Where do people live longer?

DOI:
10.1016/j.rie.2020.10.010

Document Version
Accepted author manuscript

Where do people live longer?*

Olayinka Oyekola
INTO University of Exeter
Stocker Road, Exeter, EX4 4PY, UK
E-mail address: o.oyekola@exeter.ac.uk

ABSTRACT

Can historical exposures of non-European countries to European migrants explain part of their current health outcomes? We find that higher European share of the colonial population robustly raised life expectancy and reduced both fertility and infant mortality rates of present-day population in these former colonies. Specifically, after controlling for other plausible determinants, our baseline results imply that, on average, countries at the 95th percentile of the European share of the colonial population, compared to those at the 5th percentile, live 17 years longer, have 1 less child, and experience 54 fewer infant deaths per 1000 live births. A causal interpretation is given to these results by considering various identification strategies. Overall, our results indicate that health fortunes around the world, on average, improved because of European colonial settlers and that differences in the current levels of health performance can be traced back to differential levels of European colonial settlements, where countries that experienced higher influx of colonial Europeans have better health prosperity nowadays than countries with lower inflow of colonial Europeans. A puzzlement arises, however, as countries with no colonial European settlements have outperformed countries with low colonial European settlements. Thus, explaining this phenomenon and exploring how historical migration holds such an enduring influence on the health of nations today opens up an important avenue for future research.

Keywords:
History, Migration, European colonial population, Longevity, Fertility, Mortality

JEL codes:
B15, F63, I10, J10, N30, O10

*Comments and suggestions from Michele Boldrin (Editor-in-Chief) that have helped to sharpen the focus of the paper are gratefully acknowledged.

What has been is what will be, and what has been done is what will be done, and there is nothing new under the sun.

Qoheleth (The Holy Bible, Ecclesiastes 1: 9)

In economics, history is destiny.

Joel Mokyr (1990, The Lever of Riches, p. 3)

1. Introduction

Was colonisation a blessing or a curse? In a recent paper, Easterly and Levine (2016), henceforth EL, have documented factors, such as local population density in 1500 CE, latitude, unfavourable disease environment, and indigenous mortality, which helped in shaping regions that were attractive to European settlers during the colonial epoch. They presented evidence that current income per capita of these countries still register the direct effects of the European share of the colonial population. To do this, EL constructed a new database on the European share of the colonial population for each year, starting in 1540, by combining data from various historical sources on European migratory successes to non-European countries.

Using this data, they find that the economic development of countries with a sizeable historic European share of the population is higher by about 40 per cent than it would have been otherwise. This result is strikingly large and appears to lend support to a corollary of the modernisation hypothesis (Lipset 1959), which claims that colonisation has boosted fortunes of the former colonies by integrating them into global economic, educational, socio-political, and technological systems. Moreover, EL submits that subsequent benefits from these exposures far outweigh the detrimental effects of limited European presence, leading to more extractive policies, that accompany colonial activities in some locations, as documented in the existing literature (e.g., Engerman and Sokoloff 1997, 2002; Acemoglu et al. 2001, 2002; Nunn 2008; Bruhn and Gallego 2012).

As income is only a facet of development, and conveys a people's living standards and welfare levels only in part, we focus on health performance, which is useful in moving away from the unbridled emphasis on income and in shining more light on another dimension of development: human biological well-being.[1] Our perspective is similar to that expressed by the United Nations Development Programme, when they wrote in the first Human Development Report (HDR 1990, p. 9) that: "People are the real wealth of a nation. The basic objective of development is to create an enabling environment for people to enjoy long, healthy and creative lives… [as] income is not the sum total of human life."

[1] Throughout this paper, we use the words "development", "human development", "health outcomes", "health status", and "health performance" interchangeably to refer to any (or all) of life expectancy, fertility rate, and infant mortality rate, except we explicitly state otherwise.

It is also constructive to point out that our focus on health should not be seen as an assault on the relevance of wealth, rather we contend that it is useful to investigate more closely one of the "ends" for which the "means" are sought. On this, Cutler et al. (2006, p. 97) argue that: "The pleasures of life are worth nothing if one is not alive to experience them," and Buchanan and Ellis (1955) also took this view as they correlated post-war economic outcomes and indicators of health performance—which include life expectancy, infant mortality, and food energy intake (i.e., calories consumed per person, per day).[2]

Following this tradition, we ask different but complementary questions to EL: Was the health of non-European countries determined during their historical encounters with the European invaders? Is there a role for the share of Europeans in a population during colonisation to affect the health outcomes of nations today? Why, for example, do the Canadians, and the Japanese live longer, have fewer children and witness more newly-born babies survive till adulthood, while citizens of Ghana and Sierra Leone die younger, have more children, and are accustomed to being bystanders as their infants flatline before their first birthdays? If a country's position on the comparative health development spectrum depends on the extent of European colonial settlement, which colonised countries can be grateful they were?[3]

We find that the European share of the colonial population is strongly positively related to life expectancy, whereas it is highly negatively associated with both fertility and infant mortality rates. Specifically, the direct effect of European share of the colonial population in the baseline regressions, after controlling for British legal origin, year of independence, and ethnic diversity, reveals that: (i) increasing the European share of the colonial population from the 5th percentile, a value of 0.001 for a country like Ghana, to the 95th percentile, a value of 0.659 for a country like New Zealand, is associated with (a) a 21% jump in life expectancy, (b) a 0.4% reduction in fertility rate, and (c) a 73% slump in infant mortality rate; (ii) increasing Ghana's European share of the colonial population to match the frontier value of 0.905 (computed for Canada) implies that the former's (a) life expectancy would rise by 17 to 75 years, (b) fertility rate would drop by 3 to 2 children, and (c) infant mortality rate will fall by 54 to 10 deaths per 1000 live births; and (iii) looking at the two continents with the least (Africa) and most (Oceania) European share of the colonial population, our finding indicates that 23% of the disparities in current life expectancy, 34% of the disparities in current fertility rate, and 29% of the disparities in current infant mortality rate between these two areas can be explained by their European share of the colonial population.

[2] Using health outcomes to capture development is, therefore, not new (see also Lewis 1955 and Morris 1979).
[3] This approach yields results that stress the power of European share of the colonial population in explaining current differences in development, economic or human.

Moreover, noticing that Asia performed much better on all counts of health outcomes, though having on average European share of the colonial population far less than that in Africa, we divide countries in our sample into those with high, moderate, low, or no European share of the colonial population, and ask whether the magnitude of the European share of the colonial population matters for it to be rewarding or harmful. We find that: (i) countries with high European share of the colonial population enjoy greater life expectancy today than countries with moderate, low, and no European share of the colonial population, of 9.22%, 19.5%, and 13.1% respectively; and (ii) countries with low European share of the colonial population have fared worse than countries with no European share of the colonial population, having lower life expectancy, of 6.4%, higher fertility rate, of 14.7%, and a greater infant mortality rate, of 24.5%. This finding, thus, supports the view that distinctions in the forms of distributions of European share of the colonial population at the onset have had an enduring effect that is still detectable in the contemporary dispensation of health measures across the non-European countries.

Although the estimates indicate substantial economic consequences, we may not conclusively interpret them as evidence that European share of the colonial population causes higher life expectancy, lower fertility rate, and reduced the infant mortality rate. A number of reasons can be provided for this: (i) it is plausible that some of the non-European countries were already on the path to the levels of health outcomes that we see today prior to European infiltration, such that dissimilarities in a variety of observed and unobserved country characteristics drive the differences in health development; (ii) measurement errors may also be biasing our estimates; and (iii) including outlying observations may importantly influence our results. We affirm that our results are robust to controlling for an array of observable country characteristics, assessing the bias from unobservable country heterogeneity, omitting outliers, employing an alternative estimation technique, and using different dates to measure health outcomes and European share of the colonial population.

In seeking answers to the above queries, we believe that our paper contributes to two streams of economic literature. First, our paper is related to the stream that endeavours to assimilate the influences of historical and prehistorical factors in causing differences in the levels of current socio-economic development (Diamond 1997; Engerman and Sokoloff 1997, 2002; La Porta et al. 1997, 1998; Acemoglu et al. 2001, 2002; Galor and Moav 2007).[4] The work of Galor and Moav (2007), with which our paper accords well, for example, advances an evolutionary theory that socio-economic and environmental changes taking place in the time elapsed between the Mesolithic and the Copper periods impacted on the nature of the environmental hazards

[4] We review this literature further in the next Section and develop the hypotheses we want to test around it.

confronted by the human population, and that this episodic experience activated an evolutionary process that has had a significant long-term effect on contemporary variations in life expectancy. They then provide empirical analysis showing that a substantial portion of the current differences in longevity across countries (1.6 to 1.9 years, after controlling for national geographical and continental characteristics, as well as income, education, and health expenditure per capita) can be traced back to differences in the time passed since the ancestors of the population in a country underwent the Neolithic revolution.

Our paper adds to this theoretical stance by investigating a different source of historical persistence affecting the health of nations today. In particular, we study whether there is (or not) a marked positive and consistently robust interconnectedness between European colonial settlers and present-day health outcomes. For this purpose, we combine EL's dataset on European share of the colonial population with three standard cross-nationally-available measures of a country's health status (life expectancy, fertility rate, and infant mortality rate) and explore what the long-term effects of Europeans selecting to settle in other regions of the world, starting around 500 years ago, are on the host countries' health outcomes today.

Additionally, our research fits with the economic literature on integration. This stream of research has conventionally focussed on integration through contemporary trade openness (or migration) and its effects on, for example, economic performance (e.g., Grossman and Helpman 1991; Frankel and Romer 1999; Greenaway et al. 2002; Dollar and Kraay 2003; Ortega and Peri 2014; Oyekola 2020). In a parallel stream, which has grown in intensity in more recent times, interest has been focused on other forms of integration across both time and space and their impacts on economic progress, offering a new breed of fundamental determinants of economic development (e.g., Comin and Hobijn 2004; Hibbs and Olsson 2004; Giavazzi and Tabellini 2005; Spolaore and Wacziarg 2009; Cervellati et al. 2018; Özak 2018). EL, Acemoglu et al. (2001, 2002), and Putterman and Weil (2010) mirror brightest the spirit of our work.[5]

The rest of the paper is structured as follows. In Section 2, we review some background literature and place our paper within the existing literature. Section 3 describes the data and offers preliminary correlations between the European share of the colonial population and our indicators of health development. In Section 4, we present our main results and provide

[5] As discussed above, EL constructed historical population share of Europeans that migrated to non-European countries and used it to establish that colonialism has been a blessing to the colonised countries, even in countries where colonisers have previously been associated with leaving diabolical legacies. Acemoglu et al. (2001, 2002), in their empirical studies of the colonial origins of comparative development and the theory of reversal of fortunes, computed and utilised the share of the population in 1900 and 1975 that descended from Europe (we discuss their contributions in more details in Section 2). Putterman and Weil (2010) found a positive long-term relationship between movements of people some 500 years ago and countries' economic outcomes, which has endured until now.

estimates from several extensions and sensitivity checks. Finally, we offer concluding remarks in Section 5, including a suggestive agenda for future research.

2. Background literature and hypotheses development

Economic literature has increasingly confronted how historical legacies have impacted on the state of development for some of the poorest countries in the world. This has led many to seek to explain the origins of key symptoms of underdevelopment. Influential early contributions to the historical determinants of divergent economic trajectories are provided by North and Thomas (1973) and North (1981). In more recent times, Engerman and Sokoloff (1997, 2002), La Porta et al. (1997, 1998), and Acemoglu et al. (2001, 2002), in their seminal works, promote history as the fundamental cause of large differences in cross-country income per capita, though they emphasise different intermediating factors.[6]

The widespread wisdom in all these papers regarding the influence of colonialists on the colonies is that it depends on the former's migration strategy; that is, whether they were *settling* or *exploiting*. In terms of settling, which promotes inclusive activities, such neo-European countries have since gone on to great heights because they have been able to replicate, and sometimes improve on, their homeland's institutional provisions and qualities (with Canada and the United States in the Americas, and Australia and New Zealand in Oceania, providing overriding examples). With regards to exploiting, which endorses extractive activities, on the other hand, many of these former colonies have undergone one of the most gruesome episodes in human history, and despite independence from their colonial masters, it appears that the quotation of Joel Mokyr at the beginning of this text is being fulfilled by the majority of them (with cardinal examples offered by countries from Africa).

An institutionalist view of this is that these two strategies of the colonizing countries birthed a critical juncture in history, after which countries enter into enduring divergent development trajectories (Engerman and Sokoloff 1997, 2002; Acemoglu et al. 2001, 2002, 2009). Accordingly, some countries become wealthy and healthy; whereas, others become poor and poorly. Meanwhile, a human capitalist view would be that what matters at this critical juncture is the quantity and quality of infrastructure available to increase education levels of the colonies' local population, as well as its level (and rate) of technology adoption (Barro 1999; Glaeser et al.

[6] Acemoglu et al. (2005), Nunn (2009), and Spolaore and Wacziarg (2013) provide excellent reviews of this literature.

2004; Gennaioli et al. 2014). On this account, countries that expand access to schooling enrolment and engagement will perform better than those that did not.

A popular approach in these studies is to trace how pre-colonial history impacts on current performance through some post-colonial developments. For example, the style in Acemoglu et al. (2001) is to emphasise the causal effect of current institutions protecting property rights on current economic development and, given both the possibility of feedback from income to institutions and simultaneous impact of omitted variables on both income and institutions, use the expected mortality rates of the first European settlers (soldiers, bishops, and sailors) in the former colonies between the 17th and 19th centuries as an instrument for present-day institutions, in a model specified to determine comparative economic outcomes amongst previously colonised non-European countries.[7]

The maintained hypothesis in this line of research is that pre-colonial and colonial activities of Europeans no longer exert direct effects on current development, but rather works indirectly through institutions (Engerman and Sokoloff 1997, 2002; Acemoglu et al. 2001, 2002, 2005) and other bequests of colonialism, such as human capital accumulation, which is the mechanism underlined by, for instance, Glaeser et al. (2004) and Galor (2011), and socio-cultural values, which is the mechanism highlighted by, for example, Weber (1905), Landes (1998), Barro and McCleary (2003), and Tabellini (2010).[8]

The schematic summary of the above framework for analysing differences in development across countries can be shown as:

$$\begin{aligned} &\text{Historical factors} \rightarrow \text{European settlements} \\ &\qquad \rightarrow \text{Early outcomes (inequality, human capital, institutions, culture)} \\ &\qquad \rightarrow \text{Current outcomes (inequality, human capital, institutions, culture)} \\ &\qquad \rightarrow \text{Current socioeconomic development} \end{aligned} \tag{1}$$

This line of research has piloted a new stream of studies committed to discovering whether there remain any direct long-term effects of history on current outcomes. In terms of economic development, for example, the existing literature has considered the roles of historic

[7] Similarly, Hall and Jones (1999) established that differences in physical capital accumulation, productivity, and output per worker across countries can be explained by differences in institutions and government policies (captured by measures of social infrastructure), instrumenting for the latter using distance from the equator and the extent to which the primary languages of Western Europe are spoken as countries' first languages nowadays.

[8] Further insights have also been contributed with studies by Manning (1990), Bairoch (1993), Grier (1999), Englebert (2000), Bertocchi and Canova (2002), Lange (2004), and Nunn (2008) on the socio-political and development effects of colonialism and slavery.

geographical endowments (Jones 1981; Diamond 1997; Nunn and Puga 2012), population changes (Kremer 1993; Putterman and Weil 2010), state antiquity (Bockstette et al. 2002), long-ago cultural values (Spolaore and Wacziarg 2009; Tabellini 2010), technology adoption before 1500 CE (Comin et al. 2010), and pre-modern genetic diversity (Ashraf and Galor 2013).[9]

Building on the logic of this fast-growing strand of analysis, EL showed that the existing literature is missing a vital historical determinant in a measure that represents how many Europeans permanently relocated to non-European countries beginning around 500 years ago. The ramification of European settlers during colonisation on the post-colonial discrepancy in the amount of prosperity in non-European countries have recently been widely explored, but none of the prior studies measured the European share of the host colonies' population. The construction of this crucial measure is the main contribution of their research.

With this data to hand, EL begin their evaluation of the importance of European share of the colonial population by first providing evidence of the pre-colonial determinants of European settlements. They avail themselves of a simple cost-benefit economic structure of European settlements and make use of some already established historical sources of European colonisation, according to the existing literature (e.g., pre-colonial population density, latitude, and discordant disease environment, among others), as well as construct a measure to capture the degree to which local population died from exposure to European-borne diseases.[10] They empirically show that: (i) European settlement is highly unlikely in countries with high population density in 1500 CE and those already immune to European diseases because of prior contacts with Eurasia—thereby providing a massive support for the theory of a "reversal of fortunes" proposed by Acemoglu et al. (2002)—and (ii) the farther away from the equator a country is, the higher is the probability of being subdued by Europeans.

Having identified what determined European settlement, they turn to its consequences for economic performance, finding: (i) support for both the human capital and the political institutions stances on the relationship between European colonial settlements and current levels of economic development; (ii) that Europeans during colonisation have more material

[9] Other related studies ask: What factors are responsible for the histories we hold, or do not hold, so dear either as their primary question (e.g., Ertan et al. 2016) or as part of their enquiries (e.g., Bruhn and Gallego 2012; Easterly and Levine 2016).

[10] This is a relevant exercise to carry out, as without it, EL's claim that the European share of the colonial population robustly impacts on current economic development may be bogus, especially if these other determinants of European settlement are also associated with the levels of economic development today through other channels safe European share of the colonial population.

effect on economic development today than the proportion of non-European countries in current period that is European descendant; and (iii) asserting that:

> 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. (Easterly and Levine 2016, p. 253 [emphasis added]).

To test these hypotheses in terms of human development, we note that one can also depict the conceptual framework in this second line of research schematically as:

$$\text{Historical factors} \rightarrow \text{European settlements} \rightarrow \text{Current socioeconomic development} \qquad (2)$$

Our paper fits better with the second schematic.[11] The bulk of our analysis follows Eq. (2), which allows us to concentrate on the long-term persistence of the effects of European adventures and misadventures the world over. Specifically, we focus on the link between historical European settlements and contemporary health outcomes in these non-European countries. Based on this, we empirically test the following three hypotheses in this paper:

1. Differences in European colonial settlements can explain differences in current health outcomes in non-European countries.

2. Countries with high European share of the colonial population enjoy better health status presently than countries with moderate, low, or no European share of the colonial population.

3. Countries with low European share of the colonial population experience worse health development now than countries with no European share of the colonial population.

[11] In the working paper (Oyekola 2019), we made some tentative efforts to explore the links within Eq. (1), which an interested reader may consult. Our focus here is on the direct effect. We thank the Editor for pointing out that the empirical results and related puzzle are interesting and challenging enough to leave the equally important task of uncovering the plausible causal mechanism to future research. We come back to this issue in the final section.

3. European colonial settlers and the health of nations today: some empirical evidence

We now give context to our study by presenting our main variables and discuss basic correlations between our outcome measures and the key explanatory variable.[12] To explore our questions, one must find a measure representing the fraction of a country's population that was made up of Europeans during the colonial era. Our source for this data is the recently assembled database of EL.[13] Employing forty-six primary and secondary data sources on the history of colonisation and European migration during this period (as recorded by colonial administrators), these authors compile data on the European share of population in countries around the globe during the colonial era (henceforth European share), ensuring that their definition reflects mainly the group of people that experienced permanent emigration from Europe to other parts of the world during this period.

As the data on the European share is available from 1540 (Chile) to 1990 (US) and since there are no time series entries for each country in the sample, EL identified choosing a date in which to measure European share as a major obstacle. To resolve this, they recognise that European colonial activities happened in a staggered fashion across the world and in pursuance of this, proceed to assign different dates to different countries for the computation of each country's European share of the population during colonisation. A simple rule adopted in doing this is to measure European share in a way that the referenced best year (or range of years) is: (i) at least a century after initial European contact; (ii) at least 50 years before independence; and (iii) take average of European shares of population that are close together.[14]

More particularly, EL found data for European colonial settlements in 69 countries covering all non-European regions of the world and for the remaining countries in our sample, they

[12] We focus here on the main variables and discuss the other variables as they come up. Descriptive statistics for all variables used for analysis are provided in the Appendix (see Table A.1).

[13] Interested readers are advised to see Section 2 of their paper and its Data Appendix for a more detailed description and information on original data sources.

[14] In the Data Appendix to their paper, EL explains their decision: "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 recognise that there is not a precise definition of "the" formative period of colonisation. Nevertheless, influential studies of comparative economic development emphasise 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 colonisation evolved differently there." In a robustness exercise, however, we alternatively make use of the European share measured as the average over the following three different uniform periods: (i) 1500-1800; (ii) 1801-1900; and (iii) 1500-1900.

allocate a value of zero under the strong assumption that no records of "whites" in a population of non-European countries meant that they were not present and validate this by using information from Acemoglu et al. (2001) on the European share of population at the turn of the twentieth century. On average, European share is 7%, with a standard deviation of 17%. If we remove the fifty-eight countries that were assigned a value of zero for European share, the minimum value for the remaining sample of countries is 0.001 (obtained for Rwanda, Uganda, and Ghana), while the maximum value is 0.905 (Canada).

Although we have data on European share for 129 non-European countries, five countries (Bermuda, Hong Kong, Macao, New Caledonia, and Puerto Rico) drop out from our study due to lack of data on our health variables. Among the remaining 124 countries, 51 countries are from Africa, 34 countries are from the Americas, 33 countries are from Asia, and the remaining 6 countries are from Oceania. Fig. 1 shows the distribution of Europeans that were domicile in the 124 non-European countries during colonisation. We ordered countries according to their quartiles in the distribution of European share. Countries with the least European share that fall in the first quartile are coloured in blue, while those with the most European share fall into the fourth quartile and are coloured in red. The intermedial European shares are grouped into the second and third quartiles and are coloured in green and yellow, respectively. Based on this classification, countries with the most European share of the colonial population are found in the Americas and Oceania, while the middle belt of Africa and Asia have the least European share.

Given that our quest involves examining whether the historical share of European population within a country plays a crucial role on its current levels of health outcomes, one also needs appropriate indicators for a country's health performance. The three measures that we use to represent this are life expectancy (henceforth Longevity), fertility rate (henceforth Fertility), and infant mortality rate (henceforth Mortality). Longevity is the number of years a new-born infant is expected to live if the prevailing mortality patterns at the time of its birth persisted throughout its life; Fertility is the age-appropriate fertility rates that correspond to a woman if she lived to the end of her childbearing years; and Mortality is the number of infants per thousand live births in a given year dying before reaching the age of one.[15] All three variables are measured as the average of their natural logs over the 1995-2005 period.[16] The mean of Longevity is 4.15, with a

[15] These variables are based on data from World Bank's World Development Indicator (WDI) that we download from the University of Gothenburg's Quality of Government Institute data link—https://qog.pol.gu.se/.

[16] In a robustness exercise, we alternatively make use of health outcomes measured as the average of their natural logs over the following three periods: (i) 1985-1995; (ii) 2005-2015; and (iii) 1985-2015. We also checked for the time consistency of European share's impact on the health outcomes for every year from 1960 to 2015. Further, we

standard deviation of 0.17; the mean of Fertility is 1.26, with a standard deviation of 0.45; and the mean of Mortality is 3.55, with a standard deviation of 0.88. The minimum values for Longevity, Fertility, and Mortality are 3.663, 0.296, and 1.097, respectively, and the corresponding maximum values are 4.394, 2.038, and 4.949.

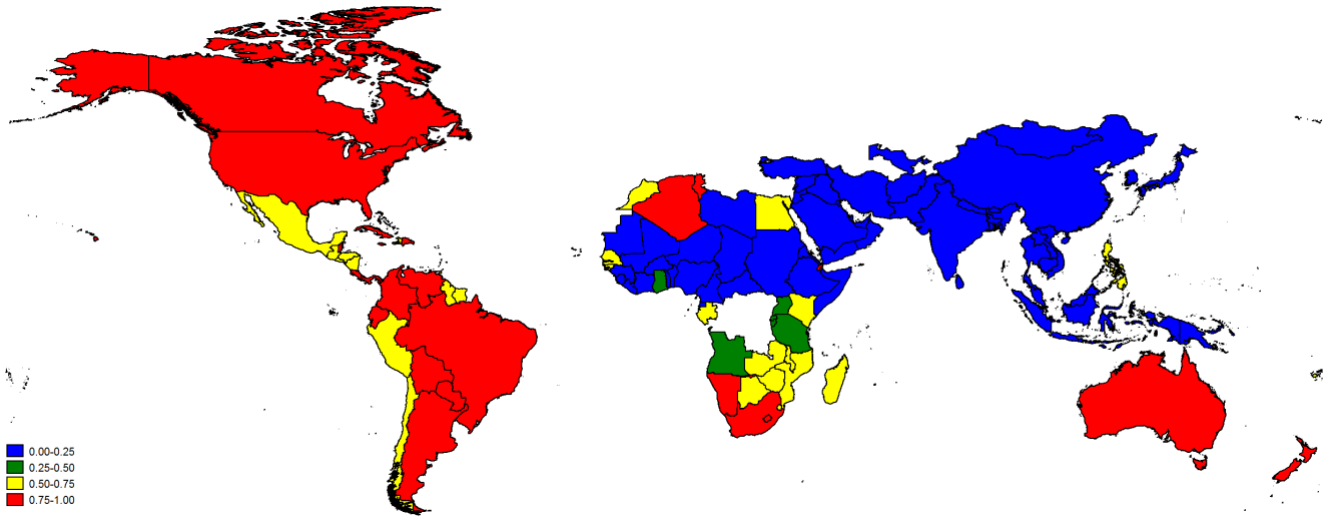

Fig. 1. European settlements around the world during the colonial epoch

Notes: The figure plots the distribution of Europeans that were domicile in the 124 non-European countries during colonisation ordered into quartiles. Colours blue, green, yellow, and red, respectively, refer to countries with European share that fall in the first, second, third, and fourth quartiles. Variable definitions and sources are provided in the text.

According to our baseline health measures, the evidence shows that Sierra Leone today has Longevity of 39 years (the lowest amongst the sample of non-European countries being studied), Fertility of 6 children per woman (one of the highest in the sample),[17] and Mortality of 141 deaths per 1000 live births (the highest in the sample).[18] At the good end of this scale, meanwhile, are the neo-European countries (Australia, Canada, New Zealand, and the United States), as well as the remaining countries from the Americas, Asia, and Oceania, which largely correspond to regions with better health outcomes. For example, Japan features prominently as the country with the most life expectancy, with an average value of 81 years, and also have the lowest fertility rate of "just over" 1 child per woman. Further, Singapore has the lowest infant mortality rate of 3 deaths per 1000 live births.[19]

---

obtain estimates of the effects of European share on a composite measure of human development, Human Development Index from the United Nations Development Programme. We find that our results survive all these checks.

[17] The eleven countries having higher fertility rates are also largely from Africa (Angola, Burkina Faso, Burundi, Chad, Ethiopia, Mali, Niger, Somalia, and Uganda), except for Afghanistan and Yemen.

[18] As already been shown for fertility rate, Sierra Leone is generally flanked by other countries from sub-Saharan Africa in having the worst health outcomes. For instance, the top thirty countries with regards to lowest life expectancy are from Africa, while nineteen of the 20 countries in terms of the prevalence of infant mortality rate are also from this region.

[19] These numbers and the supposed regional clustering of good and bad health outcomes suggest that there is a high correlation between our three measures of health performance. Specifically, the correlations between life

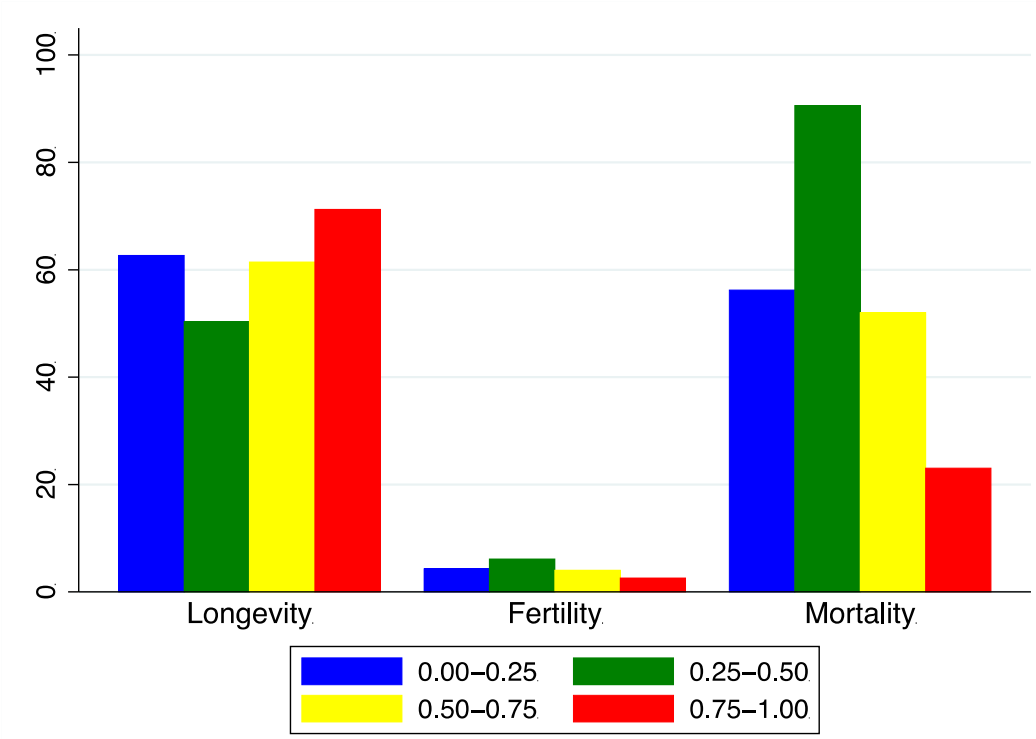

Fig. 2. Distribution of current health outcomes by quartiles of European share of the colonial population

Notes: The figure plots the distribution of Longevity, Fertility, and Mortality by quartiles of European share in 124 non-European countries during colonisation. Colours blue, green, yellow, and red, respectively, refer to countries with European share that fall in the first, second, third, and fourth quartiles. Variable definitions and sources are provided in the text.

We inspect the whole sample and discover two noticeable patterns emerging: (i) countries with higher and no European share tend to be healthier today; and (ii) countries in the middle of the spectrum appear less so. This association is confirmed in Fig. 2, where we use the same quartile categories as for Fig. 1 to obtain averages for each measure of health outcome. The diagram supports our previous observation, portraying a U-shape relationship between European share and Longevity, but a hump-shaped association between European share and both Fertility and Mortality. While the primary hypotheses to be tested in the rest of this paper are that life expectancy on average increases with the European share and that both fertility and infant mortality rates on average decrease with it, we also investigate how our measures of health development respond to different levels—high, moderate, low, or no—of European share.

expectancy and fertility rate, life expectancy and infant mortality rate, and fertility rate and infant mortality rate are respectively -0.82, -0.88, and 0.85.

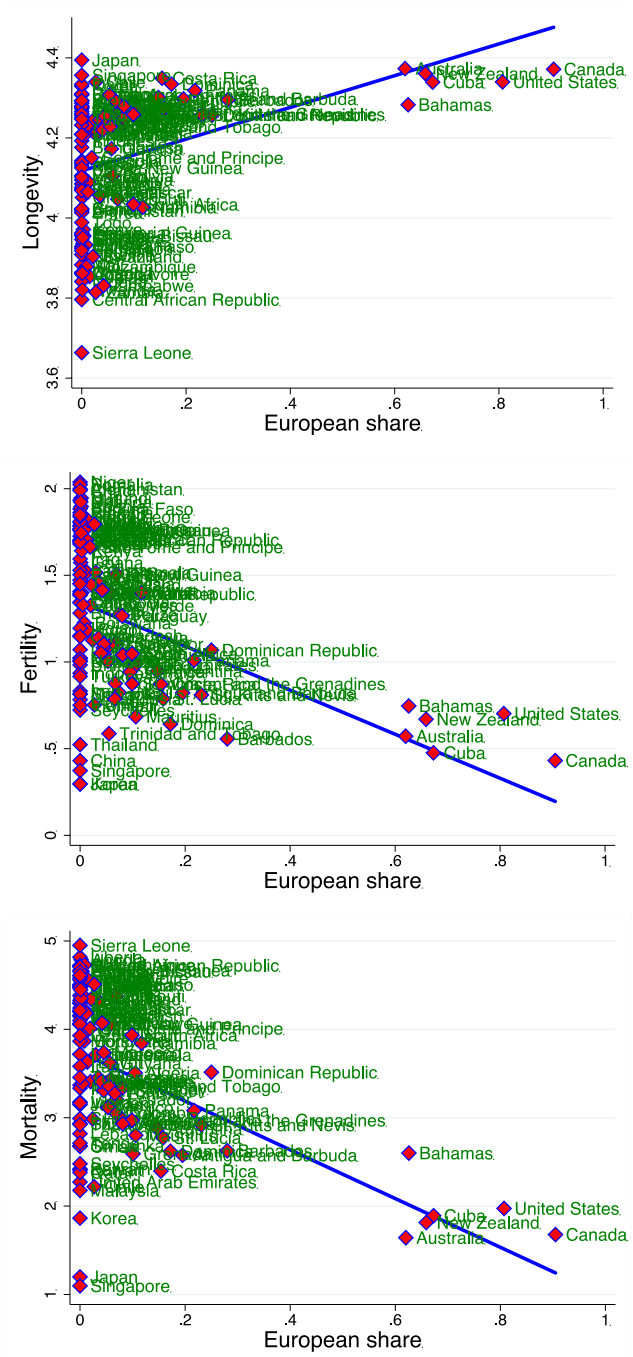

Fig. 3. Relationship between European share of the colonial population and health outcomes

Notes: The figure plots each of Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel) against European share of colonial population. The simple regression fitted values are: (top panel) $\hat{\beta} = 0.396$, p-value = 0.00, $N = 124$, R-squared = 0.141; (middle panel) $\hat{\beta} = -1.270$, p-value = 0.00, $N = 124$, R-squared = 0.204; and (bottom panel) $\hat{\beta} = -2.757$, p-value = 0.00, $N = 124$, R-squared = 0.246. Variable definitions and sources are provided in the text.

Before presenting our more structured model specifications in the next Section, we ask here whether European colonial settlements make an extra life. Fig. 3 provides the scatter plots and the fitted values from the simple regressions of our health outcomes on European share. The graphs support the foundational exposition presented so far that there is a strong positive relationship between Longevity and European share (top panel) and strong negative associations between both Fertility and Mortality and European share (middle and bottom

panels). Additionally, the regression results corroborate these patterns, with estimated coefficients of 0.396, -1.270 and -2.757 on European share with respect to Longevity, Fertility, and Mortality, respectively, and all are statistically significant with a p-value of 0.00. The corresponding R-squared values are indicative that European share of colonial population alone explains about 14%, 20% and 25% of the variations in life expectancy, fertility rate and infant mortality rate, respectively.

It must be noted that this association is unconditional on any other potential determinants of our outcome variables and European share. In the more rigorous examinations discussed below, we control for these other country characteristics that may be potentially important for current national health outcomes, as well as being correlates of European share. Nevertheless, the impressions we are mostly left with are those of marked upward trend in Longevity and downward trends in both Fertility and Mortality vis-a-vis rising European share.

4. Life regressions

Our argumentation so far is in line with the thesis that countries with larger European share should today have higher Longevity, lower Fertility, and less Mortality than countries with smaller or no European share. To test this idea and inspect whether European share is associated with our measures of health outcomes after controlling for other pertinent influences, we postulate a life-gifting regression model for each country i that takes the following form:

$$\ln y_i = \alpha + \beta \text{ European share}_i + \mathbf{C}_i' \gamma + \varepsilon_i \qquad (3)$$

where $\ln y_i$ is a proxy for health development—which is alternatively represented by Longevity, Fertility, and Mortality; European share_i is the proportion of European settlers in a non-European country measured between 1650 and 1962; $\mathbf{C}_i$ is a vector of conditioning country characteristics that are included to capture differences in history and population heterogeneity (and later expanded to include differences in geography, topography, and climate); and ε_i is an unobserved error term. α, β and γ are parameters; β is the critical parameter of interest, which captures the long-term contribution of European share on the health of nations today.

4.1. Baseline results

In order to begin to ground our sense of the depth of the links between European colonial settlements as a share of colonies' population and these former colonies' current human

development, we introduce the same baseline controls as in EL and document these baseline results—estimated by using OLS with heteroskedasticity-consistent standard errors—in Table 1. In column (1), we include only British legal origin (a dummy variable indicating British common law origin) to control for the historic influence of the United Kingdom relative to other colonial superpowers in advancing socio-political institutions and installing common law legal tradition, which have been established in existing literature to be more promising for development (e.g., North 1990; La Porta et al. 1997, 1998). This variable is taken from La Porta et al. (1998).

Column (2) of Table 1 adds only Independence year (the fraction of years since 1776 that a country has been independent) to capture the potential influence of post-colonial country-specific experiences of statehood in running its affairs (Beck et al. 2003; Easterly and Levine 2003). We then condition on only Ethnic diversity (the probability that two randomly selected individuals from a country are from different ethnolinguistic groups) in column (3) to account for how country differences in the degree of citizens' heterogeneity may impact on the health outcomes of nations. The data for both Independence year and Ethnic diversity come from Easterly and Levine (1997). Finally, we run regressions of a specification that simultaneously embed all three country characteristics and document the results in column (4).

For all specifications, the top panel of Table 1 shows that there is a positive and statistically significant association between Longevity and European share. Based on column (4), though the magnitude of the estimated coefficient fell to 0.293 as more controls are added, a decrease of 0.103 relative to the bivariate relationship depicted in the top panel of Fig. 3, it still is statistically significant at the 1% level. The effect is more extensive than those of British legal origin (-0.010 with a *p*-value of 0.74), Independence year (0.096 with a *p*-value of 0.00), and Ethnic diversity (-0.256 with a *p*-value of 0.00).

Surprisingly, the negative coefficient, albeit not significant, on the British legal origin is opposite to what one would expect if British institutional values and legacy, an example of which is its legal tradition, have had a better impact on Longevity in its former colonies relative to those of other European colonisers, as put forward by, for example, North (1981, 1990). On the other hand, the coefficients on Independence year and Ethnic diversity are as one would predict (see, for example, Easterly and Levine 1997, 2003; Montalvo and Reynal-Querol 2005): Independence year positively and significantly correlates with Longevity, which suggests that having greater experience of self-rule may be encouraging for average life expectancy; whereas, Ethnic diversity is a negative and statistically significant covariate of Longevity, which indicates that life expectancy tends to fall with increased country ethnolinguistic divisions.

Table 1: European share of colonial population and health outcomes (baseline results)

	(1)	(2)	(3)	(4)
		Longevity		
European share	0.428***	0.354***	0.309***	0.293***
	(0.00)	(0.00)	(0.00)	(0.00)
British legal origin	-0.0452			-0.0102
	(0.14)			(0.74)
Independence year		0.148***		0.0962***
		(0.00)		(0.01)
Ethnic diversity			-0.281***	-0.256***
			(0.00)	(0.00)
Observations	124	114	112	107
R-squared	0.158	0.225	0.421	0.46
		Fertility		
European share	-1.264***	-1.170***	-0.997***	-0.798***
	(0.00)	(0.00)	(0.00)	(0.00)
British legal origin	-0.00916			-0.136*
	(0.9)			(0.09)
Independence year		-0.327**		-0.266**
		(0.01)		(0.04)
Ethnic diversity			0.643***	0.607***
			(0.00)	(0.00)
Observations	124	114	112	107
R-squared	0.204	0.28	0.382	0.449
		Mortality		
European share	-2.731***	-2.580***	-2.387***	-1.999***
	(0.00)	(0.00)	(0.00)	(0.00)
British legal origin	-0.0369			-0.217
	(0.81)			(0.16)
Independence year		-0.513*		-0.434*
		(0.05)		(0.07)
Ethnic diversity			1.181***	1.164***
			(0.00)	(0.00)
Observations	124	114	112	107
R-squared	0.246	0.284	0.443	0.477

Notes: This table reports the OLS estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel). Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Variable definitions and sources are provided in the text.

Additionally, the middle and bottom panels show that there is a negative and statistically significant association between Fertility and Mortality and European share, even in multivariate

regression specifications. Once again, we observe that the estimated effect of European share falls in magnitude, but not in importance, as we condition on other country characteristics. In both these panels, we find that the additional controls, while not always significant, have their expected signs (see, in particular, column (4)). These results thus confirm that European share has a positive effect on health outcomes today.

The estimated coefficients for the effects of European share on health outcomes are not only statistically significant at the 1% level in all cases, but are also economically meaningful. To be more concrete, consider the following exercise using the magnitudes of the coefficients in column (4), which additionally control for British legal origin, Independence year, and Ethnic diversity. Suppose, during the colonial era, that Ghana, a country with European share of 0.001, instead had European share similar to that experienced by New Zealand (0.659). In such circumstances, the estimated life expectancy for Ghana may have been 70 years rather than its current lower level of 58.[20] The estimated 70 years implies that Ghana would have enjoyed a level of life expectancy much closer to that recorded for New Zealand (78 years) had the former's European share been nearer to the 95th percentile (0.666), instead of being at the 5th percentile (0.001).

Further, we find that the estimates imply that increasing Ghana's European share to match the frontier level of 0.905 (obtained for Canada) would have raised Ghana's Longevity to 75 years, thereby closing the gap between Canada and Ghana from 22 years to 4 years. With regards to implications for regional differences in Longevity, our results suggest that current differences in Longevity between Africa, the continent with the least average European share, and Oceania, the continent with the most average European share, is 0.06 [0.293×(0.217-0.016)].[21] That is, 23% of the difference in current Longevity between Africa and Oceania can be explained using European colonial migration strategy, as reflected in the regions where they chose to settle.[22]

In terms of Fertility and Mortality, the estimates in column (4) indicate that Ghana would today have a fertility rate of around 3 per woman of childbearing age and 17 deaths per 1000 live births, which are again more comparable to New Zealand's 2 and 6 (respectively).[23] Both

[20] This is calculated using: $\Delta \ln y = 0.293 \times \Delta$European share, where y is life expectancy. Ghana's European share is equal to 0.001, whereas New Zealand's is equal to 0.659. This gives: $\Delta \ln y = 0.293 \times 0.658 = 0.193$. Given, therefore, that Ghana's natural log of life expectancy is 4.05, the corresponding new level would be 4.243 (= 4.05 + 0.193). Taking exponential of this value yields approximately 70. This would have ranked Ghana in the low fifties or the 60th percentile among countries with the highest scores for life expectancy and not 89 of 124 countries or in the 30th percentile as it currently stands.

[21] The averages of European share for Africa, the Americas, Asia, and Oceania are 0.016, 0.179, 0.000, and 0.217, respectively.

[22] The average Longevity in Africa and Oceania are 3.995 and 4.257, respectively, such that $(0.06 \div 0.262) \times 100 \approx 23\%$.

[23] Using data for the same two countries, we gauge now the economic significance of European share for Fertility and Mortality. These are calculated using: $\Delta \ln y = \hat{\beta} \times \Delta$European share, where y is either fertility rate, with $\hat{\beta} =$

numbers signal huge improvements for both fertility and infant mortality rates when compared to Ghana's actual values of 5 and 64 (respectively). Supposing again that Ghana's European share is identical to the frontier value, Ghana's current Fertility would have been 2 and its Mortality 10. Besides, 34% and 29% of the differences in current Fertility and Mortality between Africa and Oceania can be accounted for using European share.[24]

The results in Table 1, while showing strong correlations between European share and Longevity, Fertility and Mortality, are not conclusive. In the remaining parts of this section, we, therefore, implement various robustness checks: (i) to assess the sensitivity of the baseline results for the time consistency of the effects of European share; (ii) to omitting influential observations; (iii) to controlling on additional observable determinants; (iv) to assessing the potential bias from unobservable factors; and (v) to examining whether our baseline results would hold up against alternative estimation techniques and model specifications. Finally, we provide a summary of several additional results (not reported to conserve on space but are available on demand), which involves alternative sample compositions, an appraisal of which Europeans—colonial or coetaneous—matter, conditioning on more observed—pre-colonial, colonial, and post-colonial—country characteristics, and to differences in the periods over which our main explanatory variable is measured.

4.2. Time consistency of the relationship between European share of the colonial population and health outcomes

Our baseline estimates are based on health outcomes obtained as a cross-sectional average over the 1995-2005 period.[25] We first checked the robustness of our baseline estimates to using health outcomes measured over three different alternative periods: 1985-1995, 2005-2015, and 1985-2015. This extension allows us to test whether the afore-discussed results are driven by the years we use to define "current" health outcomes. The results in Table 2 answer in the negatory.

-0.798, or infant mortality rate, with $\hat{\beta} = -1.999$ (see column (4) of Table 1). Ghana's European share is equal to 0.001, whereas New Zealand's is equal to 0.659. This gives: $\Delta \ln y = -0.798 \times 0.658 = -0.525$ for Fertility and $\Delta \ln y = -1.999 \times 0.658 = -1.315$ for Mortality. Given, therefore, that Ghana's natural logs of fertility and infant mortality rates are 1.573 and 4.156, the corresponding new levels would be $1.048 (= 1.573 - 0.525)$ and $2.841 (= 4.156 - 1.315)$. Taking their exponentials yields approximately 2.85 and 17.13.

[24] These values are computed as follows: (i) current differences in Fertility between Africa and Oceania that is associated with historical differences in European share is -0.160 $[-0.798 \times (0.217 - 0.016)]$, and the average Fertility in Africa and Oceania are 1.595 and 1.129, respectively, such that $(-0.160 \div -0.466) \times 100 \approx 34\%$; (ii) current differences in Mortality between Africa and Oceania that is associated with historical differences in European share is -0.402 $[-1.999 \times (0.217 - 0.016)]$ and the average Mortality in Africa and Oceania are 4.226 and 2.821, respectively, such that $(-0.402 \div -1.405) \times 100 \approx 29\%$.

[25] We have used a decadal average of the health measures to lessen the effects of short-term fluctuations and missing observations.

Hence, European share is strongly positively (negatively) correlated with Longevity (both Fertility and Mortality) regardless of whether these outcome measures are defined over 1985-1995, 1995-2005 (baseline), 2005-2015, or 1985-2015.

Furthermore, we ask: Is the impact of European share on health development consistent over the post-colonial years? Taking 1960 as the end of the colonial reign, we test for the time consistency of the influence of European share on current health performance of non-European countries for each year with available data until 2015. Fig. 4 depicts the changes in the size and significance of European share on Longevity (top panel), Fertility (middle panel), and Mortality (bottom panel). As shown, the estimated coefficient for European share on Longevity, which is consistently positive, is monotonically decreasing as we move closer to the present, starting with a value of 0.49 in 1960 and ending with a value of 0.19 in 2015.[26]

For both Fertility and Mortality, we find that the estimated coefficients over the 56 years are consistently negative, supporting the baseline estimates. We can, however, see that there is a U-shaped relationship between these two measures of health outcomes and European share, in which both Fertility and Mortality initially decrease and then increase as we travel further away from the year when European share was measured. Given these results, we need not worry about health measures averaged over 1995-2005 as our main outcome variables.

4.3. Influential observations

A source of concern regarding our estimates in Table 1 is that they may be biased because of the inclusion of some particularly influential observations. A visual inspection of Fig. 3 reveals that some countries are notable outliers in terms of their European share (e.g., Australia, the Bahamas, Canada, Cuba, New Zealand, and the United States). We adopt three strategies to verify whether our baseline results will hold up in the absence of these countries.

First, we discuss in Section 3 that 55 of the countries that enter into our parsimonious bivariate regressions (see Fig. 3) were assigned a value of zero for their European share because there was no evidence to the contrary. To ensure that this arbitrary data imputation does not drive our results, we repeat the regressions of Table 1 in columns (1)-(4) of Table 3 after dropping countries with European share that equals zero. Looking at the two sets of estimates, we confirm that dropping these countries makes little (no) difference to the magnitude (significance) of our coefficients.

[26] We note that all the coefficients are statistically significant at the 1% level, including those for Fertility (except for the year 2015 regression for which the coefficient is significant at the 5% level) and Mortality.

Table 2: European share of colonial population and health outcomes (health outcomes measured over different time periods)

	1985-1995				2005-2015				1985-2015			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Longevity						
European share	0.427***	0.388***	0.347***	0.297***	0.344***	0.274***	0.245***	0.224***	0.398***	0.339***	0.300***	0.271***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	124	114	112	107	122	112	110	105	124	114	112	107
R-squared	0.157	0.198	0.381	0.402	0.161	0.256	0.422	0.482	0.165	0.235	0.428	0.465
						Fertility						
European share	-1.410***	-1.311***	-1.208***	-0.962***	-1.012***	-0.884***	-0.760***	-0.531***	-1.227***	-1.129***	-0.989***	-0.767***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	124	114	112	107	122	112	110	105	124	114	112	107
R-squared	0.304	0.364	0.398	0.476	0.137	0.237	0.36	0.442	0.218	0.298	0.388	0.462
						Mortality						
European share	-2.597***	-2.564***	-2.408***	-2.022***	-2.539***	-2.304***	-2.057***	-1.736***	-2.615***	-2.475***	-2.278***	-1.916***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	124	114	112	107	124	114	112	107	124	114	112	107
R-squared	0.285	0.291	0.446	0.477	0.21	0.281	0.439	0.484	0.25	0.289	0.453	0.486

Notes: This table reports the OLS estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel). Columns (1), (5), and (9) control for British legal origin; columns (2), (6), and (10) control for Independence year; columns (3), (7), and (11) control for Ethnic diversity; and columns (4), (8), and (12) control for all three variables jointly. Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Variable definitions and sources are provided in the text.

Table 3: European share of colonial population and health outcomes (influential outliers)

	omit if European share = 0				omit if European share ≥ 0.125				omit if |DFBETA| >2/√N			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Longevity						
European share	0.486***	0.356***	0.340***	0.381***	1.386***	1.274***	1.243***	1.175***	0.554***	0.473***	0.362***	0.326***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	66	66	65	65	109	99	98	93	120	112	109	103
R-squared	0.338	0.363	0.383	0.455	0.09	0.167	0.377	0.426	0.137	0.224	0.449	0.494
						Fertility						
European share	-1.258***	-1.195***	-1.075***	-0.936***	-4.299***	-4.036***	-3.669***	-3.392***	-1.527***	-1.524***	-1.203***	-1.013***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	66	66	65	65	109	99	98	93	122	109	110	100
R-squared	0.39	0.408	0.46	0.472	0.102	0.208	0.32	0.401	0.184	0.264	0.376	0.527
						Mortality						
European share	-2.951***	-2.761***	-2.627***	-2.359***	-5.562***	-4.930**	-5.281***	-4.300**	-3.549***	-3.117***	-2.972***	-2.594***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	66	66	65	65	109	99	98	93	120	110	108	102
R-squared	0.529	0.556	0.57	0.587	0.047	0.099	0.3	0.343	0.229	0.218	0.495	0.535

Notes: This table reports the OLS estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel). Columns (1), (5), and (9) control for British legal origin; columns (2), (6), and (10) control for Independence year; columns (3), (7), and (11) control for Ethnic diversity; and columns (4), (8), and (12) control for all three variables jointly. Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Variable definitions and sources are provided in the text.

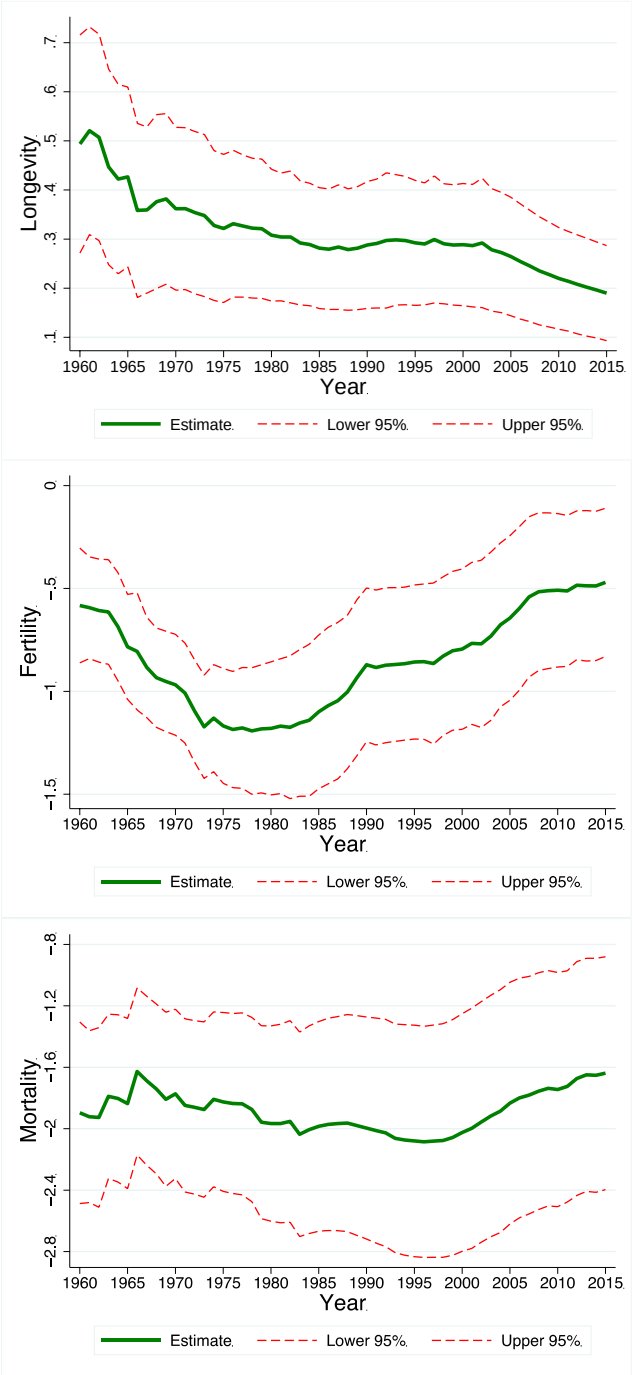

Fig. 4. Relationship between European share of the colonial population and health outcomes over time

Notes: The figure plots the OLS point estimates and the 95% confidence intervals for European share of colonial population, with each of Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel) obtained for each year from 1960-2015. Variable definitions and sources are provided in the text.

Second, we check whether our results are driven by a few former colonies, which by accident or providence just happened to have a disproportionately sizeable European presence in their colonial population and are also currently experiencing a favourably higher quality of health. To explore this, we follow EL in removing countries with a large fraction of Europeans, which they

conservatively pegged as any country with European share greater than 12.5%. Columns (5)-(8) of Table 3 document the results from using the sample of former colonies with European share of the colonial population less than 12.5%. Comparing these results with the ones in Table 1, we observe that using the restricted sample only strengthens our results further. In particular, the estimated effects of European share on Longevity, Fertility, and Mortality are more than twice as substantial in all cases. This thus resolves the concern that the positive impact of European share on our health outcomes today may have been solely due to the dominant influence of settler colonies, such as Australia, Canada, New Zealand and the United States.

Third, we compute DFBETA for each regression specified in Table 1 based on Belsley et al. (1980) and report the results from excluding all observations with $|\text{DFBETA}| > 2\sqrt{N}$, where N refers to each regression sample size, in Columns (9)-(12) of Table 3. As shown, the sign, size and significance of the estimated coefficients of European share are retained in all specifications.

4.4. Controlling on observables

The OLS estimates reported so far reveal that there is a strong association between health outcomes in non-European countries today and their European share of the colonial population. However, we cannot yet attribute to them a causal interpretation. Many historical accounts and empirical evidence demonstrate that Europeans selected to settle in countries with better health quality (or prospect). In *The Spirit of the Laws*, for instance, Montesquieu (1748), in discussing the influence of soil on laws, writes that it, "[I]s natural for a people to leave a bad soil to seek a better; and not to leave a good soil to go in search of a worse. Most invasions have, therefore, been made in countries which nature seems to have formed for happiness…" (2015, p. 217). We think that this is also congruent with the highly acclaimed "extractive vs inclusive states" hypothesis of Acemoglu et al. (2001). Could it be that these countries just happened to continue to enjoy (or eventually realise their promising) higher quality of health post-colonial era? If this was the case, it is possible, therefore, to observe the positive effect of European share on Longevity and its negative effect on Fertility, and Mortality.

Thus, a further possible source of bias in our baseline results is that some other country characteristics, which may be correlated with the share of Europeans in a colonial population, and contemporary health outcomes in non-European countries have been omitted from the analyses. In Table 4, we address this issue by re-estimating the regressions in Table 1 after introducing seven additional controls, whose omissions are likely to end in incorrectly ascribing their impacts to European share. To do this, we follow EL in controlling for more observables and formally estimate a life-gifting model of the form:

Table 4: European share of colonial population and health outcomes (observable factors)

	(1)	(2)	(3)	(4)
			Longevity	
European share	0.206**	0.167*	0.139	0.165*
	(0.03)	(0.08)	(0.1)	(0.05)
British legal origin	-0.0345			-0.0296
	(0.33)			(0.47)
Independence year		-0.0346		-0.0726
		(0.36)		(0.19)
Ethnic diversity			-0.126**	-0.121**
			(0.02)	(0.04)
Observations	62	61	62	61
R-squared	0.775	0.768	0.795	0.801
			Fertility	
European share	-0.929***	-0.856***	-0.750***	-0.795***
	(0.00)	(0.00)	(0.00)	(0.00)
British legal origin	0.0885			0.0274
	(0.26)			(0.76)
Independence year		-0.0329		0.0367
		(0.86)		(0.86)
Ethnic diversity			0.353**	0.339*
			(0.05)	(0.07)
Observations	62	61	62	61
R-squared	0.773	0.758	0.795	0.786
			Mortality	
European share	-2.422***	-2.335***	-2.188***	-2.093***
	(0.00)	(0.00)	(0.00)	(0.00)
British legal origin	0.0448			-0.0281
	(0.8)			(0.87)
Independence year		0.157		0.265
		(0.64)		(0.44)
Ethnic diversity			0.771**	0.836*
			(0.05)	(0.06)
Observations	62	61	62	61
R-squared	0.715	0.718	0.749	0.756

Notes: This table reports the OLS estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel). All the regressions also control for Biogeography, Latitude, London, Malaria ecology, Precious metals, Settler mortality and continent fixed effects. Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Variable definitions and sources are provided in the text.

$$\ln y_i = \alpha + \beta \text{ European share}_i + \mathbf{C}_i' \gamma + \mathbf{X}_i' \delta + \varepsilon_i \quad (4)$$

where everything is as in Eq. (3), except that we now have an added vector of conditioning country characteristics, **X**. To be more specific, the components of **X** are: Latitude (absolute value of the distance of the colony from the equator, normalized to lie between zero and one from CIA World Factbook), Precious metals (an indicator of whether the region has valuable minerals from Easterly (2007)), London (a measure of the distance of the colony from London calculated as the average of the oceanic distance from London to all seaports in a country from Easterly and Levine (2016)), Biogeography (an index of the prehistoric availability of storable crops and domesticable animals from Hibbs and Olsson (2004)), Malaria ecology (an index of the stability of malaria transmission in a region from Kiszewski et al. (2004)), Settler mortality (a measure of historical deaths per annum per 1000 European settlers from Acemoglu et al. (2001)), and continent fixed effects (a dummy variable indicating a country's continental association to be either Africa, the Americas, Asia, Europe, or Oceania using UN classifications).[27]

Looking first at the coefficient of our central focus, we see that European share continues to retain its significance impressively. This occurs in 11/12 regressions: the estimated coefficients of the effects of European share on Fertility and Mortality did indeed continue to be statistically significant at the 1% level in all eight cases and are statistically significant at least at the 10% level in three of the four Longevity specifications. As can be anticipated, the magnitude of the estimated coefficient of European share on health outcomes falls as more controls are added (except in column (4) of the bottom panel, where it is documented that the size effect of European share on Mortality is marginally strengthened with the expanded control set). Understandably, the model now explains 71.5%-80.1% of the variations in our health outcomes, rather than 15.8%-44.9% in the baseline regressions.

Turning to the other variables also included in the baseline specification, we find that both British legal origin and Independence year are now unimportant when controlling on these additional factors, whereas Ethnic diversity is consistently negatively and significantly related to Longevity and consistently positively and significantly associated with both Fertility and Mortality. In terms of the newly added controls in **X** (not reported to conserve on space but are available on demand), we find the distance from London, which was a centre of invaluable

[27] While including these additional factors meant that our sample size falls from 124 in column (1) of Table 1 to 62 in column (1) of Table 4, they help to ascertain the robustness of our baseline estimates to the roles that mineral resources, cost of migration, (un)friendly climate, adverse disease environment, and pre-colonial variables may play in causing European settlement, thereby in affecting comparative health development.

influence historically, to be the only determinant entering significantly in all specifications, having a positive effect on Longevity and negative impacts on both Fertility and Mortality. Could this be an indication that many countries have performed better irrespective of their geographical proximity to London?

Other variables with essential effects on our health outcomes are Biogeography (which enters positively in models of Longevity and negatively in models of Fertility and Mortality) and Settler mortality (which enters negatively and positively in respective specifications, where Longevity and Mortality are the dependent variables, but insignificant in the Fertility specifications). We also find that the continent fixed effects for the Americas are more critical for explaining Longevity, while those for Oceania are more relevant for explaining Fertility, with both entering positively in these regressions. Regarding the remaining variables, we find that Latitude, Precious metals, Malaria ecology, and fixed effects for Africa are not significant in all twelve specifications.

4.5. Accounting for unobservables

It may also be instructive to examine the extent to which unobserved heterogeneity in country characteristics may be biasing our estimates in Table 1. To get around this problem, we exploit the methodology of Altonji et al. (2005), as also applied by, for example, Nunn and Wantchekon (2011) and Fenske (2015). In particular, we compute Altonji et al.'s (2005) ratios using $\hat{\beta}^F/(\hat{\beta}^R - \hat{\beta}^F)$, where $\hat{\beta}^F$ designates the estimated coefficient for European share from a model with a full set of controls and $\hat{\beta}^R$ denotes the estimated coefficient for European share from a model with a restricted set of controls. According to Altonji et al. (2005), having $|\hat{\beta}^F/(\hat{\beta}^R - \hat{\beta}^F)| > 1$ could mean one of two things: the first is that the effect of unobserved country characteristics must be higher relative to observed country characteristics for it to explain away our baseline results, and the second is that adding controls sometimes goes towards improving the baseline estimates. Both these implications are in large parts confirmed below.

We present the results in Table 5, where we consider two sets of full controls and five sets of restricted controls. In terms of models with full controls, we consider first the baseline controls of British legal origin, Independence year, and Ethnic diversity used in column (4) of Table 1, and the second is an expanded set of controls, which extends our baseline controls to include Biogeography, Latitude, London, Malaria ecology, Precious metals, Settler mortality and continent fixed effects from column (4) of Table 4. In terms of models with restricted controls, the five options we look at are: (i) No controls; (ii) British legal origin; (iii) Independence year;

(iv) Ethnic diversity; and (v) British legal origin, Independence year, and Ethnic diversity. As our fifth set of restricted controls is identical to the baseline controls, their combination drops out from the analysis.

Table 5: European share of colonial population and health outcomes (unobservable factors)

Controls in the restricted set	Longevity	Fertility	Mortality	*N*
	Controls in the full set: baseline controls			
None	1.918	1.553	2.172	107
British legal origin	1.438	1.585	2.167	107
Independence year	3.151	2.059	2.866	107
Ethnic diversity	25.79	4.117	6.411	107
	Controls in the full set: expanded controls			
None	0.626	1.893	2.59	61
British legal origin	0.606	2.091	3.053	61
Independence year	1.046	4.711	4.92	61
Ethnic diversity	1.291	6.68	7.162	61
British legal origin, Independence year, and Ethnic diversity	14.33	-2.898	-4.9	61

Notes: This table reports Altonji, Elder and Taber (2005) statistics computed as $\hat{\beta}^F/(\hat{\beta}^R - \hat{\beta}^F)$, where $\hat{\beta}^F$ is the estimated coefficient for European share from a model with a full set of controls and $\hat{\beta}^R$ is the estimated coefficient for European share from a model with a restricted set of controls. Baseline controls are British legal origin, Independence year, and Ethnic diversity from column (4) of Table 1. Expanded controls are baseline controls plus Biogeography, Latitude, London, Malaria ecology, Precious metals, Settler mortality and continent fixed effects from column (4) of Table 4. Variable definitions and sources are provided in the text.

Based on the overall pattern shown on the table, we conclude that European share does an excellent job in explaining our health outcomes, especially Fertility and Mortality, even after considering the likely effects of unobservable country-specific factors.

4.6. Alternative estimation technique

By construction, meanwhile, we do not need to worry about reverse causation as discussed in the Introduction. Another well-known snag to the accuracy of the OLS estimates is measurement error. Conceivably, EL's European share is well-resourced and expertly compiled, but it remains that it is only as accurate as the available data sources allowed. Indeed, several data sources and non-standardised statistical methods used by different colonial regimes over different periods and covering different geographic areas are without doubt themselves a source of coding errors. Moreover, the authors also wrote that because of "data limitations, we cannot

always adhere to our own guidelines for choosing the date on which to measure Euro share" (Easterly and Levine (2016, p. 231)), which is likely to initiate further mismeasurement.

To tackle this concern, we rely on identification through heteroskedastic covariance restrictions (IHCR) that has been shown to be present in many models with mismeasured regressors.[28] We briefly summarise this procedure here, which is based around sub-section 1.2 of Lewbel (2012, p. 68). Suppose that our goal is to estimate the following linear econometric model (similar to the one between our measures of health outcomes and European share):

$$Y_1 = X' \beta_1 + Y_2^* \gamma_1 + V_1 \tag{5}$$

where V_1 is mean zero and independent of the elements of X and Y_2^* (the mismeasured regressor), such that we can define:

$$Y_2 = Y_2^* + U \tag{6}$$

where U is the classical measurement error, with usual assumptions: $E(U) = 0$; $U \perp X, Y_1, Y_2^*$. Further suppose that V_2 is the residual from linearly projecting Y_2^* on X, such that:

$$Y_2^* = X' \beta_2 + V_2 \tag{7}$$

where $E(XV_2) = 0$. Then, re-writing Eq. (6) as $Y_2^* = Y_2 - U$ and substituting this into Eq. (5) yields:

$$Y_1 = X' \beta_1 + Y_2 \gamma_1 + \varepsilon_1 \tag{8}$$

where $\varepsilon_1 = -\gamma_1 U + V_1$ is an unobserved error term. It follows also that by combining Eq. (6) and (7) we get:

[28] This sort of problem is conventionally dealt with by using instrumental variables (IV) that are correlated with the explanatory variable (European share in our case) but are uncorrelated with other country characteristics. Such instruments would isolate the exogenous source of variation in the variable of interest, thereby making it easier to discern the causal impact running from it to the relevant outcome variables, which in our case are life expectancy, fertility rate and infant mortality rate. Additionally, IV estimates are likely to be more consistent in the face of measurement error in the European share of the colonial population. The challenge is in finding valid external instruments. This leads us to adopt IHCR as our baseline IV methodology since it mainly uses internal instruments based on the heteroskedasticity of the structural shocks. For comparison and to check the efficiency of our estimates, we have also included as an additional robustness check, the results from: (i) using only an external instrument; and (ii) using IHCR and an external instrument.

$$Y_2 = X' \beta_2 + \varepsilon_2 \quad (9)$$

where $\varepsilon_2 = U + V_2$ is an unobserved error term.

We note that the structural Eq. (8) is analogous of our regression specification in Eq. (3), and in Eq. (9), we now have an instrument equation. An attractiveness of IHCR is that we do not require any external instruments for identification purposes. The important constraint that governs whether the parameters of our structural model are identified is that V_2 is heteroskedastic (i.e., $\text{cov}(X, \varepsilon_2^2) \neq 0$), together with other standard assumptions (basically that $E(X\varepsilon_1) = E(X\varepsilon_2) = \text{cov}(X, \varepsilon_1\varepsilon_2) = 0$).

In Table 6, we document the results from using IHCR in columns (1)-(4). We can see that using IHCR approach mostly does not affect the coefficients on European share. All the estimates continue to have expected sign and are statistically significant at the 1% level in 9/12 regressions. In terms of magnitude, certain IHCR estimates are close to their OLS counterparts, though the latter is generally larger. We interpret the lack of a significance in column (2) for Longevity and Fertility as likely due to a problem of a weak internal instrument in a model with only Independence year available for use to generate the heteroskedastic-based instrument.

For comparison, we employ subsistence mode in the pre-modern world as an external instrument for European share, such that our instrument equation can now be specified as:

$$\text{European share}_i = a + b\ \text{Subsistence mode}_i + \mathbf{C}_i' c + d_i \quad (10)$$

where Subsistence mode is an index of the extent to which sedentary agriculture was practised in 1000 CE within a region delineated by a country's modern international borders, which is taken from Ashraf and Galor (2013); and d_i is an unobserved error term.

The exclusion restriction is that Subsistence mode does not enter Eq. (3), which is credible insofar as subsistence mode in 1000 CE only affected health outcomes today by affecting European share and not through any other channels. Subsistence mode as an instrument is also consistent with the historical accounts of Engerman and Sokoloff (1997, 2002) and Acemoglu et al. (2001, 2002) that Europeans founded settler colonies in locations, where it is practicable to carry out small-scale agriculture. The results from these two-stage least squares estimations that use only external instruments are provided in columns (5)-(8) of Table 6. As shown, the estimated effects of European share have expected signs concerning each health outcome and are statistically significant at the 1% level in all specifications. We also observe that estimated

effects from columns (5)-(8) are larger than those from columns (1)-(4) and are closer to the baseline OLS estimates.

Lastly, as suggested by Lewbel (2012), that having an outside instrument can potentially improve efficiency, we present in the final four columns of Table 6 estimates from exploiting heteroskedasticity in the model's residuals and incorporating our external instrument (Subsistence mode) as instruments. The results are virtually the same as those in columns (5)-(8). In general, the results from the instrumental variables (IV) regressions ratifies the OLS estimates of a positive link between European share and current life expectancy in non-European countries, as well as the negative ties between the former and both fertility and infant mortality rates in these countries.

Moreover, we test for instrument quality by reporting the first-stage Kleibergen and Paap (2006) F-statistic for weak identification test. More particularly, we find that the test statistic from all our regressions is above the rule-of-thumb of 10. Further, we confirm that we can reject the null of weak instruments in all models, as all the obtained test statistics lie above all the Stock and Yogo (2005) critical values, including the most stringent, with one exception. This occurs in column (2), where the reported Kleibergen-Paap F-statistic of 13.43 lies between the 25% max IV size of 5.53 and the 10% max IV size of 16.38. We already alluded to this being a source of concern above.

4.7. *Have the good, the bad, and the ugly European colonial migration strategies been salient for non-European human development?*

The narrative to date is that following an increase in a country's share of Europeans in its population during colonisation, its current life expectancy rises, and both its fertility and infant mortality rates fall. Regarding economic development as measured by current income per capita, EL interpreted this as the diminishing marginal long-run development impact of European share. The possibility of a non-linear association between European share and our health variables are perhaps already suggested in Fig. 2, in which we recognize a U-shaped effect of European share on Longevity and hump-shaped effects of European share on Fertility and Mortality. We now conjecture that this influence could be due to the differences in diversity and homogeneity that different former colonies were exposed to, as Europeans implemented their different location-specific migration strategies.

Table 6: European share of colonial population and health outcomes (alternative estimation technique)

	IHCR only				External instrument only				IHCR with external instrument			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Longevity						
European share	0.351***	0.109	0.245***	0.194***	0.399***	0.204***	0.297***	0.224***	0.386***	0.210***	0.288***	0.229***
	(0.00)	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	124	114	112	107	102	92	92	87	102	92	92	87
R-squared	0.153	0.172	0.418	0.451	0.172	0.281	0.419	0.491	0.17	0.282	0.416	0.49
						Fertility						
European share	-1.009***	-0.225	-0.723***	-0.646***	-1.163***	-0.794***	-0.989***	-0.698***	-1.055***	-0.676***	-0.988***	-0.497**
	(0.00)	(0.52)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)
Observations	124	114	112	107	102	92	92	87	102	92	92	87
R-squared	0.196	0.162	0.374	0.439	0.172	0.326	0.371	0.476	0.166	0.316	0.368	0.467
						Mortality						
European share	-2.448***	-1.404**	-1.955***	-1.994***	-2.545***	-1.906***	-2.302***	-1.768***	-2.544***	-1.892***	-2.292***	-1.917***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	124	114	112	107	102	92	92	87	102	92	92	87
R-squared	0.244	0.24	0.437	0.471	0.218	0.304	0.413	0.471	0.218	0.302	0.413	0.463
						First-state regression						
KP F test	16.6	13.43	69.8	150.91	57.12	67.95	58.95	67.6	68.4	35.06	106.72	290.23
SY 10% max IV size	16.38	16.38	16.38	22.3	16.38	16.38	16.38	16.38	19.93	19.93	19.93	24.58
SY 25% max IV size	5.53	5.53	5.53	7.8	5.53	5.53	5.53	5.53	7.25	7.25	7.25	8.31

Notes: This table reports the IV estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (second panel) and Mortality (third panel). Columns (1), (5), and (9) control for British legal origin; columns (2), (6), and (10) control for Independence year; columns (3), (7), and (11) control for Ethnic diversity; and columns (4), (8), and (12) control for all three variables jointly. The results from the first-stage regressions are in the bottom panel. The KP F test is the Kliebergen and Paap (2006) rk Wald F statistic and corresponds to a test of the null of jointly weak instruments. SY are the Stock and Yogo (2005) critical values for i.i.d. errors. Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Variable definitions and sources are provided in the text.

There exists a plethora of research on the differential effects of population heterogeneity, especially as it relates to diversities of ethnicity, language, religion, culture, and genetics and their consequences for, or against, promoting innovation, fostering development, providing public goods, and forging social cohesion, as well as building trusting, cooperative, and healthy polity (e.g., Easterly and Levine 1997; Alesina et al. 1999; Alesina et al. 2003; Alesina and La Ferrara 2005; Ashraf and Galor 2013; Alesina et al. 2016). It is, therefore, unsurprising that pioneering European settlers in the former colonies manifested both costs and benefits to their host societies. We next formally appraise this likelihood for human development in non-European countries today using two approaches.

In the first instance, we introduce the quadratic term for European share in our baseline regressions, thereby estimating the following life-gifting model:

$$\ln y_i = \alpha + \beta_1 \text{ European share}_i + \beta_2 \text{ European share}_i^2 + \mathbf{C}_i' \gamma + \varepsilon_i. \quad (11)$$

Doing this helps to establish whether any European presence is sufficient to aid human development. In the first half of Table 7 (columns (1)-(4)), we show the estimates for both the linear and quadratic terms of Eq. (11) in models similar to those in the baseline regressions of Table 1. For both Longevity and Fertility, we find that both European share and European share^2 are statistically significant. Concerning Longevity, the estimated coefficient for European share is positive, while the estimated coefficient for European share^2 is negative. Meanwhile, these associations are reversed for Fertility. As for Mortality, the evidence obtained suggests that the linear European share term continues to be consistently negative, but the positive association with the quadratic European share term breaks down once Ethnic diversity is included in the regressions (columns (3)-(4)).[29]

[29] In unreported regression, we confirm that for Mortality, the effect of European share^2 may be running through Ethnic diversity, as it becomes positively statistically significant again when we drop Ethnic diversity from the model in column (4) but continues to include British legal origin and Independence year.

Table 7: European share of colonial population and health outcomes (differential migration policies)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Longevity			
European share	1.180***	1.156***	0.850***	0.890***				
	(0.00)	(0.00)	(0.00)	(0.00)				
European share squared	-1.067***	-1.132***	-0.747***	-0.819***				
	(0.00)	(0.00)	(0.01)	(0.01)				
High European share					0.166***	0.156***	0.127***	0.131***
					(0.00)	(0.00)	(0.00)	(0.00)
Moderate European share					0.0617*	0.0417	0.0610*	0.0388
					(0.09)	(0.3)	(0.05)	(0.24)
Low European share					-0.115***	-0.0908**	-0.0755*	-0.0637
					(0.00)	(0.03)	(0.06)	(0.12)
Optimal European share	0.55	0.51	0.57	0.54				
Confidence interval	0.49-0.68	0.46-0.62	0.48-0.90	0.46-0.83				
Observations	124	114	112	107	124	114	112	107
R-squared	0.21	0.283	0.446	0.489	0.263	0.312	0.484	0.513
					Fertility			
European share	-3.361***	-3.384***	-2.486***	-2.386***				
	(0.00)	(0.00)	(0.00)	(0.00)				
European share squared	2.975***	3.124***	2.058**	2.179**				
	(0.00)	(0.00)	(0.03)	(0.02)				
High European share					-0.524***	-0.535***	-0.428***	-0.383***
					(0.00)	(0.00)	(0.00)	(0.00)
Moderate European share					-0.203**	-0.136	-0.151*	-0.115

					(0.03)	(0.21)	(0.09)	(0.24)
Low European share					0.242***	0.179**	0.179**	0.147*
					(0.00)	(0.04)	(0.05)	(0.09)
Optimal European share	0.57	0.54	0.6	0.55				
Confidence interval	0.49-0.72	0.47-0.70	0.49-1.26	0.45-1.05				
Observations	124	114	112	107	124	114	112	107
R-squared	0.26	0.347	0.409	0.48	0.298	0.371	0.444	0.512
				Mortality				
European share	-5.259***	-5.199***	-4.124***	-3.621**				
	(0.00)	(0.00)	(0.00)	(0.01)				
European share squared	3.585***	3.698***	2.402	2.224				
	(0.00)	(0.01)	(0.12)	(0.18)				
High European share					-0.952***	-0.943***	-0.849***	-0.706***
					(0.00)	(0.00)	(0.00)	(0.00)
Moderate European share					-0.205	-0.0336	-0.185	-0.0523
					(0.23)	(0.87)	(0.25)	(0.77)
Low European share					0.481***	0.394**	0.27	0.245
					(0.00)	(0.03)	(0.13)	(0.18)
Optimal European share	0.73	0.7	0.86	0.81				
Confidence interval	0.61-1.17	0.58-1.21	$-\infty + \infty$	$-\infty + \infty$				
Observations	124	114	112	107	124	114	112	107
R-squared	0.267	0.307	0.453	0.485	0.257	0.298	0.435	0.48

Notes: This table reports the OLS estimates of the effect of European share on health outcomes, using cross-sectional data with one observation per country. The dependent variables are Longevity (top panel), Fertility (middle panel) and Mortality (bottom panel). Columns (1) and (5) control for British legal origin; columns (2) and (6) control for Independence year; columns (3) and (7) control for Ethnic diversity; and columns (4) and (8) control for all three variables jointly. Each regression includes a constant, but not reported. *p*-values are reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Confidence interval is for the Fieller 90% confidence region. Variable definitions and sources are provided in the text.

We also report the optimal European share of the colonial population needed to have a gainful impact on the modern-day health performance of non-European countries and a 90% Fieller confidence intervals for these estimates. In the three panels, the minimum optimal value for European share is 51% of the colonial population, which is rather puzzling, particularly when we consider that EL estimated that any European share greater than 4.8% is sufficient to have a positive impact on current income per capita. In our sample of non-European countries, 118 countries (95% of the 124 countries) have values of European share lower than 0.51, with the exceptions being the usual suspects of Australia, Canada, New Zealand, and the United States, as well as the Bahamas and Cuba. While this approach is informative in revealing that the relationship between European share and our health variables may be non-monotonic, it does not say much about whether it was necessary to have European explorers (exploiters) in the first place. It is this we turn to next.

Given the above evidence, we sense the need to scrutinize further the marginal contributions of a finer classification of European share distribution. In a second strategy, we follow Bruhn and Gallego (2012), who argued and empirically confirmed that the widely varying levels of development in the Americas are a consequence of the different economic activities that the colonizers engaged in different regions of the continent. Thus, we construct a discrete version of European share by splitting the sample into four groups: (i) No European share if European share is zero; (ii) Low European share if European share falls below the 33rd percentile of those countries with European share greater than zero; (iii) Moderate European share if European share falls between 33rd and 67th percentiles of those countries with European share greater than zero; and (iv) High European share if European share falls above the 67th percentile of those countries with European share greater than zero. Then, we create for each group a dummy variable that equals to one if a country's European share belongs to its classification and is zero otherwise.

Using these dummy indicators,[30] we more formally test for the consequences of the differential European colonial migration strategies:

$$\ln y_i = \alpha + \beta_h \text{ High European share}_i + \beta_m \text{ Moderate European share}_i + \beta_l \text{ Low European share}_i + \mathbf{C}_i' \gamma + \varepsilon_i \quad (12)$$

[30] We view our high, moderate, and low European share of the colonial population similarly to Bruhn and Gallego's (2012) three dummy variables that captured what they termed good, bad, and ugly colonial activities, with regards to the Americas based on studies by economic historians Engerman and Sokoloff (1997, 2002).

where No European share is the omitted category, such that one can interpret the estimated coefficients on the High, Moderate, and Low European share relative to countries with No European share. The second half of Table 7 (columns (5)-(8)) presents the results. As shown, High European share and Moderate European share yield similar results as the bundled European share estimates in the baseline specifications, being positive for Longevity and being negative with regards to both Fertility and Mortality. We, however, observe that Low European share is negatively correlated with Longevity and positively associated with both Fertility and Mortality.

Based on our full specification in column (8), we find that countries with High European share enjoy significantly greater Longevity today than countries with No European share, Moderate European share, and Low European share of 13.1%, 9.22%, and 19.5%, respectively. Importantly, the results further show that countries with Low European share experience lower Longevity today than countries with No European share of about 6.4%, though this effect is not statistically significant. This pattern is preserved in both the middle and bottom panels of Table 7, where Fertility and Mortality are respectively the dependent variables. Moreover, we notice that countries with Low European share have a statistically significant higher Fertility of 14.7% than countries with No European share.

Looking across columns (5)-(8) of Table 7, our findings indicate that, relative to No European share (countries that were either not explored historically or that successfully fought off European invaders), countries with Low European share have fared worse on our measures of health performance, while countries with High European share have done better. While similar to the evidence on High European share, the results concerning Moderate European share is less convincing. We submit that these results are consistent with Engerman and Sokoloff (1997) and Acemoglu et al. (2001). Therefore, we see that countries that experienced higher influx of colonial Europeans have better health prosperity nowadays than countries with lower inflow of colonial Europeans.

4.8. Further evidence

We bring this section to a close by summarising a number of additional results (not reported because of space but are available on demand). First, we concurrently drop countries with European share that is equal to zero and greater than 0.125. Although this exercise reduces the sample size from well over 100 (depending on the specification in Table 1) to just 51, our estimated coefficients retained their level of significance, while they appear to have undergone very appreciable increases of around 5-folds in most cases. Then, we run regressions similar to

our baseline specifications, but also control for European share equal to zero fixed effects, following EL—the estimated coefficients of the European share are qualitatively unaffected.

Second, we repeat the regressions of Table 1, but also control for European share of the current population, as measured in the year 2000 (Putterman and Weil 2010). Our primary objective has been to evaluate the plausibly enduring effect of European settlement outside of Europe during the colonial period on contemporary health outcomes. For this reason, we have aptly captured European settlement using the proportion of a country's population during colonisation that is European. It will, however, not be too big a stretch to ask whether it was the European ancestors or their descendants that mattered for prevailing health outcomes in non-European countries. In this exercise, our baseline results for European share are mostly retained, with a single exception in the regression of Longevity on European share, the European share of the current population, and British legal origin. We note that, while European share of the current population appears to have stronger effects both statistically and economically in the specification where Longevity is the dependent variable, European share continues to be of most influence when it comes to both Fertility and Mortality. All in all, we surmise that "old" European share still matters.

Third, our baseline results may still be biased if there are any known factors which could be correlated with European share and account for its estimated effect on current health outcomes. Thus, we next use the same sets of explanatory variables as in the baseline specifications,[31] but also successively add measures of early economic, political, and technological development (e.g., population density, distance to regional frontier, agricultural transition, state history, and technology adoption, with all measured at or prior to 1500 CE), as well as sequentially controlling for other geographical, climatic, historical, and disease environment covariates (e.g., temperature, precipitation, genetic distance to global frontier in 1500 CE, arable land area, soil fertility, distance to nearest waterway, ex-colony fixed effects, and percentage of population living in tropical zones). The results show that European share still performs well in terms of the sign, significance, and size, which implies that the association uncovered between European share and current health outcomes is robust to the inclusion of these additional observed country characteristics.

Fourth, we look at controls that are contemporaneous to our health outcomes (e.g., trade, government size, and investment size) and others that capture other dimensions of population heterogeneity (e.g., cultural diversity and religious and language fractionalizations). In all cases,

[31] The results correspond to the model specification in column (4) of Table 1, which regresses each health outcome on European share, British legal origin, Independence year, and Ethnic diversity. We have also repeated the regressions for the less structured models of columns (1)-(3) of Table 1 and find that our results hold.

our results on the importance of European share for health improvements are virtually unaffected.

Fifth, we have investigated the robustness of our baseline results to using European share of the colonial population measured over different periods: 1500-1800, 1801-1900, and 1500-1900. If our baseline estimates survive this experimentation, this extension helps to allay the concern that using a peculiar reference best year, or a range of years, in constructing countries' European share of the colonial population, may have exacerbated mismeasurement problem. We found that the baseline results are essentially unchanged with regards to the sign and significance, but the size of the coefficients are slightly smaller now.

Sixth, we had discussed above that EL's methodology for coding European share required that there should be at least 50 years between the reference year when the measure is computed for a country, and when it gained independence; however, due to data unavailability, this rule was not consistently applied. Thus, in a further robustness exercise, we adjusted EL's European share using Acemoglu at al.'s European settlers in 1900 for countries for which this rule is violated and their European share obtained post-1900.[32] The obtained estimates after this data splicing are largely in line with the reported values.

Seventh and finally, we estimate the effect of European share on an average composite measure of human development, which is being published by the Human Development Report Office (HDRO) of the United Nations Development Programme (UNDP 2018): Human Development Index (HDI). The HDI is a geometric mean of normalized indices for three critical dimensions of human development: (i) health, as captured by life expectancy at birth; (ii) education, as captured by the average years of schooling for adults aged 25 years or more and expected years of schooling for children of school entering age; and (iii) living standards, as captured by gross national income per capita. We find that European share is strongly linked to HDI in all specifications considered.

5. Concluding remarks

So, in answer to the question posed by the title of this paper—Where do people live longer?—we say in countries with ample European presence. To reach this conclusion, we have examined empirically whether present-time human development outside of Europe has benefitted from its colonial-period European subjugation and settlement. To advance this thesis, we have employed

[32] The 26 countries concerned are: Belize, Burundi, Cape Verde, Djibouti, Fiji, Gabon, Ghana, Guinea-Bissau, Kenya, Lesotho, Madagascar, Malawi, Morocco, Mozambique, Namibia, Nicaragua, Rwanda, Sao Tome and Principe, Senegal, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, and Zimbabwe.

a newly compiled historical data set by EL on the number of Europeans that migrated to, and settled in, other regions of the world as a proportion of the host countries' populations during the colonial epoch. This is an important contribution to the economic and historical literature on the fundamental and proximate causes of development. EL provide a measure of European settlers in non-European countries from the time period in which Europeans and non-Europeans first had contacts to when Europeans hegemonized the non-Europeans, while painstakingly avoiding the post-colonial times. Using this data, they show that European share (of the colonial population) is a robust intermediating channel between pre-colonial factors and differences in post-colonial income per capita across countries.

Our paper exploits this explanatory power of European share in accounting for current socio-economic development, offering new insights into the long-term effects of European settlements, with a particular focus on another dimension of development—human biological well-being. Notably, the main contribution has been to show that distant European migration strategies, as captured by EL's European share of the colonial population, are also a robust predictor of human development, as designated by Longevity, Fertility, and Mortality. Thus, our baseline finding is that, controlling for the influence of the United Kingdom's legal endowment (British legal origin), the fraction of years a country has been independent since 1776 (Independence year), and contemporaneous population heterogeneity (Ethnic diversity), countries with higher share of European migrants in colonial times tend to enjoy more health benefits in present times.

The effects are both economically and statistically meaningful and are found to be robust across time, to expunging outliers, to conditioning on an expanded set of observables, to evaluating the probable role of unobservable factors, to employing alternative estimation technique, to an explicit definition of what European colonial migration policies promoted, and to various alternative dates for measuring our health variables and European share. Moreover, to our best knowledge, this paper is the first to systematically use an actual share of Europeans as a fraction of a host colony's population during the colonial age to study a possible persistent effect of European colonial migration policy on current health outcomes of non-European countries.

To sum up, this paper has offered an important insight into the significant role European colonial settlements play in the post-colonial health performance of non-European countries. As it stands now, however, more cross-country research is needed in order to construct a convincing theory to explain *why* and *how* historical European share affects health outcomes today. The need to understand the causal mechanisms for this relationship is all the more

pertinent when we consider our finding that (while having high European share is the most beneficial) it is better for a country to have no European colonial settlers than a low one.[33] We therefore propose that an important task of future research will be that of uncovering plausible causal mechanisms, which take into account the within-country health implications of European share of the colonial population.

Declaration of Competing Interest

The author has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5), 1369-1401.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, 117(4), 1231-1294.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of Economic Growth*, 1, 385-472.

Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2009). Re-evaluating the modernization hypothesis. *Journal of Monetary Economics*, 56(8), 1043-1058.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4), 1243-1284.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155-194.

Alesina, A., & Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3), 762-800.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101-138.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113(1), 151-184.

Ashraf, Q., & Galor, O. (2013). The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review*, 103(1), 1-46.

Bairoch, P. (1993). Economy and World History: Myths and Paradoxes. Wheat sheaf, Brighton.

Barro, R. J. (1999). Determinants of democracy. *Journal of Political Economy*, 107(S6), S158-S183.

Barro, R., & McCleary, R. (2003). Religion and Economic Growth across Countries. *American Sociological Review*, 68(5), 760-781.

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2003). Law, endowments, and finance. *Journal of Financial Economics*, 70(2), 137-181.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). Identifying influential data and sources of collinearity. Regression Diagnostics.

[33] For instance, whilst the health measures for Asia are, on average, not as impressive as for the Americas and Oceania, the evidence suggests that Asia is healthier than Africa, even though Asia had 17.9% less European share of the colonial population. Clearly other factors are important in explaining the differences in the currently observed health of nations.

Bertocchi, G., & Canova, F. (2002). Did colonisation matter for growth?: An empirical exploration into the historical causes of Africa's underdevelopment. *European Economic Review*, 46(10), 1851-1871.
Bibles, C. A. E. (2008). The Holy Bible: English Standard Version. HarperCollins UK.
Bockstette, V., Chanda, A., & Putterman, L. (2002). States and markets: The advantage of an early start. *Journal of Economic Growth*, 7(4), 347-369.
Bruhn, M., & Gallego, F. A. (2012). Good, bad, and ugly colonial activities: do they matter for economic development?. *Review of Economics and Statistics*, 94(2), 433-461.
Buchanan, N. S., & Ellis, H. S. (1955). Approaches to economic development (No. 338.9 B8).
Cervellati, M., Naghavi, A., & Toubal, F. (2018). Trade liberalization, democratization, and technology adoption. *Journal of Economic Growth*, 23(2), 145-173.
Comin, D., Easterly, W., & Gong, E. (2010). Was the wealth of nations determined in 1000 BC?. *American Economic Journal: Macroeconomics*, 2(3), 65-97.
Comin, D., & Hobijn, B. (2004). Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics*, 51(1), 39-83.
Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives*, 20(3), 97-120.
De Montesquieu, C. [1748] (2015). Montesquieu: The spirit of the laws. Book on demand limited.
Diamond, J. (1997). Guns. Germs and Steel: The Fates of Human Societies, Vintage.
Dollar, D., & Kraay, A. (2003). Institutions, trade, and growth. *Journal of Monetary Economics*, 50(1), 133-162.
Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2), 755-776.
Easterly, W., & Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4), 1203-1250.
Easterly, W., & Levine, R. (2003). Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics*, 50(1), 3-39.
Easterly, W., & Levine, R. (2016). The European origins of economic development. *Journal of Economic Growth*, 21(3), 225-257.
Engerman, S. L., & Sokoloff, K. L. (1997). Factor endowments, institutions, and differential paths of growth among new world economies. How Latin America Fell Behind, 260-304.
Engerman, S. L., & Sokoloff, K. L. (2002). Factor endowments, inequality, and paths of development among new world economics (No. w9259). National Bureau of Economic Research.
Englebert, P. (2000). Pre-colonial institutions, post-colonial states, and economic development in tropical Africa. *Political Research Quarterly*, 53(1), 7-36.
Ertan, A., Fiszbein, M., & Putterman, L. (2016). Who was colonized and when? A cross-country analysis of determinants. *European Economic Review*, 83, 165-184.
Fenske, J. (2015). African polygamy: Past and present. *Journal of Development Economics*, 117, 58-73.
Frankel, J. A., & Romer, D. H. (1999). Does trade cause growth?. *American Economic Review*, 89(3), 379-399.
Galor, O. (2011). Unified growth theory. Princeton University Press.
Galor, O., & Moav, O. (2007). The neolithic origins of contemporary variations in life expectancy. *Available at SSRN 1012650*.
Gennaioli, N., La Porta, R., De Silanes, F. L., & Shleifer, A. (2014). Growth in regions. *Journal of Economic Growth*, 19(3), 259-309.
Giavazzi, F., & Tabellini, G. (2005). Economic and political liberalizations. *Journal of Monetary Economics*, 52(7), 1297-1330.
Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do institutions cause growth?. *Journal of Economic Growth*, 9(3), 271-303.

Greenaway, D., Morgan, W., & Wright, P. (2002). Trade liberalisation and growth in developing countries. *Journal of Development Economics*, 67(1), 229-244.
Grier, R. M. (1999). Colonial legacies and economic growth. *Public Choice*, 98(3-4), 317-335.
Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, 35(2-3), 517-526.
Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others?. *The Quarterly Journal of Economics*, 114(1), 83-116.
Hibbs, D. A., & Olsson, O. (2004). Geography, biogeography, and why some countries are rich and others are poor. *Proceedings of the National Academy of Sciences*, 101(10), 3715-3720.
Jones, E. L. (1981). The European Miracle: Environments. Economies and Geopolitics in the History of Europe and Asia. Cambridge: Cambridge University Press.
Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. (2004). A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene*, 70(5), 486-498.
Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1), 97-126.
Kremer, M. (1993). Population growth and technological change: One million BC to 1990. *The Quarterly Journal of Economics*, 108(3), 681-716.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1997). Legal determinants of external finance. *The Journal of Finance*, 52(3), 1131-1150.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1998). Law and finance. *Journal of Political Economy*, 106(6), 1113-1155.
Landes, D. (1998). The wealth and poverty of nations. Abacus. London.
Lange, M. K. (2004). British colonial legacies and political development. *World Development*, 32(6), 905-922.
Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67-80.
Lewis, W. A. (1955). The theory of economic growth. London: Alien and Unwin.
Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53(1), 69-105.
Manning, P. (1990). Slavery and African life: Occidental, oriental, and African slave trades (Vol. 67). Cambridge University Press.
Mokyr, J. (1990). The Lever of Riches: Technological Creativity and Economic Progress. Oxford University Press, USA.
Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development Economics*, 76(2), 293-323.
Morris, M. D. (1979). Measuring the condition of the world's poor: the physical quality of life index. Pergamon.
North, D. C. (1981). Structure and change in economic history. Norton.
North, D. C. (1990). Institutions, institutional change, and economic performance. Oxford University Press, USA.
North, D. C., & Thomas, R. P. (1973). The rise of the western world: A new economic history. Cambridge University Press.
Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, 123(1), 139-176.
Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, 1(1), 65-92.
Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1), 20-36.
Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7), 3221-52.

Ortega, F., & Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of International Economics*, *92*(2), 231-251.
Oyekola, O. (2019). Where Do People Live Longer?. *Available at SSRN 3475072*.
Oyekola, O. (2020). A cross-country analysis of the roles of border openness, human capital and legal institutions in explaining economic development. *Available at SSRN*.
Özak, Ö. (2018). Distance to the pre-industrial technological frontier and economic development. *Journal of Economic Growth*, 23(2), 175-221.
Putterman, L., & Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *The Quarterly Journal of Economics*, 125(4), 1627-1682.
Spolaore, E., & Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics*, 124(2), 469-529.
Spolaore, E., & Wacziarg, R. (2013). How deep are the roots of economic development?. *Journal of Economic Literature*, 51(2), 325-69.
Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. Identification and inference for econometric models: Essays in honour of Thomas Rothenberg, 109-120.
Tabellini, G. (2010). Culture and institutions: economic development in the regions of Europe. *Journal of the European Economic Association*, 8(4), 677-716.
UNDP. (various years). Human Development Report. New York: Oxford University press.
Weber, M. (1905). The Protestant ethic and the spirit of capitalism. Trans. Talcott Parsons.

Appendix

Table A.1: Descriptive statistics

	Mean	Std. Dev.	Min	Max	*N*
		Health outcomes			
Longevity, baseline: 1995-2005	4.14	0.17	3.66	4.39	124
Fertility, baseline: 1995-2005	1.26	0.45	0.3	2.04	124
Mortality, baseline: 1995-2005	3.56	0.88	1.1	4.95	124
Longevity, 1985-95	4.11	0.17	3.59	4.37	124
Fertility, 1985-95	1.46	0.42	0.45	2.09	124
Mortality, 1985-95	3.84	0.81	1.54	5.12	124
Longevity, 2005-15	4.2	0.14	3.86	4.42	122
Fertility, 2005-15	1.12	0.44	0.18	2.01	122
Mortality, 2005-15	3.23	0.87	0.79	4.68	124
Longevity, 1985-2015	4.15	0.15	3.72	4.39	124
Fertility, 1985-2015	1.28	0.42	0.32	2.03	124
Mortality, 1985-2015	3.54	0.84	1.2	4.89	124
Human development index, 1990-2005	0.57	0.16	0.24	0.89	123
		European settlements			
European share of the colonial population	0.066	0.16	0	0.9	124
European share, 2000	0.12	0.24	0	0.9	113
European share, 1500-1800	0.2	0.22	0.017	0.9	33
European share, 1801-1900	0.22	0.26	0.0011	0.97	32
European share, 1500-1900	0.2	0.23	0.0011	0.93	40
		Controls and instrument			
British legal origin	0.4	0.49	0	1	124
Independence year	0.26	0.32	0	1	114
Ethnic diversity	0.39	0.32	0	1	112
Latitude	0.2	0.12	0.011	0.67	124
Precious metals	0.29	0.46	0	1	124
London	5345	2268	1381	11359	117
Biogeography	-0.017	1.31	-1.02	3.79	82
Malaria ecology	5.16	7.31	0	31.5	113
Settler mortality	4.73	1.18	2.15	7.99	79
Dummy for European share = 0	0.47	0.5	0	1	124
Population in 1500 CE	12.4	2.57	5.85	18.5	120
Population density in 1500 CE	0.52	1.32	-3.82	3.84	120
Distance to regional frontier in 1500 CE	7.7	1.38	0	9.29	124
Genetic distance to Ethiopia in 1500 CE	2.34	0.61	0	3.34	121
Agricultural transition years in 1500 CE	34.4	24.3	0	100	98

State history in 1500 CE	0.28	0.32	0	1	98
Technology adoption in 1500 CE	0.36	0.26	0	0.88	98
Indigenous mortality	0.32	0.47	0	1	124
Former colonies	0.92	0.28	0	1	110
Arable land area	9.04	2.74	2.23	14.4	124
Percentage of population living in tropical zones	0.43	0.43	0	1	106
Land suitability for agriculture	-1.65	1.45	-5.86	-0.083	104
Soil fertility	-0.72	0.4	-1.82	-0.052	106
Percentage of land near a waterway	0.39	0.36	0	1	106
Mean distance to the nearest waterways	0.35	0.38	0.008	1.72	106
Land distance	0.11	0.22	0	0.99	98
Navigation distance	6.43	3.69	0.96	14.1	98
Temperature	3.75	0.18	2.53	3.89	122
Precipitation	4.26	1.05	1.07	5.94	122
Mean elevation	0.53	0.46	0.001	2.42	122
Terrain roughness	0.21	0.21	0.0035	1.24	122
Orientation of continental axis	1.31	0.58	0.62	3	78
Total factor productivity	-0.025	0.084	-0.22	0.35	67
Trade	4.23	0.51	3.01	5.88	121
Government size	-1.91	0.47	-3.08	-0.72	108
Investment	3.03	0.38	1.67	4.14	118
Market reform index	0.54	0.21	0	0.9	103
Tax	0.17	0.079	0.054	0.5	108
Population growth	1.99	1.06	-0.096	6.15	124
Cultural diversity	0.33	0.22	0	0.73	105
Religious fractionalization	0.44	0.25	0.0023	0.86	124
Language fractionalization	0.43	0.3	0.0021	0.92	116
Birthplace diversity	0.13	0.19	0.00007	0.86	79
Genetic diversity	0.72	0.032	0.63	0.77	79
Yellow fever present today fixed effects	0.63	0.49	0	1	79
Tuberculosis per 100,000 inhabitants	197	207	6.3	969	79
Subsistence mode	-0.031	0.17	-1.5	0	102