Online Appendix: State History and Economic Development: Evidence from Six Millennia

Oana Borcan
University of East Anglia

Ola Olsson
University of Gothenburg

Louis Putterman
Brown University

November 6, 2017

Appendix A - Variables Description and Data Sources

A1 State history variables

Statehist. The extended statehist is the normalized aggregate index of state history. This index is defined as the sum of all 50-year period state history scores, adjusted by a discount factor, divided by the maximum value of a discounted index, corresponding to a state with a score of 50 in every half century between 3500 BCE and 2000 CE The index can be calculated at various times (we also calculate it at 1500 CE and indicate accordingly where this is used in the paper) and using various discount rates to put more weight on recent history than on the more distant past. Throughout the paper we use the 1% discount rate.

Statehist 1-1950 CE This is the statehist computed according to the initial version of the index in Bockstette et al. (2002), considering only the period 1 – 1950 CE This is also a normalized index (with respect to a virtual state that would have full scores for every half century between 1 and 1950 CE). In this paper we use a discount factor of 1% for this index.

Ancestry – Adjusted Statehist of 1500 CE This is the average accumulated state history to 1500 CE of the year 1500 ancestors (by territory of residence at that time) of the year 2000 population of each country. It is obtained by adjusting the extended statehist index at 1500 CE by the migration matrix of Putterman and Weil (2010), as follows: for each country i in the sample, we use the matrix to identify the share of the current population that has ancestry that can be traced to the territory of country j; for each country j we multiply its statehist score accumulated at 1500 CE (discounted by 1% and normalized with respect to a state with full scores from 3500 BCE until 1500 CE) by a weight which is the share of the current population in country i identified to have roots in country j; the sum of all weighted statehist scores thus obtained across all j is the ancestry-adjusted statehist index at 1500 CE. Using this adjustment, we account for the state experience prior to 1500 CE of other territories, brought by post-1500 migrants into a land area defined by modern country borders, in addition to the state history of the country's own territory.

State Age. This variable, measured in millennia, represents the total amount of time elapsed from the first date (exact or approximate) when state experience is assigned a positive scores (the first date when the component S1 pertaining to the existence of a rule above tribal level is positive) until 2000 CE. In instances where we use historical outcomes in 1500 CE, we also recalculate and denote state age accordingly. State age does not discount the periods of state collapse (when scores revert to 0) incurred after the state emergence date.

A2 Historical controls

Agyears. This variable assembled by Putterman with Trainor (2006) records for each present-day country in a sample of 170 countries the number of millennia elapsed in 2000 CE. since the Neolithic transition of populations that lived on the territory of that country. The year of transition is assigned by cross-referencing expert opinions about the time when the population is a particular region covered more than half of their calorie intake from agriculture. In instances where we use historical outcomes in 1500 CE, we also recalculate and denote Agyears accordingly.

Origtime. This variable coded by Ahlerup and Olsson (2012) represents the time since initial uninterrupted settlement of anatomically modern humans (before 2000 CE) on the territories that now belong to modern-day countries. The variable was coded for 191 countries and the coding was based on Oppenheimer (2003) and Bradshaw Foundation (2007), as well as Encyclopedia Britannica (2007) for the island cases. In instances where we use historical outcomes in 1500 CE, we also recalculate and denote Origtime accordingly. Since the original settlements follow the paths of the migration routes out of Africa, the variable is correlated with the migration distance, and can therefore also be employed as a proxy for the latter.[1]

The Matrix of Migration since 1500 CE This matrix was developed by Putterman and Weil (2010) to describe the composition of the populations of modern-day countries in terms of ancestry at 1500 CE, before the migration flows of the colonial era. The matrix contains 165 rows (each row corresponding to a present-day country) and 165 columns (representing the same countries), where every cell records the percentage of current population in country on row i that traces its ancestry to the population in the source country on column j, such that the sum of all cells on each row is 1. In their paper, Putterman and Weil (2010) obtained ancestry-adjusted measures of Statehist (1-1950 CE) and Agyears, by multiplying each row to each one of the vectors containing the values of their variables of interest at 1500 CE for each country in their sample (which amounted to a sum over the values of the variables of interest of each source country by the corresponding share of the population with ancestry in other countries). We follow the same procedure in order to obtain the extended ancestry-adjusted statehist.

[1] In our sample the correlation coefficient between the time since original human settlement and the migration distance of Ashraf and Galor (2013) is -0.51, which indicates that the shorter the migration distance to a particular territory, the earlier the first human settlement.

A3 Outcomes variables

GDP per capita in 2000. Data in current US dollars, as reported by the World Bank.

Population density in 1 CE and 1500 CE. This variable is measured in number of individuals per square kilometer. The variable is retrieved from Ashraf and Galor (2013), who employ the population size data from McEvedy and Jones (1978), and the land area from World Development Indicators. Since the territorial unit employed in McEvedy and Jones (1978) is based on 1975 country borders, in some cases, the same value of the population density is assigned to contiguous present-day countries (that may have been part of the same constituency in 1975, such as Yugoslavia).

Urbanization in 1 CE This measure of urbanization was computed by Peregrine (2003) in the Atlas of Cultural Evolution (ACE). The variable takes three values: 1 if the largest settlement on the territory defined by the borders of a given modern-day country was under 100 persons, 2 if the largest settlement was between 100 and 399 persons and 3 for settlements larger than 400 persons. We retrieved this variable from Comin, Easterly and Gong (2010), where it also used previously.

Urbanization in 1500 CE The urbanization rate for 1500 A.D. comes from Acemoglu, Johnson and Robinson (2005) and it is calculated as the percentage of a country's urban area population (for cities with at least 5,000 inhabitants).

Technology Adoption in 1 CE These variables are three indices created by Comin, Easterly and Gong (2010), henceforth CEG. The index in 1 CE is based on data from Peregrine's (2003) "ACE" in which various cultural traits of 289 prehistoric cultures are evaluated: writing, agriculture, transportation, urbanization. CEG used this and additional data to code country - level data on technology adoption in five sectors: agriculture, transportation, communications, writing, and military. The authors structured the information in "ACE" into indicators that denoted the presence or absence of a technology within a certain sector and territory, which they then averaged over to create the sector specific technology adoption index between 0 and 1 (e.g. where "ACE" codes "technological specialization" by 1 for none, 2 for pottery and 3 for metalworks, CEG marked pottery and metalwork as the two potential technologies within the "industry sector" at the time, which they coded using a binary convention 1 – if technology is present and 0 if not; the average over all these dummies within every sector is the value of the technology adoption index for that sector; this average for the industry sector in this case would be 0 if neither technology was present, 0.5 if only one was present and 1 if both were present). Then, the overall adoption level, the variable that we use in this paper, for each country, was calculated as the average of the adoption levels across sectors.

Technology Adoption in 1500 CE. For the average technology adoption measure in 1500 CE, CEG (2010) used many different sources to summarize information on 20 technologies across 4 sectors excluding agriculture (for instance, for "Industry", the two possible technologies are "presence of iron" and "presence of steel"). For the latter they used a proxy based not on technology presence, but rather on which type of agriculture was the primary source on a particular territory – e.g. pastoralism, hand or plough cultivation, or none). As with overall technology adoption in 1 CE, the overall measure in 1500 CE is obtained by averaging over the scores for each sector.

Technology in 2000 CE. The technology measure in 2000 CE is constructed in CEG(2010), based on Comin, Hobijn and Rovito (2008) and it captures (one minus) the gap in the average intensity of technology adoption for every country with respect to the US (in terms of years of usage of each technology relative to the number of years since the invention of that technology) for ten technologies: electricity, internet, PC's, cell phones, telephones, cargo and passenger aviation, trucks, cars and tractors, in per capita terms. This measure is different from the measures for 1 and 1500 CE, since it also measures adoption along the intensive margin.

A4 Geographical variables

Absolute latitude. This is the absolute value of the country's centroid latitude. The variable was retrieved from the Portland Physical Geography dataset.

Distance to coast and river. This variable represents the mean distance to the nearest coastline or sea-navigable river, measured in kilometers. The variable was retrieved from the Portland Physical Geography dataset.

Mean elevation. The mean elevation above sea level is measured in meters. The variable was retrieved from the Portland Physical Geography dataset. The original source is NOAA's National Geophysical Data Center.

Land suitability. This is a measure of land suitability for agriculture, computed at country level by Michalopoulos (2012), based on grid-cell data reported by Ramankutty et al. (2002). For details on the construction of the original index, the reader is referred to Ramankutty et al (2002). The index includes information on ecological indicators of climate and soil suitability for agriculture (such as drowing degree days, evapotranspiration, soil carbon density and soil pH).

Percentage arable land. This measures the percentage of a modern-day country's area that is arable. The source is World Bank's World Development Indicators.

Temperature. This is a mean across the average monthly temperature over time (1961-1990) in 1-degree resolution grids within a country. This variable was retrieved from Ashraf and Galor (2013), whose source is the G-ECON project (Nordhaus 2006).

Precipitation. This is a mean across the average monthly precipitation over time (1961-1990) in 1-degree resolution grids within a country. This variable was retrieved from Ashraf and Galor (2013), whose source is the G-ECON project (Nordhaus 2006).

Malaria (percentage population at risk). This variable represents the level of risk of contracting malaria (measured by the percentage population in 1994 in areas of high risk of contracting malaria, times the share of cases in the country involving fatal species of P. Falciparum). The original data was constructed by Gallup and Sachs (2001).

Landlocked. This is a dummy variable equal to 1 if the country is landlocked.

Appendix B - Additional *Statehist* Coding Information and Illustrations

Regarding the question of which 159 countries are covered by the index, our coverage is the same as that of the last version of the Statehist data (see introduction and appendix of Statehist v3.1), which covers countries having populations of at least half a million people for which there are generally accepted data on GDP and a few related indicators. To code all components of the index in a manner that is consistent across periods, we relied mainly on information in the Encyclopedia Britannica Online. We surveyed the main articles on the history of the modern-day country (e.g. "History of Azerbaijan"), but also articles connected to events in its history (e.g. "Azerbaijan-historical region", "Ancient Iran: The Sasanian period"). There were a number of instances where information in Britannica was sparse, in which cases we surveyed alternative sources, such as books or journal articles treating individual cases. For instance, in the case of Afghanistan, in addition to the Encyclopedia Britannica entries, we consulted two books: "Ancient Cities of the Indus Valley Civilization," by Jonathan Kenoyer and "Bactria: An Ancient Oasis Civilization from the Sands of Afghanistan" by Giancarlo Ligabue and Sandro Salvatori. The complete description of coding decisions and sources for all country-period observations can be found in the online Data Coding Appendix accompanying the paper.

Table B1 illustrates an example of coding based on information from Encyclopedia Britannica, covering the period 450 BCE - 1 CE for the territory of modern-day Bosnia and Herzegovina.

Table B1: Coding example - the case of Bosnia and Herzegovina, 400 BCE – 1 CE

Year BCE	Government is above tribe level? (z^1_{it})	Government domestic? (z^2_{it})	What percentage of the territory is ruled? (z^3_{it})	Weight* ($w_{it\theta}$)	s_{it}
400-201	0.75	1	0.3	1	11.25
200-151	0.75	1	0.3	0.9	12.63
(cutoff at 155)	1	0.5	1	0.1	
150-1	1	0.5	1	0.72	25

Note: $w_{it\theta}$ = number of years between period ends and cutoff, or between two cutoffs, divided by 50.

CODING INFORMATION

400-200: (0.75, 1, .3). From the 4th century BCE, along with the coming of Celtic tribes in the area, the Illyrian tribes became gradually more politically cohesive. Sources recall the existence of early indigenous petty kingdoms in Illyria on the territory of present-day Albania only. We therefore mark the occasional Illyrian tribe alliances by z_{it}^1=.75.

200-151: (0.75, 1, .3) until 155 and (1, .5, 1) from 155 onwards. Delminium (on modern-day Bosnian territory) was taken by the Romans in 155 BCE, hence $z_{it}^2 = .5$. Most of the area of Bosnia was integrated in the Roman province Dalmatia, hence z_{it}^3). The score is [0.9(0.75, 1, .3) + 0.1(1, .5, 1)]*50.

150-1: (1, .5, 1). Bosnia was under Roman occupation.

The extended Statehist includes some coding exceptions not present in the 1-1950 CE Statehist:

1) The z_{it}^1 component (existence of a supra-tribal rule) may also take the value 0.5 to indicate radical uncertainty with respect to the existence of a supra-tribal rule on the territory of country i in period t. An example is the case of Somalia, which receives a score z_{it}^1 of 0.5 between 1500 BCE and 1 CE, when its territory is believed to have been part of the Kingdom of Punt.

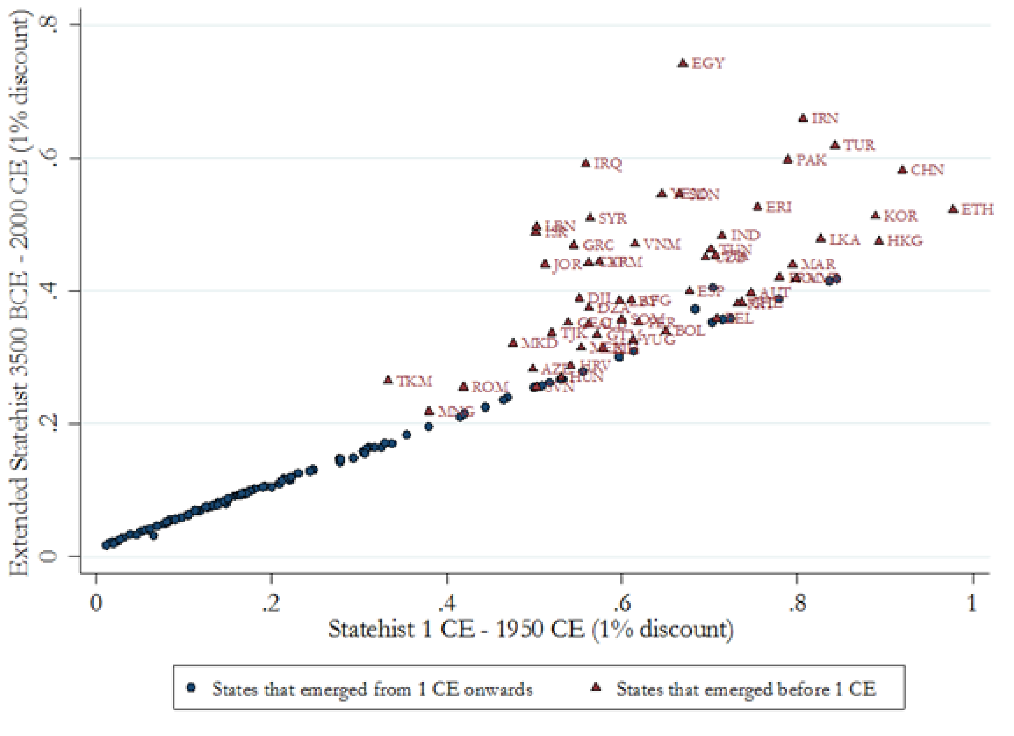
2) The z_{it}^1 component (existence of a supra-tribal rule) may also take the value 0.875 (an average of 0.75 and 1) to indicate the joint presence of a paramount chiefdom and a full-fledged state on the territory of country i. An example is the case of Armenia between 1275 and 840 BCE, when the Urartu (indigenous paramount chiefdom) and the Assyrian rule (full state) coexist on its territory.

3) The z_{it}^2 component (is the government imposed from within or from without) may take the value 0.875 (an average of 1 and 0.75) to indicate that the state is largely self-governed, with some foreign oversight. For instance, Austria qualifies for this score between 1945 and 1955, when although largely independent, it was closely observed by the Allied powers, which were still occupying some territories.

4) The z_{it}^2 component (is the government imposed from within or from without) may take the value 0.9375 (an average of 1 and 0.875) to indicate that most of the territory is governed from within, but some part of the territory is influenced or supported (at best nominally) by an external government. For instance, Cyprus receives this score from 1960 onwards, as the northern part of the island was declared the independent Turkish Republic of Northern Cyprus, recognized only by Turkey.

Figure B1 shows the comparison between the original old Statehist index for the CE-period and the new one presented in this paper. A notable feature is that the previous index failed to reflect the state history of several ancient civilizations like Egypt and Iran, countries that now receive a substantially higher score. Ethiopia, which had a full state by 1 CE but was a relative newcomer compared to those just mentioned, has the highest state history value going by the old measure, but now loses in relative terms to the older civilizations.

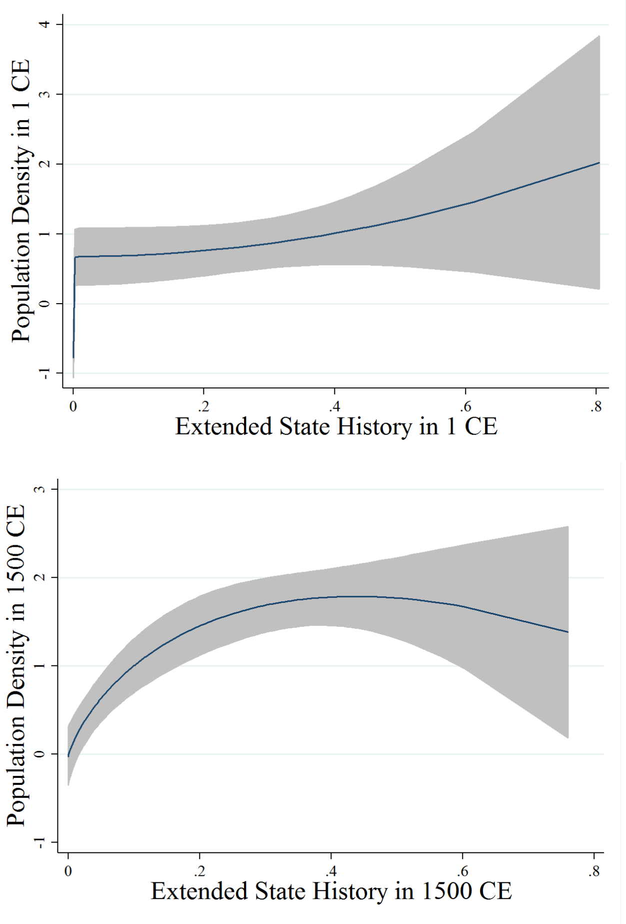
Figure B1: Extended Statehist (3500 BCE-2000 CE) vs Statehist 1-1950 CE



Note: The red, triangular observations with 3-letter isocodes show country observations where states emerged before 1 CE and whose index score changes considerably with the extended coding. Both variables use a 1 percent discount rate.

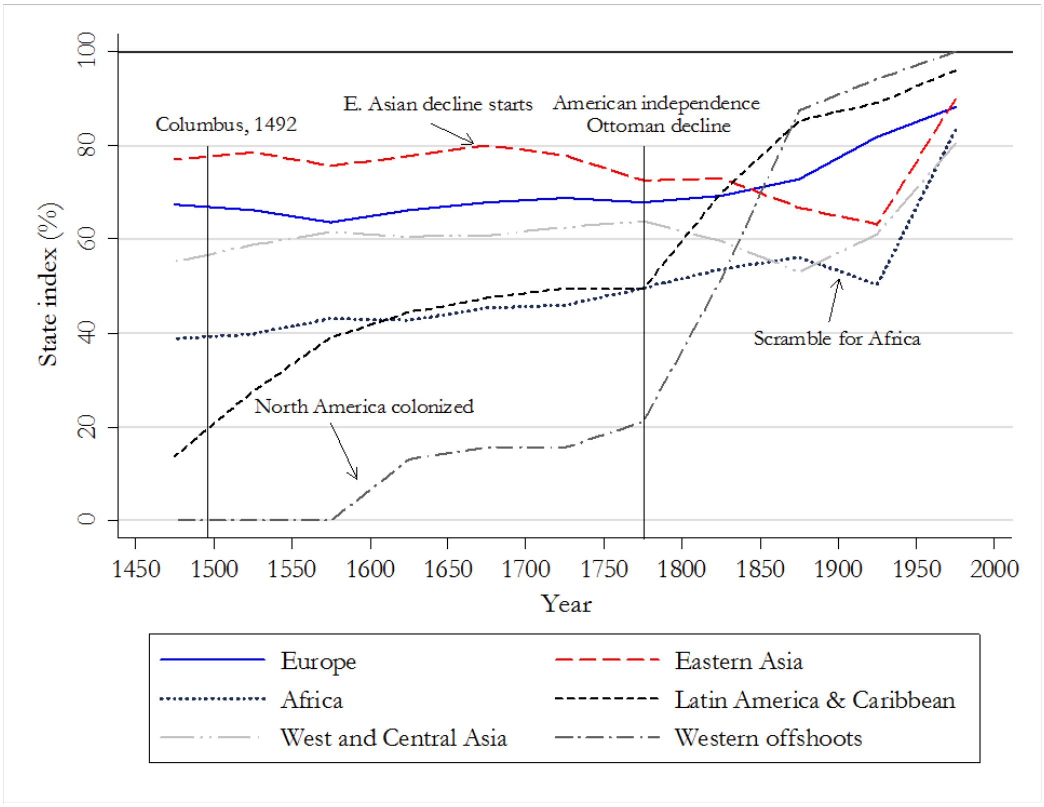
Appendix C - Supplementary Figures

Figure C1: Population density and state history in 1 CE and 1500 CE



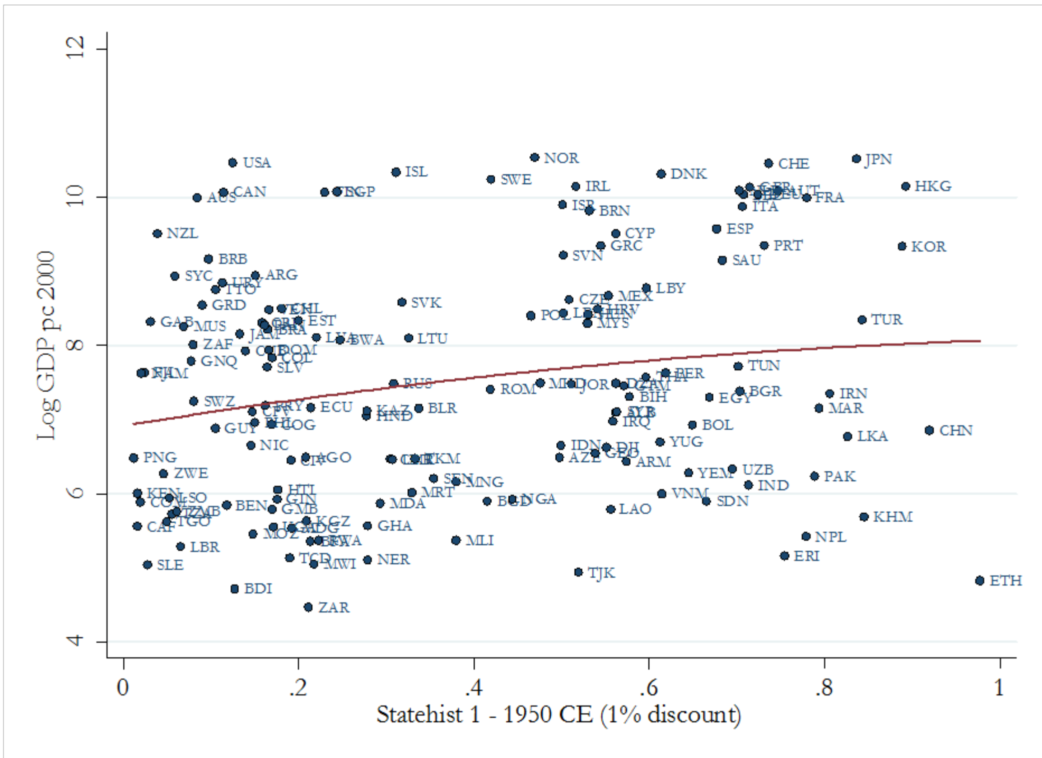
Note: The figure shows the predicted population density in 1 CE and 1500 CE fitted as a polynomial of extended *Statehist*. The shaded areas are the confidence intervals of the mean. Sources: see Appendix A

Figure C2: Emergence of states in six world regions during the colonial era, 1450-2000 CE



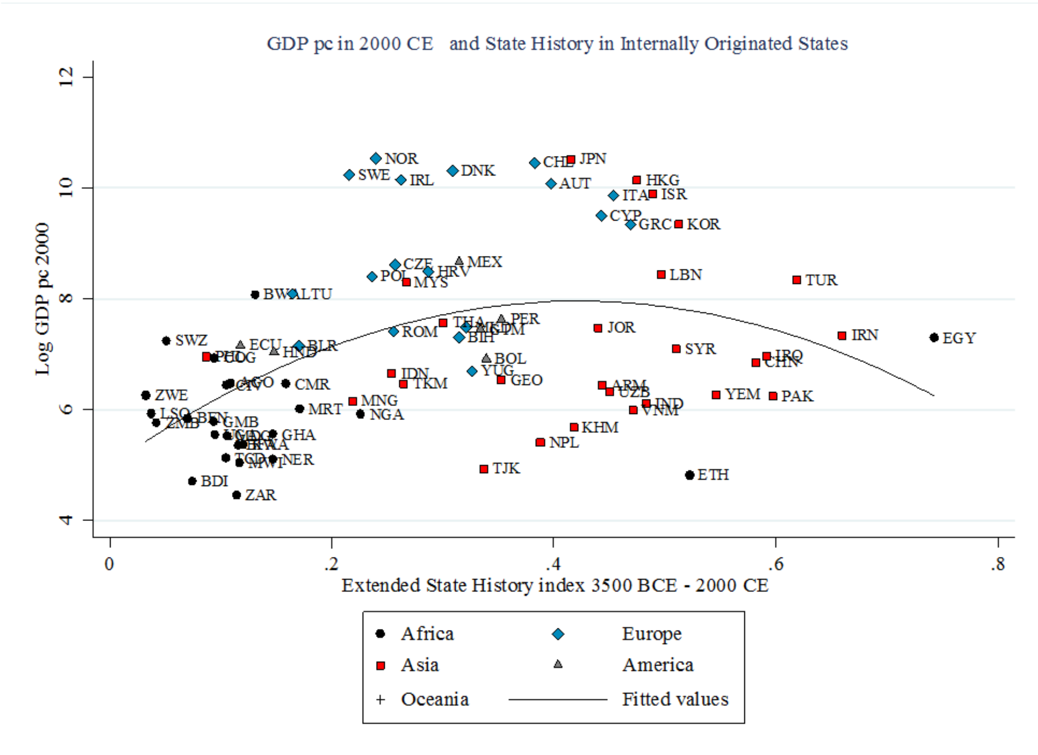
Note: The figure shows the development of the aggregated State index in Europe, Eastern Asia, West and Central Asia (including Turkey and India and the countries located in between), Latin America and the Caribbean (all countries in the Americas except Canada and USA), Africa (including North Africa), and the Western offshoots (USA, Canada, Australia, and New Zealand). Oceania is omitted. On the horizontal axis, negative values imply years BCE whereas positive values show the CE-period. Particular years with trend breaks are marked.

Figure C3: Non-linear relationship between Log GDP per capita in 2000 and Statehist 1-1950 CE



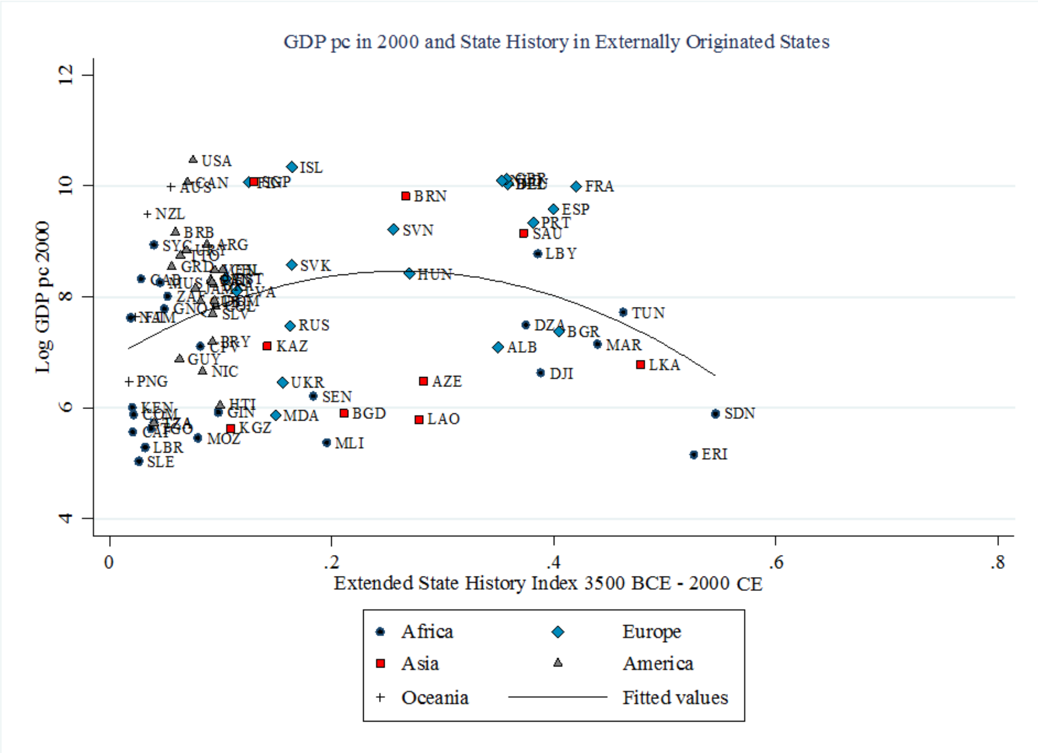
Note: The figure shows a fitted quadratic regression line corresponding to the estimates in Table 2, Panel B, column 2, with 154 country observations distinguished by 3-letter country isocodes.

Figure C4: GDP pc in 2000 CE and state history in internally-originated states



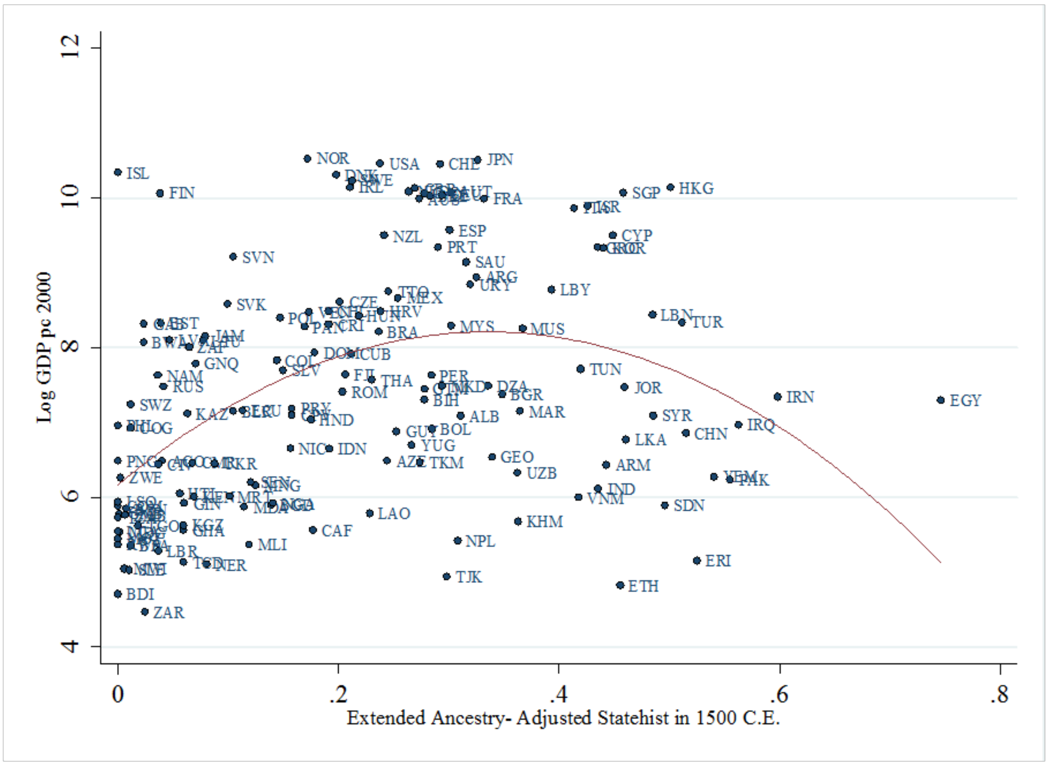
Note: The figure shows the relationship between Log GDP pc 2000 and Statehist, including a quadratic fit, for the subsample of countries where states were internally-originated.

Figure C5: GDP pc in 2000 CE and state history in externally-originated states



Note: The figure shows the relationship between Log GDP pc 2000 and Statehist, including a quadratic fit, for the subsample of countries where states were externally-originated.

Figure C6: GDP pc in 2000 CE and ancestry- adjusted state history



Note: The figure shows the relationship between Log GDP pc 2000 and ancestry-adjusted Statehist, including a quadratic fit, for all the countries in the sample.

Appendix D - Supplementary Tables

Table D1: Technology in 1500 CE and *Statehist* (different discount rates)

	(1)	(2)	(3)
	Technology Adoption in 1500 CE		
Statehist of 1500 CE (0.1% discount)	2.841***		
	(0.350)		
Statehist of 1500 CE squared (0.1% discount)	-3.359***		
	(0.738)		
Statehist of 1500 CE (1% discount)		2.643***	
		(0.328)	
Statehist of 1500 CE squared (1% discount)		-2.993***	
		(0.645)	
Statehist of 1500 CE (2% discount)			2.450***
			(0.301)
Statehist of 1500 CE squared(2% discount)			-2.601***
			(0.536)
Observations	112	112	112
R-squared	0.521	0.558	0.585
Controls	No	No	No
Continent FE	No	No	No

Note: The table presents OLS estimates of the relationship between Technology in 1500 CE and *Statehist*, from specifications similar to Table 2 in the paper, Panel A, column (2). Here we use the *Statehist* index computed using different discount rates. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D2: Technology in 1500 CE and *Statehist* (different discount rates): linear segment check of inverse-u relationship

	Technology Adoption in 1500 CE					
	< max	≥ max	< max	≥ max	< max	≥ max
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist of 1500 CE (0.1% discount)	1.813***	0.327				
	(0.175)	(0.484)				
Statehist of 1500 CE (1% discount)			1.540***	0.729		
			(0.135)	(0.452)		
Statehist of 1500 CE (2% discount)					1.333***	-0.284
					(0.151)	(0.289)
Observations	104	8	103	9	96	16
R-squared	0.462	0.049	0.504	0.158	0.477	0.023
Controls	No	No	No	No	No	No
Continent FE	No	No	No	No	No	No

Note: The table presents OLS estimates of the linear relationship between technology adoption in 1500 CE and *Statehist*, separately for the countries below and above the implied optimal level of *Statehist* (that is, the level at which the dependent variable ceases to rise and begins to fall with further increases in *Statehist*). The method is a test of whether the dependent variable exhibits both rising and declining segments, consistent with a hump-shaped or inverted-u relationship. We use the *Statehist* index computed using different discount rates. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D3: Population density and urbanization in 1500 CE and *Statehist* (different discount rates)

Panel A	Population Density in 1500 CE		
	(1)	(2)	(3)
Statehist of 1500 CE (0.1% discount)	9.559***		
	(1.666)		
Statehist of 1500 CE squared (0.1% discount)	-12.324***		
	(3.098)		
Statehist of 1500 CE (1% discount)		8.730***	
		(1.557)	
Statehist of 1500 CE squared (1% discount)		-10.385***	
		(2.663)	
Statehist of 1500 CE (2% discount)			7.513***
			(1.506)
Statehist of 1500 CE squared(2% discount)			-7.752***
			(2.433)
Observations	154	154	154
R-squared	0.254	0.269	0.276
Panel B	**Urbanization in 1500 CE**		
Statehist of 1500 CE (0.1% discount)	29.429***		
	(7.323)		
Statehist of 1500 CE squared (0.1% discount)	-25.531**		
	(12.514)		
Statehist of 1500 CE (1% discount)		25.808***	
		(7.147)	
Statehist of 1500 CE squared (1% discount)		-20.101*	
		(11.798)	
Statehist of 1500 CE (2% discount)			22.206***
			(6.909)
Statehist of 1500 CE squared(2% discount)			-14.926
			(10.820)
Observations	83	83	83
R-squared	0.311	0.323	0.329
Controls	No	No	No
Continent FE	No	No	No

Note: The table presents OLS estimates of the relationship between population density, or urbanization in 1500 CE and *Statehist*, from a specification similar to Table 3 in the paper, Panels A and B, column (2). Here we use the *Statehist* index computed using different discount rates. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D4: Population density in 1500 CE and *Statehist* (different discount rates): linear segment check of inverse-u relationship

	Population Density in 1500 CE					
	$<$ max	$\geq$ max	$<$ max	$\geq$ max	$<$ max	$\geq$ max
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist of 1500 CE (0.1% discount)	5.687***	-2.142				
	(0.937)	(1.512)				
Statehist of 1500 CE (1% discount)			4.996***	-3.083		
			(0.878)	(2.146)		
Statehist of 1500 CE (2% discount)					4.063***	-5.554**
					(0.704)	(2.633)
Observations	141	13	134	20	134	20
R-squared	0.213	0.057	0.212	0.068	0.214	0.140
Controls	No	No	No	No	No	No
Continent FE	No	No	No	No	No	No

Note: The table presents OLS estimates of the linear relationship between population density in 1500 CE and *Statehist*, separately for the countries below and above the implied optimal level of statehist (that is, the level at which the dependent variable ceases to rise and begins to fall with further increases in *Statehist*). We use the *Statehist* index computed using different discount rates. Similar checks cannot be run for urbanization, as the implied optimum level of *Statehist* is out of the observed range. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D5: GDP per capita and technology adoption in 2000 CE and *Statehist* (different discount rates)

Panel A	Log GDP per capita in 2000 CE		
	(1)	(2)	(3)
Statehist (0.1% discount)	6.571***		
	(2.331)		
Statehist squared (0.1% discount)	-10.351***		
	(3.903)		
Statehist (1% discount)		7.010***	
		(2.291)	
Statehist squared (1% discount)		-9.842***	
		(3.529)	
Statehist (2% discount)			6.099**
			(2.363)
Statehist squared (2% discount)			-7.126**
			(3.407)
Observations	154	154	154
R-squared	0.046	0.052	0.049
Panel B	**Technology Adoption in 2000 CE**		
Statehist (0.1% discount)	0.734**		
	(0.319)		
Statehist squared (0.1% discount)	-1.266**		
	(0.501)		
Statehist (1% discount)		0.842***	
		(0.318)	
Statehist squared (1% discount)		-1.285***	
		(0.452)	
Statehist (2% discount)			0.790**
			(0.318)
Statehist squared (2% discount)			-1.026**
			(0.417)
Observations	130	130	130
R-squared	0.038	0.044	0.039
Controls	No	No	No
Continent FE	No	No	No

Note: The table presents OLS estimates of the relationship between GDP per capita in 2000, or technology adoption in 2000, and *Statehist*, from a specification similar to Tables 4 and 5 (panel A) in the paper, column (2). Here we use the *Statehist* index computed using different discount rates. Robust standard errors in parentheses.***p<0.01,** p<0.05, * p<0.1

Table D6: GDP per capita, technology adoption in 2000 CE and *Statehist* (different discount rates): linear segment check of inverse-u relationship

Panel A	Log GDP per capita in 2000 CE					
	$<$ max	$\geq$ max	$<$ max	$\geq$ max	$<$ max	$\geq$ max
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist (0.1% discount)	5.023***	-1.757				
	(1.501)	(2.183)				
Statehist (1% discount)			2.234	-7.723***		
			(1.414)	(2.468)		
Statehist (2% discount)					1.835	-7.626***
					(1.252)	(2.665)
Observations	125	29	112	42	109	45
R-squared	0.084	0.012	0.020	0.157	0.018	0.120
Panel B	**Technology Adoption in 2000 CE**					
Statehist (0.1% discount)	0.468*	-0.510**				
	(0.243)	(0.240)				
Statehist (1% discount)			0.327	-0.631**		
			(0.242)	(0.236)		
Statehist (2% discount)					0.309	-0.464*
					(0.226)	(0.238)
Observations	99	31	85	45	80	50
R-squared	0.042	0.087	0.022	0.085	0.024	0.039
Controls	No	No	No	No	No	No
Continent FE	No	No	No	No	No	No

Note: The table presents OLS estimates of the linear relationship between GDP per capita or technology adoption in 2000 CE and *Statehist*, separately for the countries below and above the implied optimal level of *Statehist* (that is, the level at which the dependent variable ceases to rise and begins to fall with further increases in *Statehist*). We use the *Statehist* index computed using different discount rates. The results are very similar if we control for the time since the Neolithic transition. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D7: GDP per capita and technology adoption in 2000 CE and *ancestry-adjusted Statehist* (different discount rates)

Panel A	Log GDP per capita in 2000 CE		
	(1)	(2)	(3)
Ancestry-Adjusted Statehist (0.1% discount)	12.274***		
	(2.524)		
Ancestry-Adjusted Statehist squared (0.1% discount)	-19.432***		
	(5.477)		
Ancestry-Adjusted Statehist (1% discount)		12.888***	
		(2.183)	
Ancestry-Adjusted Statehist (1% discount)		-19.143***	
		(4.322)	
Ancestry-Adjusted Statehist (2% discount)			12.634***
			(1.811)
Ancestry-Adjusted Statehist squared (2% discount)			-17.046***
			(3.220)
Observations	148	148	148
R-squared	0.197	0.234	0.254
Panel B	**Technology Adoption in 2000 CE**		
Ancestry-Adjusted Statehist (0.1% discount)	1.179***		
	(0.299)		
Ancestry-Adjusted Statehist squared (0.1% discount)	-1.998***		
	(0.601)		
Ancestry-Adjusted Statehist (1% discount)		1.285***	
		(0.273)	
Ancestry-Adjusted Statehist squared (1% discount)		-2.024***	
		(0.505)	
Ancestry-Adjusted Statehist (2% discount)			1.304***
			(0.242)
Ancestry-Adjusted Statehist squared (2% discount)			-1.865***
			(0.405)
Observations	130	130	130
R-squared	0.112	0.140	0.154
Controls	No	No	No
Continent FE	No	No	No

Note: The table presents OLS estimates of the relationship between GDP per capita in 2000, or technology adoption in 2000, and *ancestry-adjusted Statehist*, from a specification similar to Tables 4 and 5 (panel A) in the paper, column (2). Here we use the *ancestry-adjusted Statehist* index computed using different discount rates. Robust standard errors in parentheses.***p<0.01,** p<0.05, * p<0.1

Table D8: GDP per capita, technology adoption in 2000 CE and *ancestry-adjusted Statehist* (linear segment check of inverse-u relationship

Panel A	Log GDP per capita in 2000 CE					
	$<$ max	$\geq$ max	$<$ max	$\geq$ max	$<$ max	$\geq$ max
	(1)	(2)	(3)	(4)	(5)	(6)
Ancestry-Adjusted Statehist (0.1% discount)	8.144***	-3.658				
	(1.459)	(2.522)				
Ancestry-Adjusted Statehist (1% discount)			8.450***	-1.450		
			(1.154)	(1.998)		
Ancestry-Adjusted Statehist (2% discount)					7.277***	-5.461**
					(1.057)	(2.465)
Observations	122	26	116	32	105	43
R-squared	0.222	0.043	0.325	0.007	0.322	0.088
Panel B	**Technology Adoption in 2000 CE**					
Ancestry-Adjusted Statehist (0.1% discount)	0.823***	-0.209				
	(0.199)	(0.241)				
Ancestry-Adjusted Statehist (1% discount)			0.847***	-0.406		
			(0.180)	(0.245)		
Ancestry-Adjusted Statehist (2% discount)					0.819***	-0.623**
					(0.186)	(0.254)
Observations	107	23	98	32	84	46
R-squared	0.134	0.016	0.191	0.049	0.222	0.089
Controls	No	No	No	No	No	No
Continent FE	No	No	No	No	No	No

Note: The table presents OLS estimates of the linear relationship between GDP per capita or technology adoption in 2000 CE and *ancestry-adjusted Statehist*, separately for the countries below and above the implied optimal level of *ancestry-adjusted Statehist* (that is, the level at which the dependent variable ceases to rise and begins to fall with further increases in *ancestry-adjusted Statehist*). We use the *ancestry-adjusted Statehist* index computed using different discount rates. The results are very similar (and the downward sloping part in panel B becomes significant) if we control for the time since the Neolithic transition. Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D9: Robustness checks - Log GDP per capita in 2000 and Statehist: additional controls and their squared terms.

	Log GDP per capita in 2000			
	(1)	(2)	(3)	(4)
Statehist	9.751***	6.647**	7.306**	7.306**
	(2.664)	(2.563)	(2.867)	(2.867)
Statehist squared	-6.976**	-4.685*	-4.791	-4.791
	(2.745)	(2.813)	(3.083)	(3.083)
State Age	-0.647***	-0.436**	-0.557***	-0.557***
	(0.192)	(0.183)	(0.212)	(0.212)
Agyears	-0.421**	0.006	-0.001	-0.001
	(0.166)	(0.080)	(0.091)	(0.091)
Agyears squared	0.042***			
	(0.016)			
Absolute Centroid Latitude		-0.016		
		(0.032)		
Absolute Latitude squared		0.001		
		(0.000)		
Distance from Addis Ababa			-0.064	
			(0.110)	
Distance from Addis Ababa squared			-0.001	
			(0.003)	
Predicted Genetic Diversity				22.403
				(73.470)
Predicted Genetic Diversity squared				-9.011
				(55.974)
Observations	125	125	125	125
R-squared	0.749	0.738	0.730	0.730
Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes

Note: The table presents OLS estimates of a specification similar to column 7 in Table 6, Panel A. Here we include additional controls and their squared terms. Columns (3) and (4) exclude *Origtime* (time since first human settlement) linear and squared from the set of controls. The variables "Distance from Addis Ababa" and "Predicted Genetic Diversity" are those constructed by Ashraf and Galor (2013). Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table D10: Robustness checks - Log GDP per capita in 2000 and ancestry-adjusted Statehist: additional controls and their squared terms.

	Log GDP per capita in 2000			
	(1)	(2)	(3)	(4)
Ancestry-Adjusted Statehist of 1500 CE	6.959***	6.450***	6.940***	6.940***
	(1.692)	(1.908)	(1.928)	(1.928)
Ancestry-Adjusted Statehist of 1500 CE squared	-7.759***	-7.479***	-7.503***	-7.503***
	(2.332)	(2.437)	(2.491)	(2.491)
State Age	-0.286**	-0.231	-0.331**	-0.331**
	(0.140)	(0.147)	(0.156)	(0.156)
Agyears	-0.288*	-0.024	-0.009	-0.009
	(0.153)	(0.082)	(0.088)	(0.088)
Agyears squared	0.025*			
	(0.014)			
Absolute Centroid Latitude		-0.034		
		(0.032)		
Absolute Latitude squared		0.001**		
		(0.000)		
Distance from Addis Ababa			0.007	
			(0.088)	
Distance from Addis Ababa squared			-0.003	
			(0.003)	
Predicted Genetic Diversity				71.693
				(60.609)
Predicted Genetic Diversity squared				-46.913
				(45.982)
Observations	125	125	125	125
R-squared	0.738	0.741	0.734	0.734
Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes

Note: The table presents OLS estimates of a specification similar to column (7) in Table 6. Here where we include additional controls and their squared terms. Columns (3) and (4) exclude *Origtime* (time since first human settlement) linear and squared from the set of controls. The variables "Distance from Addis Ababa" and "Predicted Genetic Diversity" are those constructed by Ashraf and Galor (2013). Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table D11: Robustness checks-Log GDP per capita in 2000 and Statehist: further region, colonial and legal origins controls

Panel A	Log GDP per capita in 2000			
	(1)	(2)	(3)	(4)
Statehist	6.790***	6.566***	6.786***	3.735*
	(2.496)	(2.193)	(2.510)	(2.042)
Statehist squared	-4.657*	-4.225*	-4.661*	-4.312*
	(2.776)	(2.348)	(2.793)	(2.259)
Colonial Dummy			0.031	
			(0.394)	
Observations	125	125	125	125
R-squared	0.734	0.730	0.734	0.833
Panel B	Log GDP per capita in 2000			
Ancestry-Adjusted Statehist	6.033***	6.977***	6.031***	4.569***
	(1.774)	(1.831)	(1.792)	(1.444)
Ancestry-Adjusted Statehist squared	-6.794***	-8.151***	-6.799***	-6.838***
	(2.332)	(2.353)	(2.332)	(2.021)
Colonial Dummy			0.011	
			(0.433)	
Observations	125	125	125	125
R-squared	0.732	0.734	0.732	0.842
Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	No	Yes	Yes
Region controls	No	Yes	No	No
Legal Origin Dummies	No	No	No	Yes
Sample	All	All	All	All

Note: The table displays OLS estimates analogous to Tables 5 (panel A) and 6, including additional controls. The dependent variable is log GDP per capita in 2000 CE and the main independent variables are the ancestry-adjusted Statehist index between 3500 BCE and 1500 CE, linear and squared. The data on colonies and legal origins was retrieved from Chanda, Cook and Putterman (2014), in turn taken from Acemoglu, Johnson and Robinson (2002) which derives these variables from La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999). Column (1) replicates column (7) in Table 6, panel A. Column (2) adds regional controls for the Middle East, North Africa and Sub-Saharan Africa. Column (3) includes a dummy for European colonies. Column (4) includes controls for legal origins. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, and percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D12: Robustness checks-Log GDP per capita in 2000 and Statehist: further sample restrictions

Panel A	Log GDP per capita in 2000				
	(1)	(2)	(3)	(4)	(5)
Statehist	6.790***	6.793***	12.138***	10.591**	-2.359
	(2.496)	(2.584)	(3.035)	(3.984)	(2.024)
Statehist squared	-4.657*	-3.095	-9.606***	-5.223	-2.058
	(2.776)	(2.902)	(3.275)	(5.472)	(1.993)
Observations	125	116	91	76	73
R-squared	0.734	0.756	0.714	0.756	0.835
Panel B	**Log GDP per capita in 2000**				
	(1)	(2)	(3)	(4)	(5)
Ancestry-Adjusted Statehist	6.033***	6.189***	9.254***	8.131**	0.534
	(1.774)	(1.810)	(2.277)	(3.247)	(1.875)
Ancestry-Adjusted Statehist squared	-6.794***	-5.000**	-9.683***	-5.189	-1.517
	(2.332)	(2.311)	(2.791)	(6.076)	(2.172)
Observations	125	116	91	76	73
R-squared	0.732	0.753	0.697	0.728	0.827
Controls	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes
Sample	All	No ME	No SSA	No SSA, ME or NA	European colonies

Note: The table displays OLS estimates analogous to Table 6, panel A, in different subsamples. The dependent variable is log GDP per capita in 2000 CE and the main independent variables are the ancestry-adjusted Statehist index between 3500 BCE and 1500 CE, linear and squared. The data on colonies and legal origins was retrieved from Chanda, Cook and Putterman (2014), in turn taken from Acemoglu, Johnson and Robinson (2002) which derives these variables from La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999). Column (1) replicates column (7) in Table 6, panel A. Column (2) presents analogous estimates in a sample excluding the Middle-east region countries. Column (3) presents analogous estimates in a sample excluding the Sub-Saharan African countries. Column (4) presents analogous estimates in a sample excluding Sub-Saharan Africa, North Africa and Middle-East countries. Column (5) presents analogous estimates in a sample excluding all countries except European colonies. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, and percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table D13: GDP per capita in 2000 CE and the Statehist components

Panel A	Log GDP per capita in 2000 CE			
	(1)	(2)	(3)	(4)
State Age	0.059	0.753**	0.121	0.151
	(0.081)	(0.289)	(0.216)	(0.241)
State Age squared		-0.148**	-0.040	-0.039
		(0.057)	(0.042)	(0.046)
Observations	154	154	125	125
R-squared	0.003	0.039	0.695	0.709
Panel B	Log GDP per capita in 2000 CE			
Average z^2 (1% discount)	-1.020	21.083***	11.408**	10.140**
	(0.998)	(7.484)	(5.030)	(4.915)
Average z^2 squared (1% discount)		-16.957***	-6.420	-5.374
		(5.884)	(3.913)	(3.786)
Observations	154	154	125	125
R-squared	0.008	0.067	0.745	0.758
Panel C	Log GDP per capita in 2000 CE			
Average z^3 (1% discount)	1.450**	7.432	3.610	5.145
	(0.601)	(5.064)	(3.592)	(4.045)
Average z^3 squared (1% discount)		-4.165	-2.433	-3.486
		(3.599)	(2.636)	(2.973)
Observations	154	154	125	125
R-squared	0.025	0.031	0.694	0.711
Panel D	Log GDP per capita in 2000 CE			
State Age	0.390***	1.494***	0.220	0.236
	(0.129)	(0.425)	(0.360)	(0.396)
State Age squared		-0.216***	-0.024	-0.012
		(0.067)	(0.057)	(0.063)
Average z^2 (1% discount)	-0.066	15.001**	11.549**	9.222*
	(1.086)	(7.585)	(5.150)	(5.172)
Average z^2 squared (1% discount)		-11.388*	-6.307	-4.477
		(6.005)	(3.975)	(3.969)
Average z^3 (1% discount)	3.767***	-0.932	-3.647	-2.183
	(0.866)	(5.211)	(3.863)	(4.171)
Average z^3 squared (1% discount)		3.620	2.862	2.068
		(4.127)	(3.083)	(3.295)
Observations	154	154	125	125
R-squared	0.083	0.180	0.748	0.760
Controls	No	No	Yes	Yes
Continent FE	No	No	No	Yes

Note: The table presents the estimated quadratic relationship between GDP per capita in 2000 and the separate components of the index: state age (years elapsed since first appearance of $z^1 > 0$, panel A), average z^2 conditional on a state existing (whether the rule was internally- or externally-based during given half centuries, panel B), and average z^3 conditional on a state existing (how much of the territory of the present-day country was subject to that rule, panel C). Note that state age is not identical to the average of z^1, since z^1 indicates whether a state existed in any given period, whereas state age adds all years since the first state, even in cases (uncommon, in practice) in which a reversion to statelessness occurred. OLS estimates corresponding to specifications similar to Table 5 (panel A, columns 1, 2, 5 and 7) in the paper. Robust standard errors in parentheses.***p<0.01,** p<0.05, * p<0.1

Table D14: State history and average technology adoption in 1 CE.

	Technology Adoption in 1 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist of 1 CE	0.735***	1.608***		0.465	0.209	0.503	0.807
	(0.109)	(0.275)		(0.381)	(0.513)	(0.342)	(0.544)
Statehist of 1 CE squared		-1.704***		-0.581	-0.397	-0.630	-0.823*
		(0.459)		(0.558)	(0.750)	(0.402)	(0.456)
Agyears in 1 CE			0.064***	0.054***	0.060***	0.023	0.031*
			(0.007)	(0.012)	(0.016)	(0.015)	(0.016)
Observations	128	128	128	128	124	124	124
R-squared	0.159	0.196	0.281	0.287	0.542	0.763	0.773
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table presents OLS estimates similar to Table 4, Panel A, where the dependent variable is the technology adoption index in 1 CE and the main independent variables are the Statehist index between 3500 BCE and 1 CE, linear and squared. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, and percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D15: Robustness checks: Technology adoption (excluding agriculture) in 1 CE and 1500 CE and statehist.

Panel A	Technology adoption in 1 CE excluding agriculture						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist of 1 CE	0.876***	1.909***		0.583	0.318	0.581	1.056*
	(0.128)	(0.324)		(0.454)	(0.559)	(0.379)	(0.625)
Statehist of 1 CE squared		-2.016***		-0.713	-0.531	-0.746*	-1.048**
		(0.544)		(0.665)	(0.826)	(0.449)	(0.525)
Agyears in 1 CE			0.075***	0.062***	0.067***	0.027	0.036*
			(0.008)	(0.014)	(0.017)	(0.017)	(0.019)
Observations	128	128	128	128	124	124	124
R-squared	0.170	0.208	0.293	0.300	0.578	0.759	0.769
Panel B	Technology adoption in 1500 CE excluding agriculture						
Statehist of 1500 CE	1.195***	2.776***		1.889***	1.607***	0.768***	1.029***
	(0.126)	(0.363)		(0.439)	(0.309)	(0.176)	(0.285)
Statehist of 1500 CE squared		-3.168***		-2.727***	-1.667***	-0.323	-0.207
		(0.700)		(0.806)	(0.560)	(0.310)	(0.293)
Agyears in 1500 CE			0.107***	0.070***	0.040***	-0.000	0.005
			(0.009)	(0.015)	(0.013)	(0.009)	(0.011)
Observations	111	111	109	109	107	107	107
R-squared	0.424	0.536	0.510	0.615	0.822	0.917	0.923
Controls	No	No	No	Yes	Yes	Yes	Yes
Continent FE	No	No	No	No	Yes	Yes	Yes

Note: The table presents OLS estimates similar to Table 2, where the dependent variable is the technology adoption index in 1 CE (panel A) and in 1500 CE (panel B), calculated to exclude the agriculture technology component. The main independent variables are the Statehist index between 3500 BCE and 1 CE (panel A) and between 3500 BCE and 1500 CE (panel B) linear and squared. The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, and percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table D16: State history, Log Population Density and Urbanization in 1 CE.

Panel A	Log Population Density in 1 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist of 1 CE	4.135***	7.856***		-0.610	2.129	2.810	4.487
	(0.729)	(1.800)		(2.121)	(2.601)	(1.771)	(2.923)
Statehist of 1 CE squared		-7.488***		-0.007	-2.153	-3.012	-3.926
		(2.770)		(3.232)	(4.232)	(2.392)	(2.909)
Agyears in 1 CE			0.455***	0.481***	0.436***	0.452***	0.460***
			(0.040)	(0.063)	(0.079)	(0.080)	(0.084)
Observations	135	135	130	130	115	115	115
R-squared	0.162	0.186	0.455	0.457	0.719	0.801	0.804
Panel B	**Urbanization in 1 CE**						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist of 1 CE	1.102***	2.726***		0.817	0.637	1.286	1.547
	(0.220)	(0.623)		(0.706)	(1.038)	(0.902)	(1.595)
Statehist of 1 CE squared		-3.172***		-1.289	-1.083	-1.260	-1.426
		(1.035)		(1.111)	(1.461)	(1.087)	(1.245)
Agyears in 1 CE			0.101***	0.089***	0.069**	-0.041	-0.038
			(0.020)	(0.027)	(0.034)	(0.039)	(0.039)
Observations	128	128	128	128	125	125	125
R-squared	0.069	0.094	0.139	0.143	0.373	0.527	0.527
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table presents OLS estimates similar to Table 3, where the dependent variable is log population density in 1 CE (panel A) and urbanization rate in 1 CE (panel B). The main independent variables are the *Statehist index* between 3500 BCE and 1 CE, linear and squared. The data for historical population density is based on population data from McEvedy and Jones(1978) and land data from World Bank's World Development Indicators. The data for urbanization rate in 1 CE is taken from Comin, Easterly and Gong (2010) and is based on Peregrine (2003) and takes three values: 1 if the largest settlement is smaller than 100 persons; 2 if it is between 100 and 399 persons; and 3 if it is larger than 400 persons. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, and percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table D17: Statehist and Ancestry-Adjusted Statehist and Log GDP pc 2000. Spatially correlated residuals.

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	1.482	8.033**		7.643**	3.868**	4.137*	7.027***
	(1.143)	(3.703)		(3.649)	(1.945)	(2.174)	(2.447)
Statehist squared		-11.183**		-10.548**	-4.717*	-4.986*	-4.038
		(5.429)		(5.179)	(2.754)	(2.775)	(2.761)
Agyears			0.124	0.030	-0.071	0.013	0.127
			(0.077)	(0.084)	(0.057)	(0.097)	(0.096)
Origtime					0.002	0.017***	0.019
					(0.003)	(0.004)	(0.013)
Origtime squared							-0.000
							(0.000)
State Age							-0.650***
							(0.185)
Observations	145	145	141	141	125	125	125
Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry-Adjusted Statehist of 1500 CE	2.603**	12.928***		12.882***	6.016***	5.532***	6.945***
	(1.177)	(3.032)		(2.811)	(1.834)	(1.890)	(1.827)
Ancestry-Adjusted Statehist of 1500 CE squared		-19.869***		-19.221***	-8.970***	-8.002***	-6.957***
		(5.644)		(5.379)	(2.478)	(2.777)	(2.653)
Agyears			0.127	-0.007	-0.054	0.034	0.139
			(0.077)	(0.078)	(0.050)	(0.089)	(0.087)
Origtime					0.003	0.015***	0.014
					(0.003)	(0.004)	(0.011)
Origtime squared							-0.000
							(0.000)
State Age							-0.438***
							(0.153)
Observations	142	142	140	140	125	125	125
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table displays OLS estimates from regressions of log per capita GDP in 2000 CE on *Statehist*, linear and squared. In panel A we have the *extended Statehist index*. In panel B the main independent variable is the *Ancestry-Adjusted Statehist index*. The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Standard errors using the Conley(1999) correction in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table D18: Statehist and Ancestry-Adjusted Statehist and Log GDP pc 2000. Spatial-autoregressive model.

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	0.936	6.280**		6.705**	4.294**	4.678**	6.774***
	(0.789)	(2.875)		(3.028)	(1.941)	(2.063)	(2.221)
Statehist squared		-8.864*		-8.376*	-5.596*	-5.364*	-4.758
		(4.589)		(4.591)	(2.985)	(3.147)	(3.091)
Agyears			0.055	-0.045	-0.043	-0.047	0.017
			(0.064)	(0.095)	(0.062)	(0.084)	(0.088)
Origtime					0.003	0.008**	0.010
					(0.003)	(0.004)	(0.010)
Origtime squared							-0.000
							(0.000)
State Age							-0.450**
							(0.183)
Observations	145	145	141	141	125	125	125
Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry-Adjusted Statehist of 1500 CE	2.314***	12.208***		13.091***	6.019***	5.514***	6.109***
	(0.767)	(2.044)		(2.087)	(1.635)	(1.709)	(1.783)
Ancestry-Adjusted Statehist of 1500 CE squared		-18.928***		-17.568***	-9.282***	-8.089***	-7.331**
		(3.670)		(3.650)	(2.772)	(2.954)	(2.983)
Agyears			0.053	-0.153*	-0.031	-0.020	0.012
			(0.063)	(0.080)	(0.059)	(0.083)	(0.087)
Origtime					0.004	0.006	0.007
					(0.003)	(0.004)	(0.010)
Origtime squared							-0.000
							(0.000)
State Age							-0.211
							(0.157)
Observations	142	142	140	140	125	125	125
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table displays SAR estimates from regressions of log per capita GDP in 2000 CE on *Statehist*, linear and squared. All results are maximum likelihood estimates based on Drukker, Prucha and Raciborski (2013) allowing for first-order spatial autocorrelation in the dependent variable. SAR models are $\mathbf{y} = \lambda \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, where $\mathbf{u}$ is a vector of i.i.d. error terms, $\mathbf{W}$ is a $n \times n$ spatial weighting matrix, n is the sample size, λ is a scalar, and $\mathbf{W}\mathbf{y}$ is a $n \times 1$ vector representing spatial lags in the dependent variable. We estimate a model where $\mathbf{W}$ is a matrix with diagonal elements equal to zero, and off-diagonal elements representing the inverse great-circle distances between geodesic centroids. In panel A we have the *extended Statehist index*. In panel B the main independent variable is the *Ancestry-Adjusted Statehist index*. The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. *** p<0.01, ** p<0.05, * p<0.1

Table D19: Statehist and Ancestry-Adjusted Statehist and Log GDP pc 2000. Spatial-autoregressive model with spatial-autoregressive disturbances

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	0.257	0.337		0.615	2.198	1.824	2.947
	(0.744)	(2.310)		(2.330)	(1.834)	(1.936)	(2.215)
Statehist squared		-0.128		0.017	-3.974	-2.419	-2.744
		(3.488)		(3.492)	(2.696)	(2.879)	(2.855)
Agyears			-0.023	-0.062	0.057	-0.059	-0.054
			(0.074)	(0.092)	(0.074)	(0.091)	(0.093)
Origtime					0.004	0.006	0.019*
					(0.004)	(0.004)	(0.011)
Origtime squared							-0.000
							(0.000)
State Age							-0.169
							(0.167)
Observations	145	145	141	141	125	125	125
Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry-Adjusted Statehist of 1500 CE	1.368*	6.434***		6.252***	4.150**	3.279*	4.038**
	(0.760)	(2.108)		(2.066)	(1.660)	(1.743)	(1.837)
Ancestry-Adjusted Statehist of 1500 CE squared		-8.805**		-7.765**	-7.428***	-5.341*	-7.043**
		(3.442)		(3.415)	(2.620)	(2.829)	(2.739)
Agyears			-0.021	-0.079	0.065	-0.023	-0.033
			(0.077)	(0.084)	(0.070)	(0.090)	(0.087)
Origtime					0.003	0.005	0.030***
					(0.004)	(0.004)	(0.011)
Origtime squared							-0.000**
							(0.000)
State Age							-0.055
							(0.126)
Observations	142	142	140	140	125	125	125
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table displays SARAR estimates from regressions of log per capita GDP in 2000 CE on *Statehist*, linear and squared. All results are maximum likelihood estimates based on Drukker, Prucha and Raciborski (2013) allowing for first-order spatial autocorrelation in the dependent variable and disturbances. SARAR models are $\mathbf{y} = \lambda\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, with $\mathbf{u} = \rho\mathbf{M}\mathbf{u} + \boldsymbol{\epsilon}$, where $\mathbf{W}$ and $\mathbf{M}$ are $n \times n$ spatial weighting matrices, n is the sample size, λ and ρ are scalars, and $\mathbf{Wy}$ and $\mathbf{Mu}$ are $n \times 1$ vectors representing spatial lags. We estimate a model where $\mathbf{W}$=$\mathbf{M}$ is a matrix with diagonal elements equal to zero, and off-diagonal elements representing the inverse great-circle distances between geodesic centroids. In panel A we have the *extended Statehist index.* In panel B the main independent variable is the *Ancestry-Adjusted Statehist index.* The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. *** p<0.01, ** p<0.05, * p<0.1

Table D20: Statehist and Ancestry-Adjusted Statehist and Log GDP pc 2000. Spatial-autoregressive model with spatial-autoregressive disturbances with proximity in terms of shared legal origins

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	1.068	5.614**		5.700**	4.266***	3.496**	4.952***
	(0.683)	(2.588)		(2.810)	(1.611)	(1.658)	(1.903)
Statehist squared		-7.816*		-9.171**	-6.546***	-4.783*	-4.837*
		(4.235)		(4.269)	(2.393)	(2.478)	(2.550)
Agyears			0.115**	0.106	0.003	-0.084	-0.037
			(0.052)	(0.081)	(0.047)	(0.057)	(0.067)
Origtime					-0.002	0.003	0.004
					(0.003)	(0.003)	(0.003)
State Age							-0.224
							(0.152)
Observations	154	154	147	147	125	125	125
Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry-Adjusted Statehist	2.866***	12.066***		10.464***	5.707***	4.692***	5.057***
of 1500 CE	(0.667)	(1.856)		(2.013)	(1.328)	(1.338)	(1.392)
Ancestry-Adjusted Statehist		-17.694***		-10.217***	-9.697***	-7.394***	-7.074***
of 1500 CE squared		(3.355)		(3.288)	(2.217)	(2.276)	(2.295)
Agyears			0.134***	-0.245***	0.012	-0.085	-0.060
			(0.051)	(0.039)	(0.042)	(0.056)	(0.062)
Origtime					-0.001	0.001	0.001
					(0.003)	(0.003)	(0.003)
State Age							-0.111
							(0.121)
Observations	148	148	144	144	125	125	125
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The table displays SARAR estimates from regressions of log per capita GDP in 2000 CE on *Statehist*, linear and squared. All results are maximum likelihood estimates based on Drukker, Prucha and Raciborski (2013) allowing for first-order spatial autocorrelation in the dependent variable and disturbances. SARAR models are $\mathbf{y} = \lambda \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, with $\mathbf{u} = \rho \mathbf{M}\mathbf{u} + \boldsymbol{\epsilon}$, where $\mathbf{W}$ and $\mathbf{M}$ are $n \times n$ spatial weighting matrices, n is the sample size, λ and ρ are scalars, and $\mathbf{Wy}$ and $\mathbf{Mu}$ are $n \times 1$ vectors representing spatial lags. We estimate a model where $\mathbf{W}$=$\mathbf{M}$ is a contiguity matrix of proximity in terms of shared legal origins: diagonal elements are zero, and off-diagonal elements (i,j) are 1 if countries i and j share the same legal origins, and 0 otherwise. In panel A we have the *extended Statehist index*. In panel B the main independent variable is the *Ancestry-Adjusted Statehist index*. The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. The squared term of Origtime is omitted due to computational restrictions. *** p<0.01, ** p<0.05, * p<0.1

References

[1] Acemoglu, D., Johnson, S. and Robinson, J. (2005). “Institutions As the Fundamental Cause of Long Run Economic Growth” in Handbook of Economic Growth, Aghion, P. and Durlauf, S. eds., North-Holland, Amsterdam.

[2] Acemoglu, D., Johnson, S. and Robinson, J. (2002). “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution ” The Quarterly Journal of Economics, Vol. 117, No. 4 (Nov., 2002), pp. 1231-1294

[3] Ahlerup, P. and Olsson, O. (2012). “The roots of ethnic diversity” Journal of Economic Growth, Springer, vol. 17(2), pp. 71-102, June.

[4] Ashraf, Q. and Galor, O. (2013). “The ‘Out of Africa’ Hypothesis, Human Genetic Diversity, and Comparative Economic Development” American Economic Review, 103(1), pp. 1-46.

[5] Bockstette, V., Chanda, A. and Putterman, L. (2002). “States and Markets: The Advantage of an Early Start” Journal of Economic Growth, Springer, vol. 7(4), pp. 347-69, December.

[6] Bradshaw Foundation (2007), available online at http://www.bradshawfoundation.com/stephenoppenheimer/.

[7] Chanda, A., Cook, A. and Putterman, L. (2014). “Persistence of Fortune: Accounting for Population Movements, There was No Post-Columbian Reversal,” *American Economic Journal – Macroeconomics*, 6(3), pp. 1 – 28.

[8] Comin, D., Easterly, W. and Gong, E. (2010). “Was the Wealth of Nations Determined in 1000 BC?” American Economic Journal: Macroeconomics, American Economic Association, vol. 2(3), pp. 65-97, July.

[9] Comin D., Hobijn, B. and Rovito, E., (2008). “A new approach to measuring technology with an application to the shape of the diffusion curves,” The Journal of Technology Transfer, Springer, vol. 33(2), pages 187-207, April.

[10] Conley, T. G. 1999. “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 92(1): 1 – 45.

[11] Drukker, D. M., Prucha, I. R. and Raciborski, R. (2013). “Maximum Likelihood and Generalized Spatial Two-Stage Least-Squares Estimators for a Spatial-Autoregressive Model with Spatial-Autoregressive Disturbances,” *Stata Journal*, 13(2), pp. 221 – 241.

[12] Gallup, J. L., and Sachs., J. D. (2001). “The Economic Burden of Malaria,” The American Journal of Tropical Medicine and Hygiene, 64(1-2): 8596.

[13] McEvedy, C., and Jones, R.. (1978). “Atlas of World Population History,” New York, NY:Penguin Books Ltd.

[14] Michalopoulos, S. (2012). “The Origins of Ethnolinguistic Diversity,” American Economic Review, 102(4): 1508-39.

[15] Nordhaus, W. D. (2006). “Geography and Macroeconomics: New Data and New Findings,” Proceedings of the National Academy of Sciences, 103(10): 35103517.

[16] La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and R. W. Vishny (1999). Unpublished appendix from "The Quality of Government," Journal of Law, Economics and Organization, XV (1999), 222-279.

[17] Oppenheimer, S. (2003). "Out of Eden: The peopling of the World," London: Constable.

[18] Peregrine, P. N. (2003). "Atlas of Cultural Evolution," World Cultures: Journal of Comparative and Cross-Cultural Research, 14(1): 175.

[19] Putterman, L. with C.A Trainor (2006) "Agricultural Transition Year - Country Data Set," http://www.econ.brown.edu/fac/louis putterman/.

[20] Putterman, L. and Weil, D. N. (2010). "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," The Quarterly Journal of Economics, MIT Press, vol. 125(4), pp. 1627-1682, November.

[21] Ramankutty, N., Foley, A. J., Norman, J. and McSweeney, K. (2002). "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," Global Ecology and Biogeography, 11(5): 377392.