estimates in nearby years (e.g., 0.03 in 1807) provided by the same source (Baron Castro, 1942). Since (1) the estimates for *euro share* over the entire period with available data from 1551 to 1950 are reasonably constant except for this one observation 1796 and (2) there seems to be a change in the definition of "white" for this particular year, we decided not to include this observation in the "euro share" worksheet. For El Salvador, therefore, we compute the average over the period from 1551-1807, excluding this 1796 outlier because of the change in definition.

Data: Countries in which Europeans did not settle

There are many countries in which Europeans did not settle to any appreciable degree. In these countries, unsurprisingly, we do not find historical sources documenting the share of Europeans during the colonial period. Thus, we face a problem: We do not want eliminate these countries from the sample when we know that European settlers were not a material part of their history, but we do not have documentation to that effect. Thus, although we cannot strictly prove that there were no Europeans, available evidence suggests that Europeans did not settle everywhere and we can incorporate this information into our analyses.[14]

We follow the following procedure. We conduct a worldwide search for colonial data on European settlement. Besides the sources listed in the workbook, we examine many additional sources in search of data. When we fail to find any mention of European settlement in any of these sources for a given nation, we coded that country as having zero European settlement. This procedure runs the risk of biasing downward European settlement for these countries. But, colonial histories seem unlikely to ignore material European settlement. We confirm our data using information from AJR (2001) on the European share of the population in 1900.

Furthermore, Figure A1, depict the scatter plot of our measure of colonial European settlement and the measure of the European share of the population in 1900 provided by AJR (2001). One of the key differences between the two measures, as one would expect, is Latin America. Many Europeans arrived in Argentina, Brazil, Chile, etc. after the colonial period.

---

[14] There are many complications. For example, India obviously had a colonial administrative staff from the U.K. Our measure of Euro Share does not include such staff, as it only counts those identified as coming for permanent settlement as private individuals rather than to serve a term in the colonial administration. In the excel file containing our data, we provide the source and page from which we obtained each data point.

Problems: A brief discussion of a few of the problems

There are many challenges associated with constructing this database on the European share of the population during the period of colonization. First, although colonists did document the number of Europeans in the total population at various points during colonization, the processes and periodicities were not standardized. We just do not know if the same colonial power used the same methods in different countries in different years, not to mention differences across European powers. For example, different colonial powers probably used different methods to estimate population numbers. In a census-based method, there could be an undercount of non-European populations, which would bias population numbers downward and European shares upward. In a sampling methodology, there is zero expected bias only if the samplings were random. Unfortunately, we have almost no information on methods followed to get these population numbers. Put simply, we do not have a continuous time series of data collected using similar measures by a centralized coordinating entity.

Second, most of the cells in the data running from the 16th century to the 20th century are empty. This means that for many countries we cannot get measures of the share of Europeans in the population within decades of the years that we would ideally want to measure *Euro share*. Although most countries do not experience huge changes in the share of Europeans, some experience substantial changes, so this is another challenge facing the construction of the database.

Third, the basic conceptual predictions about the role of Europeans during colonization and their enduring influence on economic development do not provide a concrete definition of when to measure the share of Europeans during colonization. We try to measure the share of Europeans about a century after the start of colonization and 50 years before independence, but this simply represents the articulation of a hopefully helpful empirical guideline. Without ignoring these—and other—challenges, we constructed the database and use it to provide empirical evidence on the relationship between the share of Europeans during colonization and comparative economic development.

Table A: Variable Definitions

This table provides definitions and sources of the major variables used in the analyses.

	Definition	Source
Euro share	Proportion of Europeans in colonial population	Constructed by authors. See appendix for details.
Euro 2000 P-W	Proportion of Europeans in 2000 population. Constructed from Putterman and Weil's (2010) migration database by (for each country in the sample) adding the proportion of ancestors coming from each European country.	Putterman and Weil (2010)
Current income	Ln average of GDP per capita over 1995-2005 (PPP, Constant 2005 international $)	World Bank World Development Indicators
Population density 1500	Log population per square km in 1500	AJR (2002)
Indigenous mortality	Dummy variable reflecting high rates of indigenous mortality from European diseases. De Facto turned out to be equal to one for Americas and Oceania, and zero otherwise.	McEvedy and Jones (1978), McNeil (1076), Karlen (1995), Oldstone (1998).
Latitude	The absolute value of latitude in degrees, divided by 90 to be between 0 and 1.	CIA World Factbook
Malaria ecology	An index of the stability of malaria transmission based biological characteristics of mosquitos such as the proportion of blood meals taken from human hosts, daily survival of the mosquito, and duration of the transmission season and of extrinsic incubation.	Kiszewski et al. (2004)
Settler mortality	Log of potential settler mortality, measured in terms of deaths per annum per 1,000 "mean strength" (constant population).	AJR (2001)

Biogeography	The first principal component of log of number of native plants species and log number of native animals species, where plants are defined as "storable annual or perennial wild grasses with a mean kernel weight exceeding 10 mg (ancestors of domestic cereals such as wheat, rice, corn, and barley)" and animals denotes the number of "species of wild terrestrial mammalian herbivores and omnivores weighing greater than 45 kg that are believed to have been domesticated prehistorically in various regions of the world."	Hibbs and Olsson (2004), later updated by them to add observations
British legal origin	A dummy variable indicating British legal origin	La Porta et al. (1999)
Education	Average rate of gross secondary school enrollment from 1995-2005	World Bank World Development Indicators
Independence	The fraction of years since 1776 that a country has been independent	Easterly and Levine (1997)
Government quality	The first principal component of the six governance indicators from the 2002 vintage of Kaufman et al.	Kaufman et al (2002)
Ethnicity	An index of ethnic diversity (updated).	Easterly and Levine (1997)
Precious metals	Dummy =1 if country produced any gold or silver in 1999, 0 otherwise	Easterly (2007), original source World Bureau of Metal Statistics
London	The distance of the country from London	Constructed by authors, average of oceanic distance from London to all seaports in country
Ex-colony	A dummy variable that equals one if the country is a former colony as classified by AJR and zero otherwise.	AJR (2001)
Dummy for euro share equal to zero	A dummy variable that equals one if *Euro share* equals 0 and zero otherwise.	Constructed by authors
Soil suitability	Gauges the suitability of the soil for agriculture through measures of soil carbon density and soil pH.	Ashraf and Galor (2013)
Distance to waterways	The average distance from grid cells throughout a country to the nearest ice-free coastline or sea-navigable river.	Ashraf and Galor (2013)

Table 1: Descriptive Statistics

This table provides descriptive statistics of the major variables used in the regression analyses. Table A provides variable definitions.

	Observations	Mean	Standard Deviation	Min	Max	Median
Euro share	129	0.07	0.17	0.00	0.90	0.00
Euro 2000 P-W	115	0.13	0.24	0.00	0.90	0.00
Current income	123	8.18	1.25	5.48	11.04	8.13
Population density 1500	94	0.50	1.52	-3.83	4.61	0.42
Indigenous mortality	127	0.29	0.46	0.00	1.00	0.00
Latitude	129	0.20	0.12	0.01	0.67	0.18
Malaria ecology	114	5.13	7.28	0.00	31.55	1.44
Settler mortality	80	4.70	1.20	2.15	7.99	4.51
Biogeography	82	0.01	1.31	-1.02	3.79	-0.65
British legal origin	129	0.40	0.49	0.00	1.00	0.00
Education	122	57.50	30.67	5.60	152.84	60.30
Independence	89	0.31	0.34	0.00	1.00	0.10
Government quality	128	-0.47	1.93	-4.91	4.62	-0.62
Ethnicity	115	0.38	0.32	0.00	1.00	0.33
Precious metals	130	0.28	0.45	0.00	1.00	0.00
London	122	5,410.37	2,342.12	1,381.39	11,769.29	4802
Ex-colony	112	0.92	0.27	0.00	1.00	1.00
Soil suitability	108	0.53	0.19	0.16	.95	0.51
Distance to waterways	108	0.34	0.38	0.01	1.72	0.21

Table 2: Determinants of colonial European settlement

This table presents regression results concerning which factors shaped European settlement during colonization. The sample is non-European countries. The dependent variable is *Euro share,* which gives the proportion of Europeans in the colonial population. *Population density 1500* is computed as the logarithm of population density in 1500. *Indigenous mortality* takes a value of one if the region experienced a large drop in the indigenous population from diseases brought by Europeans, and zero otherwise. *Latitude* measures the absolute value of the distance of the colony from the equator. *Precious metals* is an indicator of whether the region has valuable minerals that may have affected European settlement. *London* measures the distance of the colony from London. *Biogeography* is an index of the prehistoric availability of storable crops and domesticable animals, where large values signify higher availability. *Malaria ecology* is an ecologically-based spatial index of the stability of malaria transmission in a region, where larger values signify a greater propensity for malaria transmission. *Settler mortality* equals historical deaths per annum per 1,000 European settlers. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Euro Share				
Population density 1500	-0.0376**	-0.0272***	-0.0277***	-0.0275***	-0.0224*	-0.0270***	-0.0323**	-0.0177
	(0.02)	(0.01)	(0.01)	(0.01)	(0.05)	(0.01)	(0.01)	(0.24)
Indigenous mortality	0.140***	0.136***	0.139***	0.131***	0.104***	0.139***	0.0818**	0.0243
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	(0.65)
Latitude		0.704***	0.715***	0.728***	0.748***	0.728***	0.707***	0.767***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Precious metals			-0.0128					0.0177
			(0.66)					(0.62)
London				-2.06e-06				-5.71e-06
				(0.78)				(0.56)
Biogeography					-0.0184			-0.0481**
					(0.12)			(0.02)
Malaria ecology						0.00116		5.56e-05
						(0.37)		(0.98)
Settler mortality							-0.0146	-0.0489**
							(0.22)	(0.03)
Number of observations	94	94	94	90	71	88	72	57
R-squared	0.352	0.541	0.542	0.544	0.586	0.539	0.580	0.675

Table 3a: Current income and colonial European settlement

This table presents regression results assessing the relationship between *Euro share* and *Current income*. The sample is non-European countries. The dependent variable is C*urrent income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Current Income				
Euro share	3.623***	3.626***	0.621	3.453***	0.511	3.437***	3.089***
	(0.00)	(0.00)	(0.23)	(0.00)	(0.31)	(0.00)	(0.00)
British legal origin		-0.00240					0.144
		(0.99)					(0.52)
Education			0.0309***				
			(0.00)				
Independence				0.380			0.344
				(0.29)			(0.33)
Government quality					0.429***		
					(0.00)		
Ethnicity						-1.341***	-1.335***
						(0.00)	(0.00)
Observations	123	123	119	113	123	111	106
R-squared	0.166	0.166	0.638	0.186	0.449	0.374	0.388

Table 3b: Current income and colonial European settlement, *Euro share* < 12.5%

This table presents regression results assessing the relationship between *Euro share* and *Current income*. The sample is non-European countries with *Euro share* values of less than 0.125. The dependent variable is *Current income,* computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Current Income			
Euro share	8.378***	8.401***	-0.904	8.093***	3.612	9.846***	8.987***
	(0.00)	(0.00)	(0.69)	(0.00)	(0.14)	(0.00)	(0.00)
British legal origin		-0.0365					0.0831
		(0.88)					(0.72)
Education			0.0326***				
			(0.00)				
Independence				0.427			0.419
				(0.29)			(0.28)
Government quality					0.427***		
					(0.00)		
Ethnicity						-1.212***	-1.161***
						(0.00)	(0.00)
Observations	110	110	108	100	110	98	93
R-squared	0.047	0.047	0.600	0.065	0.361	0.244	0.260

Table 3c: Current income and colonial European settlement, Sensitivity analyses

The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *Legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. The other variables (*Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology, Settler mortality*, and the continent dummy variables) are defined in the text and Table A. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent:			*Current Income*			
Additional Control:		*Legal origin*	*Education*	*Independence*	*Government quality*	*Ethnicity*
Panel 1: Adding *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology,* and *Settler mortality*						
Table 3a sample of Non-European countries						
Euro share	1.698**	1.916**	0.673	1.766**	0.0303	1.607**
	(0.03)	(0.03)	(0.34)	(0.02)	(0.96)	(0.04)
Observations	63	63	61	62	63	63
R-squared	0.690	0.693	0.836	0.700	0.792	0.705
Panel 2: Adding *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology,* and *Settler mortality*						
Table 3b sample of Non-European countries with Euro share < 0.125						
Euro share	15.22***	15.44***	4.305	15.89***	10.29***	15.47***
	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.00)
Observations	55	55	53	54	55	55
R-squared	0.655	0.659	0.804	0.671	0.769	0.683
Panel 3: Adding *Continent fixed effects for Africa, the Americas, Asia, and Oceania*						
Table 3a sample of Non-European countries						
Euro share	3.092***	3.109***	1.110***	3.163***	0.724	2.938***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	(0.00)
Observations	123	123	119	113	123	111
R-squared	0.381	0.381	0.663	0.415	0.561	0.439
Panel 4: Adding *Continent fixed effects for Africa, the Americas, Asia, and Oceania*						
Table 3b sample of Non-European countries with Euro share < 0.125						
Euro share	13.17***	13.50***	1.246	12.69***	7.397**	12.42***
	(0.00)	(0.00)	(0.60)	(0.00)	(0.01)	(0.00)
Observations	110	110	108	100	110	98
R-squared	0.313	0.315	0.618	0.342	0.498	0.326

Table 4a: Current income and colonial European settlement, controlling for current proportion of population of European descent

This table presents regression results assessing the relationship between *Current income* and *Euro share* while controlling for the current proportion of the population that is of European descent. The sample is non-European countries. The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro 2000 P-W* is the proportion of Europeans in the 2000 population (using Putterman and Weil's (2010) migration database). *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *,**, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Current Income			
Euro share	1.905***	1.677***	0.412	2.064***	-0.380	2.104***	1.953***
	(0.00)	(0.00)	(0.48)	(0.00)	(0.55)	(0.00)	(0.00)
Euro 2000 P-W	1.358***	1.476***	0.162	1.100*	0.904**	1.115***	0.901
	(0.00)	(0.00)	(0.66)	(0.08)	(0.02)	(0.01)	(0.13)
British legal origin		0.129					0.203
		(0.62)					(0.40)
Education			0.0316***				
			(0.00)				
Independence				0.319			0.355
				(0.53)			(0.44)
Government quality					0.402***		
					(0.00)		
Ethnicity						-1.120***	-1.146***
						(0.00)	(0.00)
Observations	112	112	110	102	112	102	97
R-squared	0.187	0.190	0.638	0.209	0.435	0.374	0.386

Table 4b: Current income and colonial European settlement, controlling for current proportion of population of European descent, Euro share < 12.5%

This table presents regression results assessing the relationship between *Current income* and *Euro share* while controlling for the current proportion of the population that is of European descent. The sample is non-European countries with *Euro share* values of less than 0.125. The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro 2000 P-W* is the proportion of Europeans in the 2000 population (using Putterman and Weil's (2010) migration database). *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *,**, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Current Income			
Euro share	6.455**	6.087**	-1.709	7.471**	2.329	8.702***	8.649***
	(0.02)	(0.04)	(0.51)	(0.01)	(0.46)	(0.00)	(0.01)
Euro 2000 P-W	1.026**	1.136**	0.287	0.621	0.757	0.701	0.350
	(0.02)	(0.03)	(0.54)	(0.43)	(0.14)	(0.13)	(0.64)
British legal origin		0.102					0.162
		(0.71)					(0.51)
Education			0.0333***				
			(0.00)				
Independence				0.386			0.443
				(0.49)			(0.38)
Government quality					0.408***		
					(0.00)		
Ethnicity						-1.094***	-1.085***
						(0.00)	(0.00)
Observations	104	104	102	94	104	94	89
R-squared	0.070	0.072	0.599	0.089	0.356	0.251	0.267

Table 4c: Current income and colonial European settlement, controlling for current proportion of population of European descent, Sensitivity analyses

The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *Euro 2000 P-W* is the proportion of Europeans in the 2000 population (using Putterman and Weil's (2010) migration database). *Legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. The other variables (*Indigenous mortality*, *Latitude*, *Precious metals*, *London, Biogeography, Malaria ecology, Settler mortality*, and the continent dummy variables) are defined in the text and Table A. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *,**, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

Dependent:	(1)	(2)	(3)	(4)	(5)	(6)
			Current Income			
Additional Control:		*British legal origin*	*Education*	*Independence*	*Government quality*	*Ethnicity*

Panel 1: Adding *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology,* and *Settler mortality*

Table 4a sample of Non-European countries

	(1)	(2)	(3)	(4)	(5)	(6)
Euro share	0.748	0.830	0.470	0.852	-0.345	0.821
	(0.31)	(0.31)	(0.48)	(0.23)	(0.63)	(0.27)
Euro 2000 P-W	1.642**	1.603**	0.413	1.500**	0.860	1.409**
	(0.01)	(0.02)	(0.40)	(0.04)	(0.11)	(0.04)
Observations	63	63	61	62	63	63
R-squared	0.718	0.718	0.837	0.722	0.799	0.724

Panel 2: Adding *Indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology,* and *Settler mortality*

Table 4b sample of Non-European countries with Euro share < 0.125

	(1)	(2)	(3)	(4)	(5)	(6)
Euro share	12.22***	12.39**	3.758	12.78**	8.849**	13.00***
	(0.01)	(0.01)	(0.25)	(0.01)	(0.01)	(0.00)
Euro 2000 P-W	1.784**	1.732**	0.686	1.589*	1.059*	1.432*
	(0.03)	(0.04)	(0.27)	(0.07)	(0.10)	(0.06)
Observations	55	55	53	54	55	55
R-squared	0.686	0.686	0.809	0.692	0.780	0.702

Panel 3: Adding *Continent fixed effects for Africa, the Americas, Asia, and Oceania*

<u>Table 4a sample of Non-European countries</u>

Euro share	2.172***	2.050***	0.861*	2.038***	0.364	2.160***
	(0.00)	(0.00)	(0.07)	(0.00)	(0.50)	(0.00)
Euro 2000 P-W	1.060**	1.147**	0.196	1.351**	0.538	0.958**
	(0.02)	(0.03)	(0.70)	(0.01)	(0.18)	(0.05)
Observations	112	112	110	102	112	102
R-squared	0.389	0.390	0.659	0.425	0.553	0.432

Panel 4: Adding *Continent fixed effects for Africa, the Americas, Asia, and Oceania*

<u>Table 4b sample of Non-European countries with Euro share < 0.125</u>

Euro share	14.08***	14.04***	0.836	13.38***	8.508***	13.45***
	(0.00)	(0.00)	(0.75)	(0.00)	(0.01)	(0.00)
Euro 2000 P-W	1.124*	1.135*	0.416	1.345**	0.674	1.119*
	(0.05)	(0.07)	(0.53)	(0.04)	(0.17)	(0.06)
Observations	104	104	102	94	104	94
R-squared	0.346	0.346	0.611	0.377	0.508	0.359

Table 5: Current income and colonial European settlement: Alternative time periods for computing colonial European settlement

This table presents regression results assessing the relationship between *Euro share* and *Current income* while using a uniform time period for *Euro share*. The sample is non-European countries. The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share (1500-1800)* is the average value of the share of European settlement in the country from 1500 to 1800. Euro share (1801-1900) is the average value of the share of European settlement in the country from 1801 to 1900. Euro share (1500-1900) is the average value of the share of European settlement in the country from 1500 to 1900. *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. All specifications are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent:			*Current Income*			
Additional Control:		*British legal origin*	*Education*	*Independence*	*Government quality*	*Ethnicity*
		Panel 1: Euro share computed over 1500 - 1800				
Euro share (1500-1800)	2.435***	2.201***	1.823***	2.475***	1.096***	2.427***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Observations	31	31	28	31	31	31
R-squared	0.478	0.496	0.587	0.505	0.700	0.483
		Panel 2: Euro share computed over 1801 - 1900				
Euro share (1801-1900)	1.989***	1.829***	1.660***	2.229***	1.150***	1.914***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	30	30	29	30	30	30
R-squared	0.547	0.638	0.685	0.566	0.718	0.582
		Panel 3: Euro share computed over 1500 - 1900				
Euro share (1500-1900)	2.618***	2.281***	1.921***	2.717***	1.301***	2.564***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	38	38	35	38	38	38
R-squared	0.584	0.620	0.690	0.610	0.740	0.592

Table 6: Current income and colonial European settlement : Eliminating countries with Euro share = 0

This table presents regression results assessing the relationship between *Euro share* and *Current income* when eliminating countries in which *Euro share* equals zero. The sample is non-European countries. The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. *All* signifies that all of the control variables (*British legal origin, Education, Government quality,* and *Ethnicity*) are included simultaneously. The regressions are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent:				*Current Income*			
Additional Control:		*British legal origin*	*Education*	*Independence*	*Government quality*	*Ethnicity*	*All*
Panel 1: Sample is non-European countries with Euro share > 0							
Euro share	3.929***	4.054***	1.382***	3.625***	1.309**	3.633***	3.239***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)
Observations	64	64	61	64	64	64	64
R-squared	0.389	0.393	0.692	0.422	0.553	0.432	0.451
Panel 2: Sample is non-European countries with 0.125 > Euro share > 0							
Euro share	17.24***	17.44***	5.154*	15.89***	13.23***	16.41***	15.64***
	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	51	51	50	51	51	51	51
R-squared	0.343	0.357	0.605	0.393	0.495	0.357	0.395

Table 7: Current income and colonial European settlement: Including a fixed effect for Euro share = 0

This table presents regression results assessing the relationship between *Euro share* and *Current income* while including *Dummy for euro share equal to zero,* which equals one if *Euro share* equals 0, and zero otherwise. The sample is non-European countries. The dependent variable is *Current income*, computed as the average of the log of real per capita income from 1995 to 2005. *Euro share* is the proportion of Europeans in the colonial population. *British legal origin* takes a value of one if the country's laws are based on the United Kingdom's legal system, and zero otherwise. *Education* equals the average rate of secondary school enrollment from 1995 to 2005. *Independence* equals the fraction of years since 1776 that a country has been independent. *Government quality* is an index of the current level of government accountability and effectiveness. *Ethnicity* measures each country's degree of ethnic diversity. The regressions are estimated using OLS with heteroskedasticity-consistent standard errors. P-values are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. More detailed variable definitions and sources are provided in Table A and the Data Appendix.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent:				*Current Income*		
Additional Control:		*British legal origin*	*Education*	*Independence*	*Government quality*	*Ethnicity*
	Panel 1: Sample is non-European countries					
Euro share	3.929***	3.934***	0.856*	3.746***	0.885*	3.423***
	(0.00)	(0.00)	(0.08)	(0.00)	(0.10)	(0.00)
Dummy for euro share equal to zero	0.223	0.223	0.175	0.238	0.296	-0.0118
	(0.33)	(0.33)	(0.24)	(0.34)	(0.11)	(0.95)
Observations	123	123	119	113	123	111
R-squared	0.173	0.173	0.642	0.193	0.461	0.374
	Panel 2: Sample is non-European countries with 0.125 > Euro share					
Euro share	17.24***	17.28***	1.641	16.51***	11.88***	14.71***
	(0.00)	(0.00)	(0.53)	(0.00)	(0.00)	(0.00)
Dummy for euro share equal to zero	0.836***	0.837***	0.221	0.830***	0.773***	0.485*
	(0.00)	(0.00)	(0.22)	(0.00)	(0.00)	(0.07)
Observations	110	110	108	100	110	98
R-squared	0.113	0.113	0.604	0.132	0.417	0.273

Figure 1: Distribution of colonial European settlement and median current income

This figure shows the number of countries classified in groups according to their European shares at colonization (left axis). The median current income (in logs) for each group is also reported (right axis).
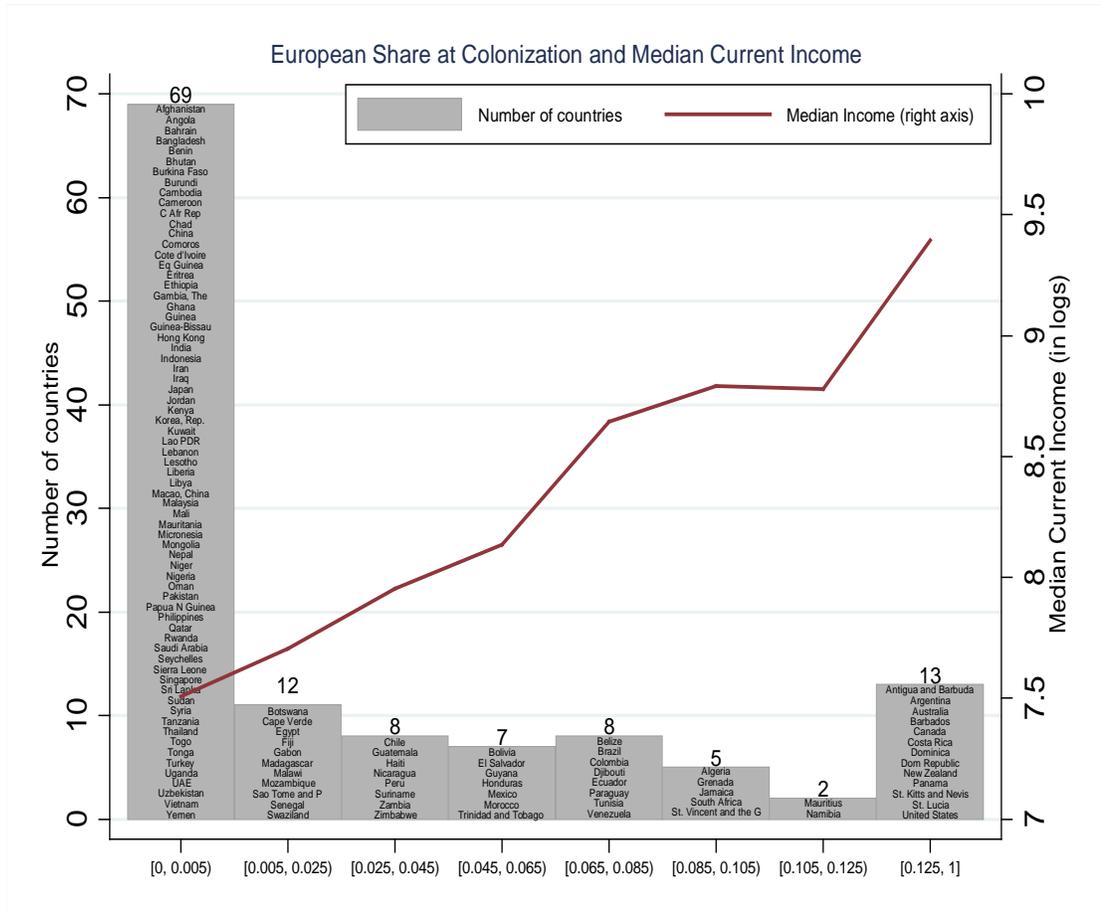
Figure 2: Current income and colonial European settlement

These three figures plot *Current income* (measured by average log of GDP per capita from 1995 to 2005) against *Euro share* (the proportion of Europeans in the colonial population). The figures also include the locally weighted scatterplot smoothing curve (*lowess*), in which simple regressions are fitted to localized subsets of data to produce the nonlinear curve. Figure 2a uses the full sample; figure 2b uses the sample of countries with Euro share < 0.125, and figure 2c uses the sample of countries with 0 < *Euro Share* < 0.125..
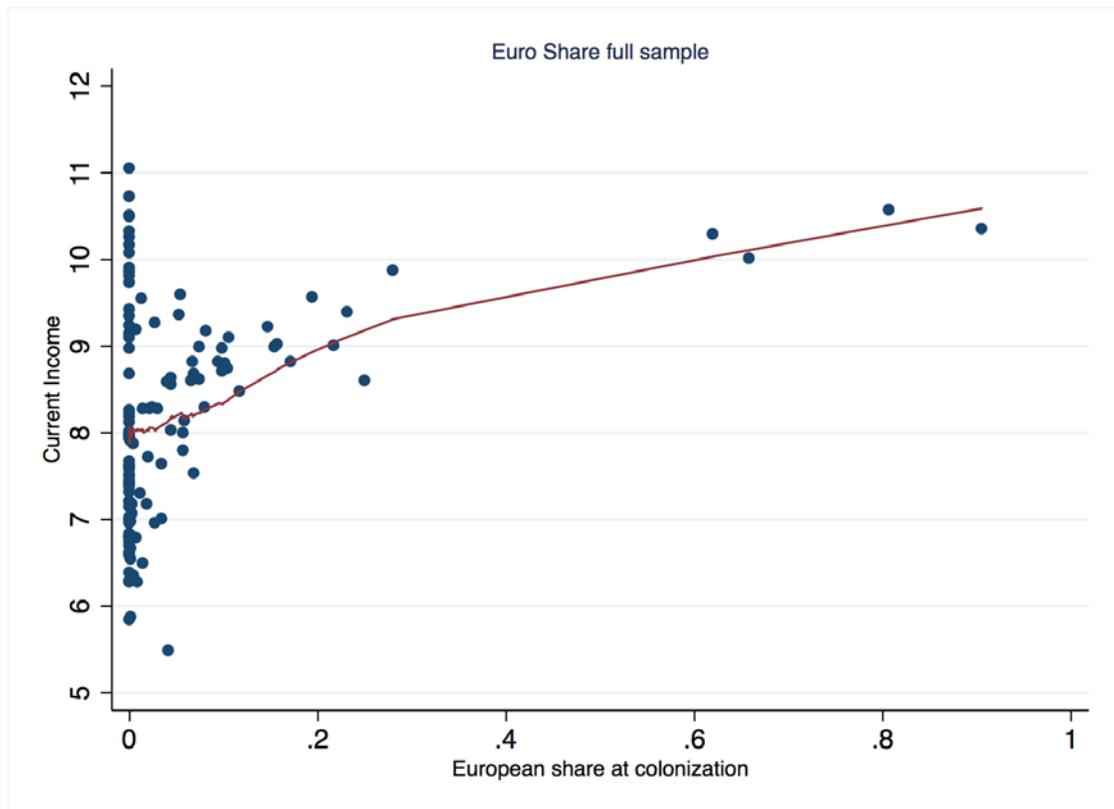
Figure 2a: Euroshare, full sample of countries

Figure 2b: Countries with Euro Share < 0.125

Figure 2c: Countries with 0 < Euro Share < 0.125

Figure 3: Colonial European share and European share today

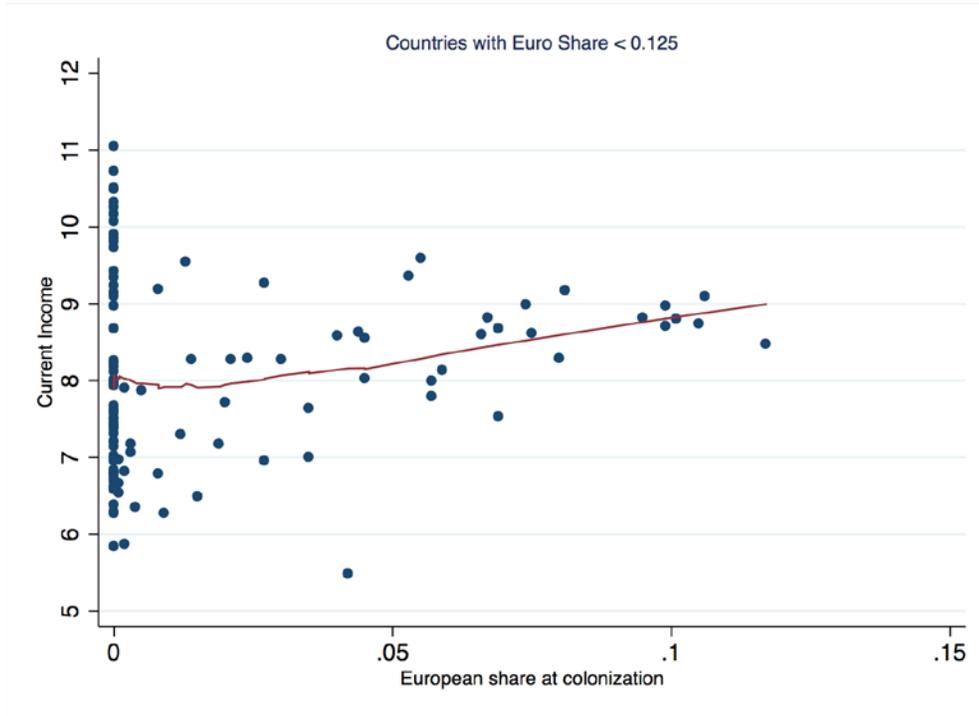This figure shows a simple scatter plot comparing the proportion of Europeans in the Colonial population with the same proportion in 2000.
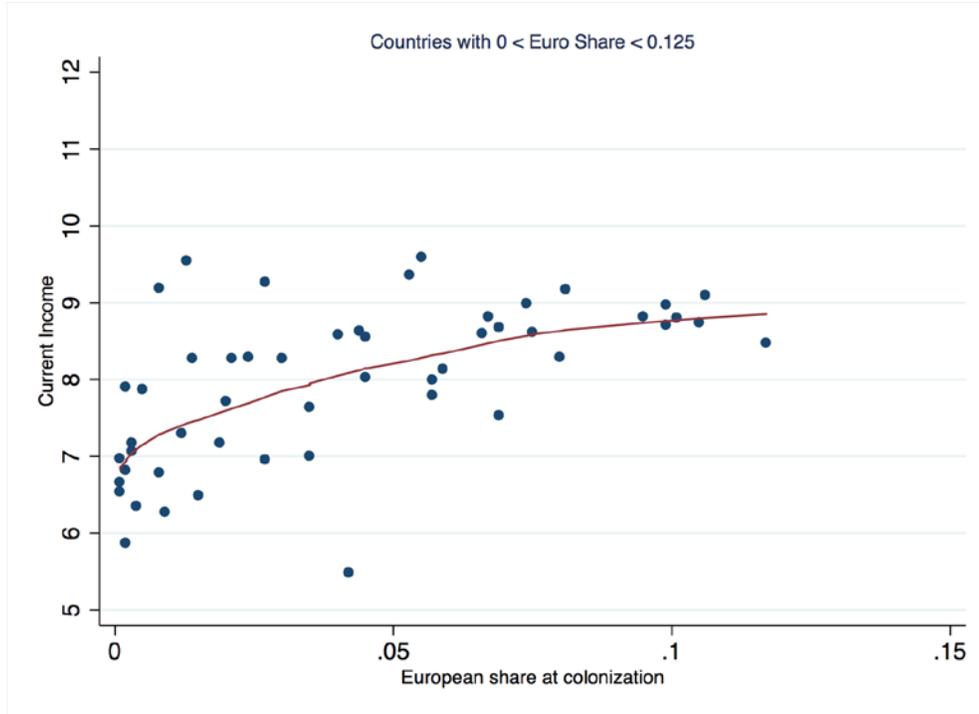
Figure A1: Colonial European share and the Population of European Descent in 1900

This figure shows a scatter plot comparing the proportion of Europeans in the Colonial population to the proportion of the population of European descent in 1900, as reported by AJR.

Colonial European Share and the Population of European Descent in 1900