# Subnational Persistence of State History: Evidence from Geolocalized Civilizations.*

C. Justin Cook
Tulane University
Murphy Institute and Dept. of Economics

January 2023

**Abstract**

This paper uses a novel data set that geolocalizes civilizations across the globe and human history. The data document roughly 1,450 civilizations over 500 periods between 3200 BCE and 2006 CE. Geolocalized data are then spatially aggregated to measure state history for pre-defined grids. Grid-cells with a prolonged history of containing a civilization have larger contemporary populations and increased nighttime lights. The state history data also allow for heterogeneity tests that have not been possible previously. My results show the sub-national relationship between state history and contemporary development is unaltered by country-level characteristics, suggesting an alternative mechanism is at play.

*JEL-Classification*: O11, O43, N10, R12
*Keywords*: State history, persistence, civilizations, GIS, development, institutions, nighttime lights

*I am grateful to Andrew Tollefson for the initial set of maps and sources. Extraordinary work in digitizing these maps was done by Erin Mutch, Amy Newsam, and the SPARC lab at UC-Merced. Helpful suggestions were provided by Dan Keniston, Ketki Sheth, Stelios Michalopoulos, the LSU Dept. of Economics, the Brown Growth Lab Seminar, the Tulane University Dept. of Economics, the Western Economic Association Conference, and the Southern Economic Association Conference.
Email: Justin Cook ccook7@tulane.edu

## 1 Introduction

People and economic output are not randomly distributed across the world; instead, they are tied to historic settlements, or spatially persistent. Indeed, cross-country correlation coefficients show a strong persistence in population densities across time, and this positive relationship between historic settlement patterns and contemporary populations and economic activity has been documented in a growing literature that explores the proposed stickiness of populations (and resulting output) both across and within countries.[1] This paper seeks to contribute to this literature by introducing a novel data set that synthesizes the cross- and within-country approaches. The data also allow for new tests of mechanisms that examine whether country-level factors like institutional quality are driving the estimated relationship between historic settlements and contemporary output.

In order to measure historical settlements, I use a new data set that geolocalizes the position of roughly 1,450 civilizations from 3200 BCE to the present day (c.2006). These data dynamically plot civilization borders, documenting expansions and contractions over time. The data are geolocated and not bound by contemporary country borders, which are often used as units of analysis for periods preceding the formation of these boundaries. The time and location of all civilizations are then collapsed to a spatially distributed measure that catalogs the accumulated historical presence of a civilization for a sub-national cell, creating a measure of state history similar in concept and construction to those of Bockstette et al. (2002) and Borcan et al. (2019).

I follow the empirical framework of (Henderson et al., 2018, hereafter HSSW), who create a spatial quarter degree squared grid of inhabitable land and show that a parsimonious set of environmental controls tied to agriculture and trade explain roughly half of the variation in nighttime lights. HSSW also argue that agricultural variables serve as a substitute for historical societal organization, but they are unable to directly show this because "[they] can't observe the detailed locations of historical agglomerations[.]" This paper constructs a proxy for these historical agglomerations and tests for an association with contemporary nighttime lights and population, conditional on HSSW's geo-climatic covariates. In other words, I take the $\approx$ 240K grid cells of HSSW, count the years a cell has been part of a civilization (i.e., state history), then test whether cells that have a prolonged exposure to civilization (or more state history) also have more contemporary economic activity–proxied by nighttime lights–and larger contemporary populations.[2]

[1] The correlation coefficient is 0.88 between 1 CE and 1500 CE population densities and 0.62 between 1500 CE and 2000 CE densities.

[2] My granular spatial state history data are also not beholden to modern day country borders, which were endogenously formed by natural or colonized means that are possibly tied to current economic conditions (Borcan, Olsson, and Putterman, 2018). Civilization borders were endogenously formed too, but these historic borders do not perfectly align with current borders and can be seen as conditionally randomized

The spatial data allow for across and *within* country estimations, where within country variation allows for the explicit control of federal institutions and other unobserved country-level differences. Across all specifications I do indeed find a longer state history is associated with richer and more populated cells today, a finding inline with prior studies (discussed in detail in Section 1.1). But I also find that controls for broad country-level differences in institutional quality make little difference in the estimated relationship, suggesting the estimated effect of state history on output/population is independent of accumulated differences in institutional quality. This finding runs counter to prior studies that argue prolonged exposure to higher-order state formation leads to learning-by-doing in state capacity that manifests in country-level differences in federal institutions and ultimately, a range of socioeconomic measures today (Borcan, Olsson, and Putterman, 2018; Lagerlöf, 2016).

The effect of state history being independent of country-level factors suggests a mechanism other than contemporary institutional quality is at play. One possibility is that the persistence of populations and economic activity are due to agglomeration effects, or hard-to-measure externalities, that contribute to the persistence of cities and states, such that populations build upon historic settlements rather than move to geoclimatically identical, but new locations. Agglomeration effects are more likely in urban locations (Glaeser, 2010). Examples of this are numerous, from Mexico City in the New World to Rome, Athens, Damascus, and Luoyang in the Old World. Historic markets, or centers of trade, serve as focal points that persist over time, creating a strong link between the historical presence of the city and contemporary population/output. With this in mind, I test whether the effect of state history is stronger in urban environments, decaying as cells mover farther away from the city-center.

Baseline estimates show areas with a more prolonged state history are richer, proxied by being more well lit, and more populated today. A standard deviation increase in state history leads to a 100% increase in HSSW's measure of nighttime lights (i.e., modified DMSP), a 110% increase in VIIRS nighttime lights, and a 140% increase in population.[3] These estimated relationships are unaffected when including or excluding country FE or direct measures for institutional quality, suggesting a mechanism that is independent of

within a country. That is, when controlling for agricultural conditions and country fixed effects, the historic placement of a civilization within a country is more randomly determined than modern country borders. Indeed, the approach I take in the current paper is explicitly laid out in BOP (p. 3): "A potential alternative to using country borders could have been to divide the world into equal-sized grid cells (or "virtual countries") and then study the history of each such cell." Depetris-Chauvin (2015) and Maloney and Valencia Caicedo (2016) also perform similar analyses; however, each is constrained to particular time periods and locations–respectively, post-1000 CE sub-Saharan Africa and post-1492 America. My approach instead considers the global distribution of all civilizations from 3200 BCE to the present.

[3]These effects are large but similar to the effects of control variables.

institutional quality. To this point, the effect of state history is influenced by the proximity of a city, becoming insignificantly different than zero roughly 600 kilometers outside of an urban location. Taken together, these heterogenetiy tests suggests some role for agglomeration that is independent of institutional quality.

The rest of the paper is organized as follows: Section 1.1 provides a background of prior studies linking state history to contemporary outcomes; Section 2.1 gives a descriptions of the primary state history variable and all other variables used in the paper; Section 2.2 compares the constructed state history measure to other similar measures; Section 2.3 gives the baseline empirical specification, the results of which are given in Sec. 3.1; heterogeneity tests of mechanisms are discussed in Sec. 3.2; and robustness exercises for the relationship between state history and contemporary outcomes are discussed in Sec. 3.3.

## 1.1 Related Literature

My measure of state history is based on a similar measure from (Bockstette, Chanda, and Putterman, 2002, hereafter BCP) and (Borcan, Olsson, and Putterman, 2018, hereafter BOP). BCP provides evidence that country-level growth in GDP per capita following decolonization in the 1960s is positively associated with state capacity, which is determined by state history. The mechanism being that longer lived states are more stable and have higher institutional quality and that this stability contributed to faster growth rates.

BOP extends this measure of state history and argues for non-linear associations between state history and state capacity. The hypothesis being that state history increases state capacity up to a point with prolonged exposure leading to over-centralization, creating a hump-shaped (inverted-U) relationship between state history and current development (Hariri, 2012; Lagerlöf, 2016). The non-linear mechanism of BOP are again primarily tied to institutional quality from longer lived states.

My approach builds off related sub-national research within Africa (Depetris-Chauvin, 2015; Michalopoulos and Papaioannou, 2013). Michalopoulos and Papaioannou use ancestral ethnic boundaries and a measure of state formation for each ethnicity to estimate the effect of higher order governance on contemporary economic output, proxied by night-time lights; indeed, they find a strong, positive, and robust relationship between ancestral state formation and contemporary development. My approach is most similar to work by Depetris-Chauvin (2015), who documents the boundaries of civilizations in sub-Saharan and connects these to modern conflict. Figure A.2 directly compares Depetris-Chauvin's measure of state history to my own for Africa, showing considerable overlaps between the independently constructed measures. My approach extends the core idea of these papers by plotting borders for civilizations across the globe, not just Africa, and extends the time frame to cover the historic universe of human settlements.

Outside of Africa, work by (Maloney and Valencia Caicedo, 2016, hereafter MVC) uses historic censuses to estimate the persistence of populations. Within the Americas, states/provinces (from contemporary borders) that contained more people per square kilometer before the arrival of Columbus have a strong positive association with contemporary population densities and incomes. While MVC's analysis is limited to North and South America, my approach broadens the proposed relationship, both globally and deeper into the past; however, the degree of intensity of a settlement, which is captured by MVC's historic population density, is lost. That said, MVC's core research question is mirrored in the current work: within a country, do people currently live and produce where they have historically lived and produced?

# 2 Data and Empirical Strategy

## 2.1 Data

**Geolocalized Civilizations** One of the primary contributions of the current work is the introduction of a new data set that geolocalizes all (or nearly all) recorded civilizations with a capitol and recording system.[4] These locations and time frames come from a number of historic almanacs and encyclopedias; sources for which can be found in the appendix. The list of included civilizations is as comprehensive as possible as is the mapping of borders over time. To my knowledge, no other source attempts to locate all civilizations throughout history. The closest analogues are the *Centennia Historical Atlas*, which plots historic state boundaries for Europe, North African, and the Middle East starting in 1000 CE (Reed, 2016; Schönholzer and Weese, 2018); Murdock's mapping of ethnic homelands (Michalopoulos and Papaioannou, 2013; Murdock, 1959); and the set of maps from Depetris-Chauvin (2015), who documents 72 civilizations across 15 time periods between 1000-1850 CE. In contrast, my data plot locations for roughly 1,470 civilizations over 500 periods from 3200 BCE to 2006 CE for all habitable continents.

These data are not perfect. Borders are likely mislabeled or effectively meaningless in many instances. That said, I believe this error to be unbiased (classical), leading to a simple attenuation in the coefficient of interest. It is highly unlikely that grid cells that were heavily populated historically as part of a civilization are mislabled; instead, the converse is almost certainly true, especially given climatic change in the regions of the most historic civilizations–North Africa and the Middle East. Indeed, Classical Egypt predates

[4]The locations, sizes, placements, and sources of these borders were originally done by Andrew Tollefson. The maps were then used for an informational YouTube video: https://www.youtube.com/watch?v=dp0tqdu7fH4. These maps were then geolocalized by the SpARC lab at UC Merced, primarily by Erin Mutch and Amy Newsam.

the growth of the Saharan Desert into North Africa (Kröpelin et al., 2008). Furthermore, the early civilizations of the Middle East may have amplified desertification of this region, creating lightly populated/low output grids with a long state history (Ruddiman, 2003). In other words, bias from incorrectly measured civilization borders have likely *attenuated* the estimated relationship between state history and contemporary populations/lights and not led to a false-positive relationship.[5]

**Sub-National State History** The sub-national measure follows a similar form as measures in BCP, BOP, and Depetris-Chauvin. Using HSSW's grid, I count the years a civilization has been within a particular cell. A simple intersection of a civilization with a cell is counted as the cell containing a civilization in that year. In other words, the civilization does not have to comprise a majority of area in a cell; if it is present, it is counted, regardless of coverage. This counting also excludes years when a particular grid is no longer part of a civilization, allowing for a more accurate counting of state history for marginal areas along the border or absences associated with a collapse of a civilization. Following BOP, I discount by 1% for 50 year intervals, sum the discounted years a cell has been part of a civilization, and set the variable relative to the maximum value.

Appendix Figure A.1 plots my measure of state history across the world. As expected, Eurasia has a longer history of a state presence; Africa shows a similar pattern to the previously collected data of Depetris-Chauvin with more state history in North Africa and the areas immediately beneath the Saharan desert; the New World shows patterns consistent with the big 3 civilizations–Mayan, Incan, and Aztec Empires; the United States shows a longer state history that is tied to its colonial origins; and Australia shows relatively recent state presence.

**Outcome and Control Variables** My hypothesis is that the historic location of civilizations have been sticky for contemporary development. To measure this association, I look at the relationship between sub-national state history and contemporary populations and output. To measure output, I follow HSSW and use their measure for the natural log of calibrated nighttime lights, but I also supplement HSSW's DMSP nighttime light data with the generally improved VIIRS nighttime light data (Gibson, 2021). Population is from SEDAC's UN adjusted population count data (CIESIN, 2018). HSSW's DMSP nighttime light and SEDAC's population are measured for 2010, while VIIRS nighttime lights are for 2014 (the first complete year of existence that corrects for stray lights; Gibson and

[5]Since my data are primarily based on historic atlases, civilizations are likely Euro-centric, affecting the included civilizations in both accuracy and the definition of a "civilization". While not ideal, this problem is present in other databases like Murdock's Ethnographic atlas.

Boe-Gibson (2021)).[6]

Control variables are also directly taken from HSSW, who use a set of 24 geoclimatic controls to account for the historic and persistent effect from agriculture and access to trade. The agricultural variables includes indicators for biome, mean temperature, mean precipitation, the number of grow days, land suitability, latitude, and elevation. Controls for trade include indicators for being on a coast, the mean distance to the coast, and indicators for being within 25km of a harbor, river, or lake. Summary statistics for all variables can be found in Table A.1.

I also use HSSW's unit of analysis: quarter degree latitude by longitude cells for all habitable locations, totalling 242K observations. These grid cells, which are roughly 770 square kilometers at the equator (or 28km $\times$ 28km), are sufficiently large to overcome concerns of spatial spillovers while also providing a large number of observations per country, allowing for precise within country estimations.[7]

### 2.2 Validating Cell-Level State History

Figure 1 takes the average state history score within a country to compare my measure to the country-level measure of BCP and BOP. A clear and strong positive association exists between the two measures. The pairwise correlation is 0.77 and the Spearman rank correlation is 0.79. There are a few differences, however, between my measure of state history and the measure of BCP/BOP. First, BCP's/BOP's measure of state history codes for units of hierarchy beneath that of a kingdom or civilization. My measure only considers civilizations that have a capitol and a writing system and does not account for chiefdoms, or other lower-order forms of state organization, that are present in BCP's/BOP's data. Second, within country weighting for the geographic distribution of a historic civilization differs from BCP/BOP. Third, BCP/BOP discount invading or non-local civilizations, while I make no such distinction (the presence of a civilization is always coded as a "1" regardless of the civilization's origins). To better highlight the differences between my measure and BCP/BOP's measure, I examine two outliers: Ethiopia and Kyrgyzstan.

**Case Analysis: Ethiopia and Kyrgyzstan** Figure 2a plots the borders of the Kingdom of Axum (or Aksum) in 1 CE for modern day Ethiopia. As seen, the Kingdom of Axum covers a very small portion of Ethiopia, yet this kingdom is used to justify the long state history date of Ethiopia in BCP/BOP. Using the modern day borders of Ethiopia, most of

[6] VIIRS nighttime light data can be found at https://eogdata.mines.edu/products/vnl/. I use Annual VNL V2.1 data for the average masked score. The granular data are then summed within quarter degree grids as specified in Henderson et al. (2018).

[7] Spatial spillovers remain a concern; however, standard error adjustments provided in Table 1 suggest spillovers are not affecting the statistical significance of estimated effects of state history.

the country has had very little exposure to state organization, with only 3% being exposed to the Kingdom of Axum. This difference leads to Ethiopia being an outlier in Figure 2 and highlights the importance of a sub-national approach.

Figure b plots the modern day border of Kyrgyzstan as well as civilization borders for 1 CE. As seen a number of civilizations–the Wusun (NE), Han (SE), Dayuan (W), and Yuezhi (S) civilizations–overlap with Kyrgyzstan's modern day borders at this early date, but BCP'/BOP's measure ascribes a recent state history Kyrgyzstan. This is likely due to discounting from occupying instead of home-grown civilizations. For my purposes, this distinction is immaterial. I simply seek to measure the persistence of people and resulting cultures from exposure to higher order state organization.

### 2.3 Empirical Strategy

The empirical strategy follows that of BCP/BOP while using the unit of observation and geoclimatic controls of HSSW. This conditional correlation approach follows the literature and has similar threats to causation. The primary estimating equation is:

$$y_i = \alpha_i + \beta_{sh} \times StateHistory_i + \beta_2' \mathbf{Base_i} + \beta_3' \mathbf{Agr_i} + \beta_4' \mathbf{Trade_i} + \gamma_{ci} + \epsilon_i \qquad (1)$$

Again, $i$ represents land area divided into $0.25° \times 0.25°$ latitude by longitude grid. I consider three outcome variables, $y_i$: (1) the natural log of DMSP nighttime lights (from HSSW), (2) the natural log of population count (SEDAC; UN-WPP adjusted), and (3) VIIRS nighttime lights (c.2014).[8] State history is my primary explanatory variable and measures the number of discounted years (relative to the maximum) a grid has been part of a civilization. The primary hypothesis is that populations and output have persisted in places that have historically been part of a civilization, implying $\beta_{sh} > 0$. The set of agricultural and access to trade variables from HSSW are respectively given by $Agr.$ and $Trade.$ Country fixed effects are included in the baseline estimation and given above by $\gamma_c$.

Assuming agriculture is a necessary condition for state organization, the inclusion of HSSW's agricultural set of controls will be collinear with state history. As noted in Table B.1, more favorable agricultural conditions are tied to state history, making it almost certain collinearity is affecting the estimated coefficient of state history. Nevertheless, favorable agricultural conditions likely also have an explanatory role for contemporary development outside of the relationship with historic development, and agricultural conditions alone do not fully explain the variation in state history. With this in mind, I include all of HSSW's geoclimatic variables as controls but piecemeal introduce these controls to show

[8] The first full year of VIIRS availability is 2014. I use this year to stay as close as possible to other outcomes measured in 2010.

changes in the coefficient of state history. Following HSSW, I also report Shapley values in Appendix Table B.2, which decompose the marginal contribution to the R-square for each set of controls. These decompositions suggest state history independently explains a greater portion of the variation in nighttime lights and population than HSSW's trade variables, but the set of agricultural variables remains the largest predictor of contemporary development. LASSO model selection includes all geoclimatic controls and state history, again suggesting an independent contribution from state history.

# 3 Results

## 3.1 Baseline Results

Base results are presented in Table 1. Panel A considers HSSW's grid measure of nighttime lights; Panel B regresses the natural log of population from SEDAC (UN-WPP adjusted); and Panel C regresses the natural log of VIIRS nighttime lights. Controls are piecemeal introduced identically for all panels. Column (1) estimates the bivariate relationship between either the natural log of nighttime lights or population. Column (2) adds country fixed effects to the estimation of column (1). Column (3) introduces HSSW's base variables–ruggedness and malaria ecology–to the estimation of column (2). Columns (4) and (5) separately include the set of agriculture and access to trade controls to the estimation of column (3). And column (6) comprises the base specification, which includes all of HSSW's geoclimatic controls and country fixed effects. Three standard errors are included: (1) country clustered, (2) three-by-three grid clustered (from HSSW), and (3) spatially adjusted for 200 km. I use the highest nested (i.e., country) and most conservative cluster for determining statistical significance. The coefficient of state history is statistically significant at the 1% level for all specifications and all outcomes.

For the bivariate estimates of column (1) with no country fixed-effects, a standard deviation increase in state history (i.e., 0.2) is associated with roughly a tripling of DMSP nighttime lights, a five-fold increase in population, and a tripling of VIIRS nighttime lights. These large effects are in part due to the large number of cells with no nighttime lights; an issue that is discussed in Sections 3.3. The estimated effect of state history is effectively unaltered when including country fixed effects in column (2), a point also discussed in greater detail below. The inclusion of HSSW's base controls–i.e., ruggedness and malaria ecology–in column (3) again do not significantly alter the coefficient of state history.

The inclusion of agricultural controls in column (4), however, causes the coefficient of state history to be reduced by roughly half for all outcomes, implying a one standard deviation increase in state history is now associated with a 100% of DMSP nighttime lights, a 140% increase in population, and a 110% increase in VIIRS nighttime lights. As discussed in Section 2.3, HSSW's agricultural controls intend to capture agricultural conditions that have led to prolonged state history, suggesting the inclusion of the agricultural set of controls splits a common effect among many coefficients. That said, the lowered magnitude coefficient of state history from the inclusion of agricultural controls remains statistically significant and the lowered magnitude still suggest a sizable economic impact from a grid having more state history. Furthermore, Appendix B provides additional evidence that state history is indeed making an independent and sizable contribution in explaining the three outcomes of interest.

Column (5) adds HSSW's trade variables to the regression specification of column (3)–i.e., country FE and the base set of controls–with no substantial changes to the coefficient of state history, and column (6) comprises the baseline specification and includes all controls. The estimated coefficient of column (6) shows the reduced magnitude associated with the introduction of agricultural variables in column (4) with point estimates and economic effects being similar in magnitude to those in column (4).

The relatively large effects of state history are inline with the estimated effects of other variables. For example, from Tables A.1 and column (6) of Tables E.1 and E.2, a one standard deviation increase in the average monthly temperature is associated with roughly a quadrupling of DMSP nighttime lights and a quintupling of population. Similarly large effects are estimated for other covariates.

### 3.2 Institutions or Agglomeration

**No Alteration from Country FE** Table 1 provides the coefficients of state history for identical estimations absent country fixed effects. The magnitudes are surprisingly similar, and a Hausman-based test shows no statistically significant differences in the coefficients of state history from a model with country fixed effects and a model omitting country fixed effects for both nighttime light outcomes; although, statistically significant differences do exist (larger when omitting country FE) when considering population and including agricultural controls (columns 4 and 6). The consistency in magnitude across specifications suggests time-invariant country-level factors like institutional quality are not playing a large role in the relationship between a cell's state history and its corresponding lights/population. In other words, if state history were working solely through country-level institutional quality, one would expect country-level fixed effects to alter (likely attenuate) the relationship between state history and nighttime lights/population. The absence of any change from the inclusion of country fixed effects suggests any country-level differences are orthogonal to the estimated grid-cell relationship.

**Institutional Quality FE** While the the estimated relationship between cell-level state history and nighttime lights/population appears to be independent of country-level differences, country FE may possibly over control for other factors not directly related to institutional quality. With this idea in mind, Table 2 directly controls for institutional quality using data from the Polity5 Project (Marshall and Gurr, 2020). The Polity5 Project documents country-level institutional changes from 1800 CE to the present and categorizes institutions based on autocratic (-10; "strongly autocratic") versus democratic (+10; "strongly democratic") governance in the "POLITY" variable. In addition to this 21-point scale, countries are take differing values based on transitions from one type of governance to another. Table 2 uses the POLITY data for 2010 as a direct control for institutional quality and performs heterogeneity tests for a varied effect of state history by institutional quality.

The autocracy/democracy variation is closely tied to the arguments of BOP. The idea being that older lived autocratic governments have persisted, serving as a (negative) mechanism connecting state history to current economic development (Hariri, 2012; Olsson and Paik, 2020). Alternatively, newer states developed inclusive, democratic institutions that led to economic prosperity; BOP's positive mechanism of state history. With these mechanisms in mind, the POLITY variable is a good candidate to measure the institutional quality mechanisms proposed by BOP.

Column (1) of Table 2 directly controls for the POLITY score by introducing each category as a fixed effect.[9] The inclusion of institutional quality fixed effects does not significantly alter the coefficient of state history for any of the three outcomes; the estimated effects are similar in magnitude to those of column (6) of Table 1, with or without country FE.

Columns (2)-(4) use a dichotomized POLITY score that serves as a democracy indicator– i.e., POLITY $> 0$. Countries with transition POLITY scores are omitted. Column (2) directly controls for the country-level indicator of democracy, leading to no large differences in the coefficients of state history. Column (3) interacts the democracy indicator with state history to test whether the effect of state history differs within democratic versus autocratic governments. For all panels, the coefficient of the interaction is insignificant, suggesting there is no statistically significant difference in the effect of state history across democratic or autocratic countries. Column (4) builds on the estimates of column (3) by including country fixed effects.[10] As with column (3), however, no statistically significant difference is seen in the effect of state history for democratic countries.

[9] Linearly controlling for POLITY does not alter the estimated coefficient of state history.

[10] Note that the democracy indicator is no longer included due to being perfectly collinear with the country FE, but the interaction is estimated.

**Relationship Driven by Cities** Tables 1 and 2 provide evidence that the relationship between state history and nighttime lights/population is unaffected by country level controls, suggesting a federal institution mechanism is absent. Instead of "proving a negative", Table 3 provides evidence for agglomeration by showing the effect of state history is driven by proximity to urban areas.[11]

Table 3 uses the mid-point of a cell to determine the distance (in 100s of kilometers) to the closest urban location. This distance is then interacted with the cell's state history score to determine how the marginal effect of state history on each of the three outcomes changes with distance from a city. Column (1) includes all baseline controls; column (2) adds country FE to the specification of column (1); and column (3) adds closest city FE to the specification of column (2). In general, the estimates of Table 3 show that the effect of state history is significantly tied to urban distance, becoming weaker in more rural areas and equal to zero roughly 500-600km away from a city. The decaying effect of state history holds for all specifications, and the inflection point, or where the marginal effect of state history is zero, is relatively consistent.

Agglomeration is much more pronounced in cities (Chanda and Ruan, 2017; Glaeser, 2010), and proximity to a city is driving the positive and significant effect of state history. This is counter to the estimates of Table 2, which show no variation in the effect of state history by federal-level institutions. The estimates from the heterogeneity tests of Table 2 and 3 and the lack of impact from country FE in Table 1 suggest a stronger role for an agglomeration mechanism than an institutional quality effect.

### 3.3 Robustness and Alternative Specifications

**Unlit/Unpopulated Cells** Figures C.1 - C.4 respectively plot the distributions for HSSW's DMSP nighttime lights, SEDAC's population count, VIIRS nighttime lights, and state history. Roughly 60% of cells are unlit (both DMSP and VIIRS) and 10% of cells are unpopulated. The sizable frequency of "empty" cells is likely to affect estimation, so Table C.1 considers alternative samples that exclude unlit/unpopulated cells, which comprises those in the bottom decile, and those in the top decile.[12]

Column (1) of Table C.1 replicates the estimation of column (6) of Table 1 while omitting the bottom decile of each outcome measure, which comprises grid cells with no lights (Panels A and C) or no population (Panel B). For nighttime lights in Panel A and C, this leads to a large reduction in the sample size (again, 60% of grids have no lights) and a reduction in

[11] Urban location data are from Reba, Reitsma, and Seto (2016), which aggregates data from Chandler (1987) and Modelski (1999).

[12] Bias from the large number of unlit/unpopulated cells is likely to be positive. Considering a positive linear fit, the large number of zero, or low value, observations will increase weight in the lower, lefthand side, potentially creating upward bias in the estimated relationship of interest.

the coefficient of state history that remains positive and statistically significant at the 1% level. For population in Panel B, the omission of unpopulated grid cells has a minor (to nonexistent) effect on the magnitude and significance on the coefficient of interest.

Column (2) of Table C.1 omits instead the top decile of nighttime lights and population. The omission of the top decile effectively removes urban areas and shows that my subnational measure of state history is not only accounting for the persistence of cities. The argument being that the persistence of both urban areas and areas in the hinterlands that feed these cities and more rural areas that are taxed and have access to public goods like roads to cities are all accounted for within a civilization's boundary. The hypothesis is not tied strictly to urban areas; although, effects are more pronounced in these areas due to agglomeration. As expected, estimates from column (3) are smaller in magnitude compared to the baseline estimates of column (6) in Table 1 but remain positive and statistically significant, suggesting contemporary urban areas are not the sole cause of the positive relationship between state history and contemporary output/populations.[13].

Column (3) omits both the top and bottom quartile of contemporary nighttime lights and population. For both nighttime light measures (Panels A and C), the coefficient of state history is reduced–primarily from unlit cells–but remains significant at the 1% level; Panel B shows a relatively smaller attenuation when considering population, which is driven by the omission of more populated cells.

To summarize, prevalent unlit/unpopulated cells do not dissipate the positive and statistically significant association between output/population and state history, but the sheer volume of these cells does significantly affect the magnitude of the relationship for nighttime lights. From columns (2) and (3) the effect of state history is not driven solely by urban areas, suggesting a rural/hinterland persistence that has not been documented in other work.

**Large Country Effects** While country FE account for country-specific time-invariant factors, the quarter degree latitude/longitude grids are mechanically tied to the area of a country. This implies larger countries will have more observations and a greater weighting on the estimated coefficient, with or without country FE. To show that this unequal weighting is not driving the relationship between state history and nighttime lights/population, Table D.1 limits the sample to within country observations for the largest countries–Russia, Canada, China, and the United States–and omits these large countries from the base sample.

Columns (1)-(4) respectively limit the sample to cells within Russia, Canada, China, and the United States. Point estimates vary from the baseline estimates of Table 1, but

[13] Similar effects are observed when omitting city locations from Table 3

this is expected given the focus on singular countries; however, the estimated coefficients from omitting these large countries in column (5) are similar in magnitude to the baseline estimates, suggesting little influence from a singular, but heavily weighted, country.

**Effect by Continent** Table D.2 examines the baseline estimation of Table 1, column (6) by continent. For nighttime lights, a positive and statistically significant effect of state history is estimated for Africa, Asia, South America, and Oceania. The largest estimated coefficient is for South America–inline with Maloney and Valencia Caicedo (2014), but this large coefficient is in part due to the lower standard deviation for state history within South America (see Table A.1). That is, a standard deviation increase in state history for South America is associated with roughly a doubling in nighttime lights, a finding very similar to the baseline estimate of column (6) in Table 1. Estimated effects for population are very similar in Panel B. To account for differences in the variation of state history across continents, Table **??** standardizes state history within continent (or Old/New World grouping) to a mean of zero and standard deviation of one. Once this adjustment is made, the effect of state history is mostly inline across continents, still remaining relatively large in South America and the Old World effect of state history is more pronounced when compared to the New World effect.

## 4 Conclusion

This paper introduces a novel data set that documents the location and time of a civilization's borders. These borders are not tied to current country borders and allow for the creation of a measure of state history by counting the years a plot of land has been exposed to any civilization. Plots are from a $0.25°$ by $0.25°$ fishnet created by Henderson et al. (2018) that provides a large number of observations per country. I then test the relationship between nighttime lights and population both across and within countries, finding a consistent (between and within country), statistically significant, and positive role for state history in measuring differences in economic output today.

The consistency in the effect of state history with and without country fixed effects suggests country-level factors, like institutional quality, are not influencing the estimated relationship. This contrasts with the mechanism posited by BCP and BOP that suggests improvements in state capacity over time lead to better economic outcomes today. Furthermore, directly controlling for institutional quality or testing for heterogeneity from institutional quality shows no role for federal level institutions. Instead, it is much more likely that the mechanism connecting state history to contemporary development is agglomeration. Indeed, I observe statistically significant heterogeneity in the effect of state history tied to the distance from an urban center, where agglomeration effects are much more pronounced. The lack of heterogeneity from institutional quality measures coupled with the strong heterogeneity from urban distances suggests agglomeration as the much more likely mechanism. That said, I cannot explicitly rule out federal institutional quality as a primary mechanism.

## References


Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and markets: The advantage of an early start." *Journal of Economic Growth* 7 (4):347–369.

Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "State history and economic development: evidence from six millennia." *Journal of Economic Growth* 23 (1):1–40.

Chanda, Areendam and Dachao Ruan. 2017. "Early urbanization and the persistence of regional disparities within countries." *Available at SSRN 2909199* .

Chandler, Tertius. 1987. *Four thousand years of urban growth: An historical census*. Mellen.

CIESIN. 2018. "Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11." https://doi.org/10.7927/H4JW8BX5.

Depetris-Chauvin, Emilio. 2015. "State history and contemporary conflict: Evidence from sub-saharan africa." *Available at SSRN 2679594* .

Gibson, John. 2021. "Better night lights data, for longer." *Oxford Bulletin of Economics and Statistics* 83 (3):770–791.

Gibson, John and Geua Boe-Gibson. 2021. "Nighttime lights and county-level economic activity in the United States: 2001 to 2019." *Remote Sensing* 13 (14):2741.

Glaeser, Edward L. 2010. "Introduction to" Agglomeration Economics"." In *Agglomeration economics*. University of Chicago Press, 1–14.

Hariri, Jacob Gerner. 2012. "The autocratic legacy of early statehood." *American Political Science Review* 106 (3):471–494.

Henderson, J Vernon, Tim Squires, Adam Storeygard, and David Weil. 2018. "The global distribution of economic activity: nature, history, and the role of trade." *The Quarterly Journal of Economics* 133 (1):357–406.

Kröpelin, Stefan, Dirk Verschuren, A-M Lézine, Hilde Eggermont, Christine Cocquyt, Pierre Francus, J-P Cazet, Maureen Fagot, Bob Rumes, James M Russell et al. 2008. "Climate-driven ecosystem succession in the Sahara: the past 6000 years." *Science* 320 (5877):765–768.

Lagerlöf, Nils-Petter. 2016. "Statehood, democracy and preindustrial development." *Journal of Economic Dynamics and Control* 67:58–72.

Maloney, William F and Felipe Valencia Caicedo. 2016. "The persistence of (subnational) fortune." *The Economic Journal* 126 (598):2363–2401.

Marshall, Monty G and Ted Robert Gurr. 2020. "POLITY5: Dataset users' manual." *Center for Systemic Peace* .

Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-colonial ethnic institutions and contemporary African development." *Econometrica* 81 (1):113–152.

Modelski, George. 1999. "Ancient world cities 4000–1000 BC: Centre/hinterland in the world system." *Global Society* 13 (4):383–392.

Murdock, George Peter. 1959. *Africa its peoples and their culture history.* McGraw-Hill Book Company.

Olsson, Ola and Christopher Paik. 2020. "A Western reversal since the Neolithic? The long-run impact of early agriculture." *The Journal of Economic History* 80 (1):100–135.

Reba, Meredith, Femke Reitsma, and Karen C Seto. 2016. "Spatializing 6,000 years of global urbanization from 3700 BC to AD 2000." *Scientific data* 3 (1):1–16.

Reed, Frank. 2016. *The Centennia Historical Atlas: Academic Research Edition.* Clockwork Mapping: Jamestown, RI.

Ruddiman, William F. 2003. "The anthropogenic greenhouse era began thousands of years ago." *Climatic change* 61 (3):261–293.

Schönholzer, David and Eric Weese. 2018. "Creative destruction in the european state system: 1000-1850." *Working Paper* .

# 5 Tables and Figures

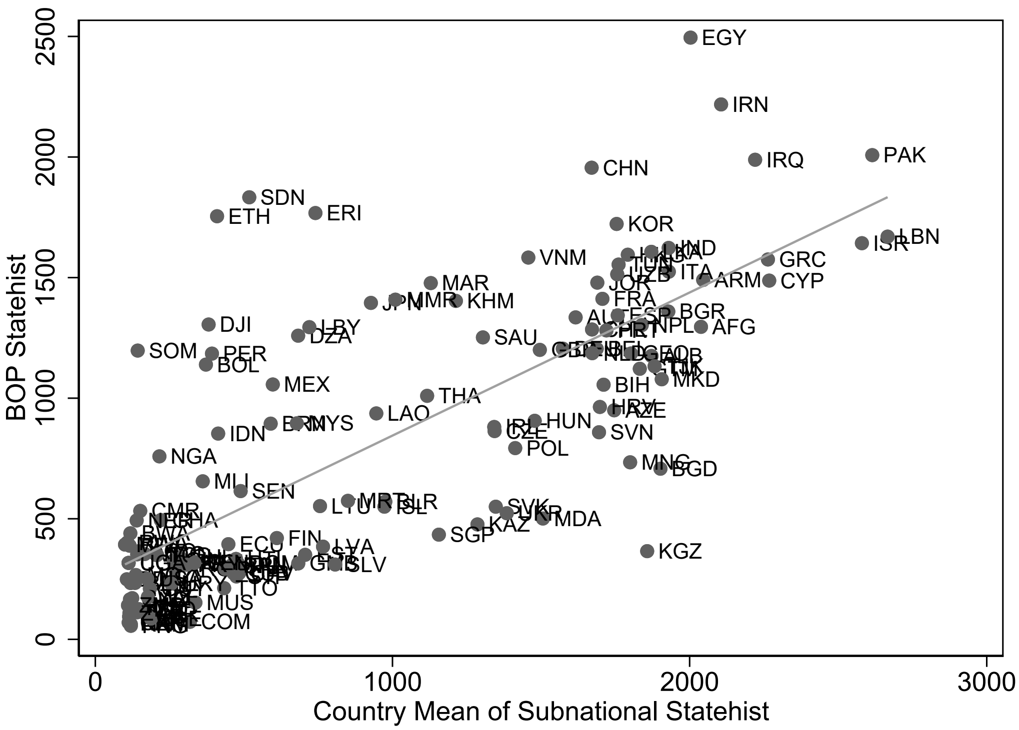
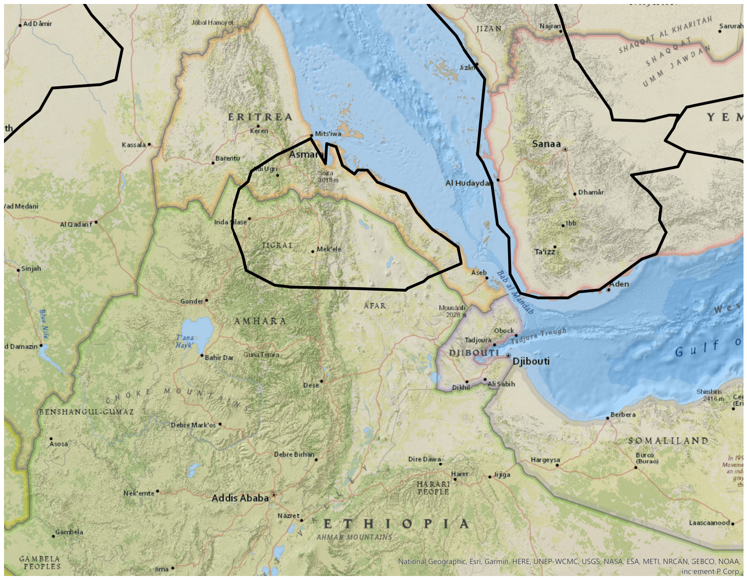
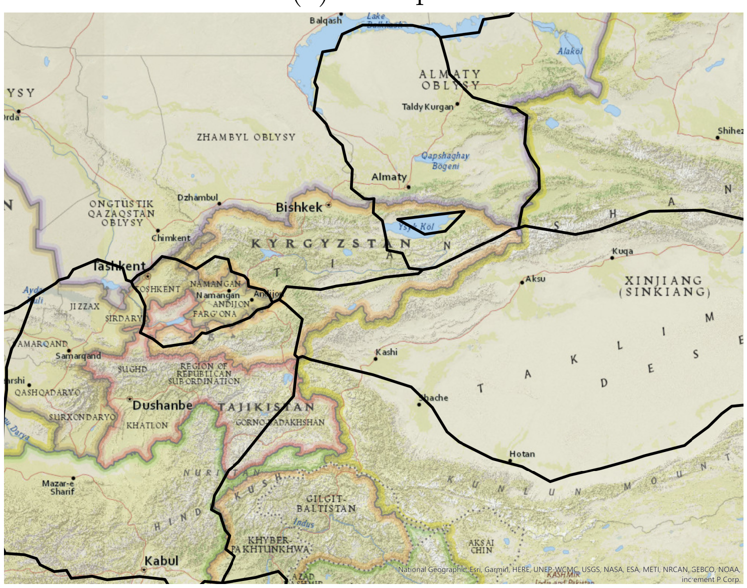



**Figure 1:** Comparing BOP's State History to Cell State History

**Notes:**
This figure plots state history from Borcan et al. (2018; y-axis) relative to a country aggregated (mean) measure of my cell-level state history measure based geolocalized civilizations (x-axis). The two measures are highly correlated with a pairwise correlation coefficient of 0.77 (p<0.000). Of note, outliers from Figure 1 illustrate the differences between my measure and that of BOP. This is discussed further in the case analysis of Section 2.1 and in Figure 2.




(a) Ethiopia




(b) Kyrgyzstan

**Figure 2:** Civilization Borders: 1 CE

**Notes:**
This figure plots geolocalized civilization borders for 1 CE relative to the modern day borders of Ethiopia (sub-figure (a)) and Kyrgyzstan (sub-figure (b)). These two countries serve as outliers in Figure 1, showing large differences between my measure of state history and that of BOP. Ethiopia has a long state history in BOP but relatively short state history in my analysis. This disparity is due to the geographic distribution of the Kingdom of Aksum, the borders of which for 1 CE are plotted on top of modern day Ethiopia in sub-figure (a). As seen, the Kingdom of Aksum only covers a small portion of modern day Ethiopia, with the vast majority of the country not belonging to a historic civilization. At the opposite end of the spectrum, BOP define Kyrgyzstan as having a relatively recent state history, while my measure shows a longer lived state presence. Sub-figure (b) plots civilizations from 1 CE around modern day Kyrgyzstan. While being in modern-day Kyrgyzstan, these civilizations are not local to Kyrgyzstan, leading to substantial discounting in BOP's measure of state history.

**Table 1:** Baseline Results: Nighttime lights, population, and state history

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Panel A. ln DMSP Nighttime Lights, 2010 | | | | | |
| State History | 5.7222*** | 8.0480*** | 8.0173*** | 3.4938*** | 7.9298*** | 3.5624*** |
| country-clustered s.e. | (0.9966) | (0.8897) | (0.8573) | (1.0009) | (0.9548) | (0.8074) |
| HSSW (3-grid) clustered s.e. | (0.0515) | (0.0989) | (0.0961) | (0.0730) | (0.0928) | (0.0705) |
| spatial–200km–clustered s.e. | (0.3559) | (0.6368) | (0.6121) | (0.3862) | (0.5881) | (0.3569) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.1408 | 0.3965 | 0.4068 | 0.5747 | 0.4196 | 0.5857 |
| $\beta^{SH}$, omitting country FE | | 5.7222*** | 5.7001*** | 2.9039*** | 6.2498*** | 3.1713*** |
| | | (0.9966) | (1.1000) | (0.8411) | (1.0816) | (0.8396) |
| p-value of difference | | 0.0898 | 0.0920 | 0.3776 | 0.2015 | 0.5204 |
| | Panel B. ln Population, 2010 | | | | | |
| State History | 8.4013*** | 9.4568*** | 9.4860*** | 4.2796*** | 9.4541*** | 4.4186*** |
| country-clustered s.e. | (1.8863) | (1.4901) | (1.4450) | (0.5559) | (1.5534) | (0.4679) |
| HSSW (3-grid) clustered s.e. | (0.0917) | (0.1662) | (0.1610) | (0.1111) | (0.1604) | (0.1090) |
| spatial–200km–clustered s.e. | (0.4546) | (0.6929) | (0.6702) | (0.3812) | (0.6588) | (0.3670) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.2100 | 0.6147 | 0.6168 | 0.7749 | 0.6238 | 0.7809 |
| $\beta^{SH}$, omitting country FE | | 8.4013*** | 9.7261*** | 6.7255*** | 9.9031*** | 6.8219*** |
| | | (1.8863) | (1.7672) | (0.9417) | (1.6190) | (0.9180) |
| p-value of difference | | 0.5940 | 0.8986 | 0.0210 | 0.8086 | 0.0135 |
| | Panel C. ln VIIRS Nighttime Lights, 2014 | | | | | |
| State History | 5.7467*** | 7.9630*** | 7.9365*** | 3.7159*** | 7.9128*** | 3.8184*** |
| country-clustered s.e. | (1.0050) | (0.9092) | (0.8898) | (1.0285) | (0.9718) | (0.8424) |
| HSSW (3-grid) clustered s.e. | (0.0773) | (0.1605) | (0.1584) | (0.1229) | (0.1555) | (0.1186) |
| spatial–200km–clustered s.e. | (0.3203) | (0.6109) | (0.5966) | (0.3929) | (0.5737) | (0.3608) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.1471 | 0.3843 | 0.3895 | 0.5426 | 0.3977 | 0.5499 |
| $\beta^{SH}$, omitting country FE | | 5.7467*** | 5.6908*** | 2.9386*** | 6.1925*** | 3.2094*** |
| | | (1.0050) | (1.1090) | (0.8713) | (1.1117) | (0.8692) |
| p-value of difference | | 0.1072 | 0.1080 | 0.4583 | 0.2029 | 0.5138 |
| Controls (for all panels): | | | | | | |
| Country FE | | Y | Y | Y | Y | Y |
| Base variables | | | Y | Y | Y | Y |
| Agricultural variables | | | | Y | | Y |
| Trade variables | | | | | Y | Y |

**Summary & Notes:** This table presents the baseline estimates showing that sub-national state history has a persistent and positive relationship with contemporary output (proxied by nighttime lights) and population. Panel A regresses the natural log of HSSW's DMSP nighttime lights; Panel B regresses the natural log of cell-level population; and Panel C regresses the log of VIIRS nighttime lights. Sets of controls and summary statistics are given in Table A.1. Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *. Spatially adjusted (200km) and HSSW's 3-grid cluster standard errors are also reported.

**Table 2:** Institutional Quality: Controlling and Testing Heterogeneity

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Panel A. ln DMSP Nighttime Lights, 2010 | | | |
| State History | 3.533*** | 3.592*** | 3.623*** | 3.250*** |
| | (0.721) | (0.800) | (0.625) | (0.590) |
| Polity Indicator, 𝟙[Democracy] | | 0.577** | 0.590* | |
| | | (0.236) | (0.352) | |
| State History × Polity Ind. | | | -0.045 | 0.448 |
| | | | (0.892) | (1.423) |
| Observations | 238,952 | 236,120 | 236,120 | 236,119 |
| R Sqr. | 0.512 | 0.500 | 0.500 | 0.584 |
| | Panel B. ln Population, 2010 | | | |
| State History | 5.113*** | 6.606*** | 3.792*** | 3.997*** |
| | (0.602) | (0.839) | (0.772) | (0.643) |
| Polity Ind., 𝟙[Democracy] | | -0.336 | -1.528*** | |
| | | (0.384) | (0.575) | |
| State History × Polity Ind. | | | 4.127*** | 0.862 |
| | | | (1.302) | (0.953) |
| Observations | 238,952 | 236,120 | 236,120 | 236,119 |
| R Sqr. | 0.707 | 0.647 | 0.656 | 0.775 |
| | Panel C. ln VIIRS Nighttime Lights, 2014 | | | |
| State History | 3.624*** | 3.760*** | 4.031*** | 3.534*** |
| | (0.725) | (0.828) | (0.628) | (0.748) |
| Polity Ind., 𝟙[Democracy] | | 0.692*** | 0.807** | |
| | | (0.233) | (0.361) | |
| State History × Polity Ind. | | | -0.397 | 0.476 |
| | | | (0.928) | (1.595) |
| Observations | 238,952 | 236,120 | 236,120 | 236,119 |
| R Sqr. | 0.487 | 0.476 | 0.476 | 0.549 |
| Controls (for all panels): | | | | |
| Polity FE | Y | | | |
| Country FE | | | | Y |
| All geoclimatic controls | Y | Y | Y | Y |

**Summary & Notes:** This table tests whether direct controls for country-level institutional quality affect the estimated coefficient of state history across the two nighttime lights measures (Panels A and C) and population (Panel B). As shown, controlling for institutional quality (done by categorical fixed effects in column 1) does not alter the estimated effect. Going further, columns (2)-(4) perform heterogeneity tests using an indicator of democracy constructed from the Polity variable used in column (1). Estimates suggest the effect of state history is not significantly affected by the type of country-level governance. This suggests a channel outside of state capacity is driving the baseline estimates of Table 1. Geoclimatic controls are from HSSW and listed in Table 1. Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

## Table 3: Decay from Urban Distance

| | (1) | (2) | (3) |
|---|---|---|---|
| | Panel A. ln DMSP Nighttime Lights, 2010 | | |
| State History | 3.3166*** | 4.2156*** | 3.7663*** |
| | (0.9302) | (0.8080) | (0.7025) |
| Distance to nearest city (in 100km) | -0.0863*** | -0.0271 | -0.0243 |
| | (0.0256) | (0.0215) | (0.0157) |
| State History × Distance | -0.5934*** | -0.7732*** | -0.6755*** |
| | (0.1733) | (0.1874) | (0.1717) |
| Inflection point | 559km | 545km | 558km |
| Observations | 242,184 | 242,180 | 242,153 |
| R Sqr. | 0.5209 | 0.5969 | 0.6551 |
| | Panel B. ln Population, 2010 | | |
| State History | 5.0229*** | 4.5366*** | 4.3271*** |
| | (0.7167) | (0.7354) | (0.4674) |
| Distance to nearest city (in 100km) | -0.1817*** | -0.0759*** | -0.0560*** |
| | (0.0403) | (0.0218) | (0.0179) |
| State History × Distance | 0.4185 | -0.5692* | -0.6357** |
| | (0.4467) | (0.3432) | (0.2509) |
| Inflection point | n.a. | 797km | 681km |
| Observations | 242,184 | 242,180 | 242,153 |
| R Sqr. | 0.6786 | 0.7911 | 0.8375 |
| | Panel C. ln VIIRS Nighttime Lights, 2014 | | |
| State History | 3.5369*** | 4.6404*** | 4.0630*** |
| | (0.9733) | (0.8471) | (0.6619) |
| Distance to nearest city (in 100km) | -0.0708*** | -0.0125 | -0.0178 |
| | (0.0241) | (0.0195) | (0.0144) |
| State History × Distance | -0.6168*** | -0.7827*** | -0.7358*** |
| | (0.1803) | (0.1958) | (0.1783) |
| Inflection point | 573km | 593km | 552km |
| Observations | 242,184 | 242,180 | 242,153 |
| R Sqr. | 0.4911 | 0.5603 | 0.6240 |
| Controls: | | | |
| Country FE | | Y | Y |
| City FE | | | Y |
| All geoclimatic controls | Y | Y | Y |

**Summary & Notes:** This table provides evidence the persistent effects of state history are largely due to urban proximity. The coefficient of the interaction shows the marginal effect of state history decays as distance from a current or historic city increases; the effect of state history is indistinguishable from zero at roughly 600 km from an urban location. Panel A regresses the natural log of HSSW's measure of nighttime lights based on DMSP satellite readings for 2010; Panel B regresses the natural log of a cell's total population in 2010 (SEDAC); and Panel C regresses the natural log of VIIRS nighttime light for 2014, the first complete year of data with stray light corrections. Geoclimatic controls are from HSSW and listed in Table 1. Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.
.

**Online Appendix Tables and Figures**

# A Summary Statistics and Additional Figures

**Table A.1:** Summary Statistics

| Variable | Obs. | Mean | Std Deviation | Min | Max |
|---|---|---|---|---|---|
| State History | 242,184 | 0.1964 | 0.2045 | 0 | 1 |
| Years, Discounted (1% per 50) | 242,184 | 665.1138 | 692.5049 | 0 | 3386.872 |
| Years, Not Discounted | 242,184 | 784.6688 | 900.0418 | 0 | 5206 |
| By continent: | | | | | |
| Africa | 41,346 | 0.1178 | 0.1723 | 0 | 1 |
| Asia | 51,258 | 0.4633 | 0.1806 | 0 | 0.9737 |
| Europe | 60,050 | 0.2252 | 0.1509 | 0 | 0.7704 |
| South America | 25,369 | 0.0860 | 0.0342 | 0.0013 | 0.2905 |
| North America | 50,941 | 0.0533 | 0.0622 | 0 | 0.6966 |
| Oceania | 13,220 | 0.0400 | 0.0051 | 0.0179 | 0.0630 |
| Outcome Variables: | | | | | |
| HSSW's ln DMSP Nighttime Lights (c.2010) | 242,184 | -3.3571 | 3.1186 | -5.6841 | 6.9410 |
| SEDAC's ln Population Count (c.2010) | 242,184 | 6.2954 | 3.7458 | 0 | 16.1906 |
| ln VIIRS Nighttime Lights (c.2014) | 242,184 | 2.3087 | 3.0636 | 0 | 14.6335 |
| Control Variables: | | | | | |
| Base set: | | | | | |
| Rugged | 242,184 | 2.7806 | 4.8515 | 0 | 95.8144 |
| Malaria Ecology | 242,184 | 1.9208 | 5.2886 | 0 | 38.0810 |
| Agriculture set: | | | | | |
| Biome 1 (tropical moist forest) | 242,184 | 0.1168 | 0.3211 | 0 | 1 |
| Biome 2-3 (tropical dry forest) | 242,184 | 0.0223 | 0.1477 | 0 | 1 |
| Biome 4 (temperate broadleaf) | 242,184 | 0.1044 | 0.3059 | 0 | 1 |
| Biome 5 (temperate conifer) | 242,184 | 0.0330 | 0.1785 | 0 | 1 |
| Biome 6 (boreal forest) | 242,184 | 0.1664 | 0.3724 | 0 | 1 |
| Biome 7-9 (tropical grassland) | 242,184 | 0.1208 | 0.3259 | 0 | 1 |
| Biome 8 (temperate grassland) | 242,184 | 0.0772 | 0.2670 | 0 | 1 |
| Biome 10 (montane grassland) | 242,184 | 0.0334 | 0.1797 | 0 | 1 |
| Biome 11 (tundra) | 242,184 | 0.1221 | 0.3274 | 0 | 1 |
| Biome 12 (Mediterranean forest) | 242,184 | 0.0242 | 0.1536 | 0 | 1 |
| Biome 13 (desert) | 242,184 | 0.1753 | 0.3802 | 0 | 1 |
| Biome 14 (mangroves) | 242,184 | 0.004 | 0.0634 | 0 | 1 |
| Temperature | 242,184 | 10.0181 | 13.7679 | -22.2861 | 30.3660 |
| Precipitation | 242,184 | 60.8169 | 59.2736 | 0.3866 | 921.9088 |
| Growing Days | 242,184 | 139.63 | 99.0428 | 0 | 366 |
| Land Suitability | 242,184 | 0.2748 | 0.3200 | 0 | 1 |
| Absolute Latitude | 242,184 | 38.3148 | 20.9350 | 0.125 | 74.875 |
| Elevation | 242,184 | 0.6045 | 0.7899 | -0.1873 | 6.1690 |
| Trade set: | | | | | |
| Coastal Ind. | 242,184 | 0.0972 | 0.2963 | 0 | 1 |
| Dist. to Coast | 242,184 | 0.4865 | 0.4811 | 0 | 2.2738 |
| Harbor w/in 25km | 242,184 | 0.0273 | 0.1629 | 0 | 1 |
| River w/in 25km | 242,184 | 0.0273 | 0.1629 | 0 | 1 |
| Lake w/in 25km | 242,184 | 0.0108 | 0.1035 | 0 | 1 |

**Notes:**
This table provides summary statistics. The unit of observation for all variables is the 0.25° by 0.25° grid cell from Henderson et al. (2018).

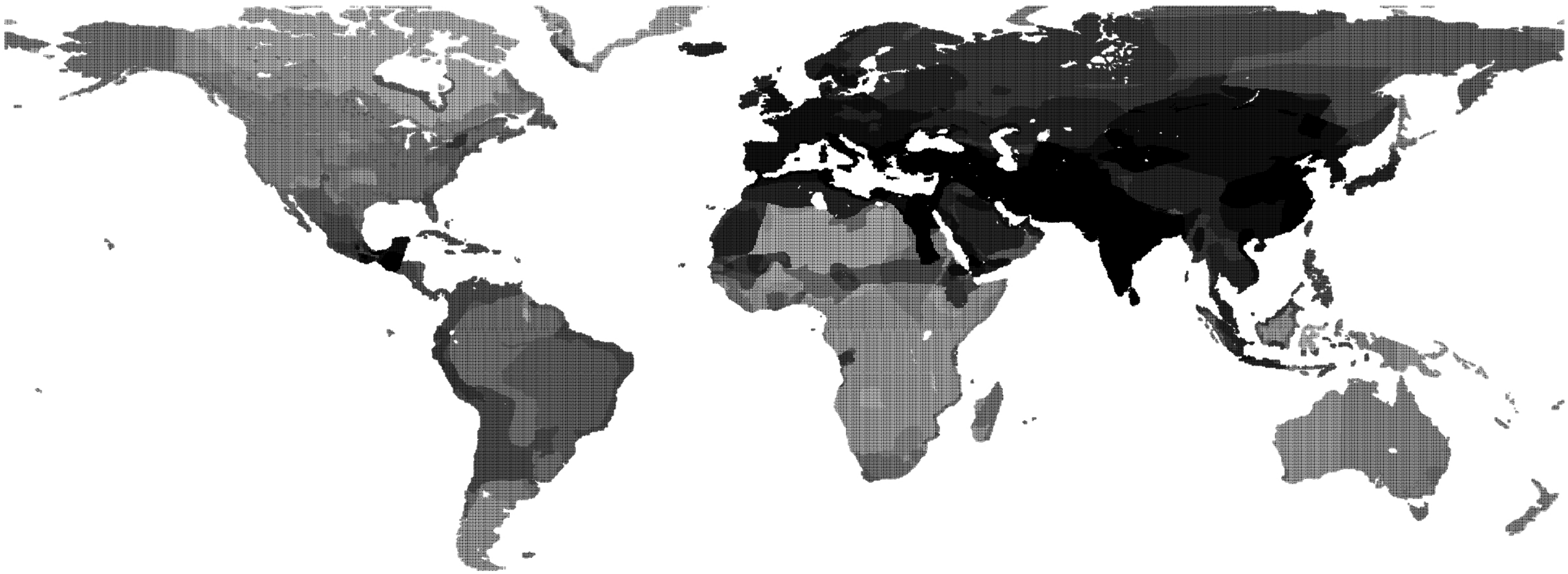
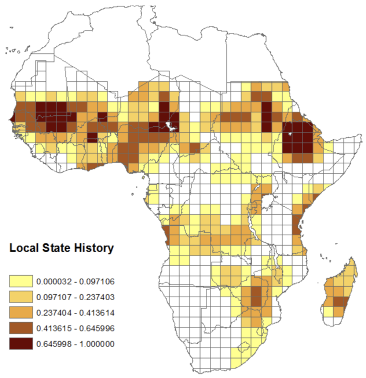
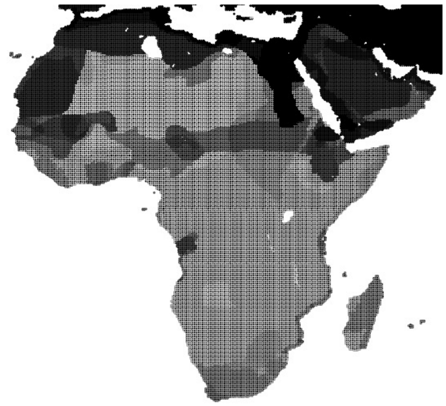


**Figure A.1:** Density Plot (Heatmap) of State History

**Notes:**
This figure plots the cell-level gradient of my state history measure. Darker areas represent longer "civilized" grids.




**(a)** Depetris-Chauvin (Figure 4; 2015)



**(b)** Civ Based

**Figure A.2:** Comparing State History Measures for Sub-Saharan Africa

<u>**Notes:**</u>
Sub-figure (a) plots Depetris-Chauvin's constructed measure of state history (Figure 4 from Depetris-Chauvin (2015)), while sub-figure (b) focuses on Africa from Figure A.1. The measures show considerable overlap.

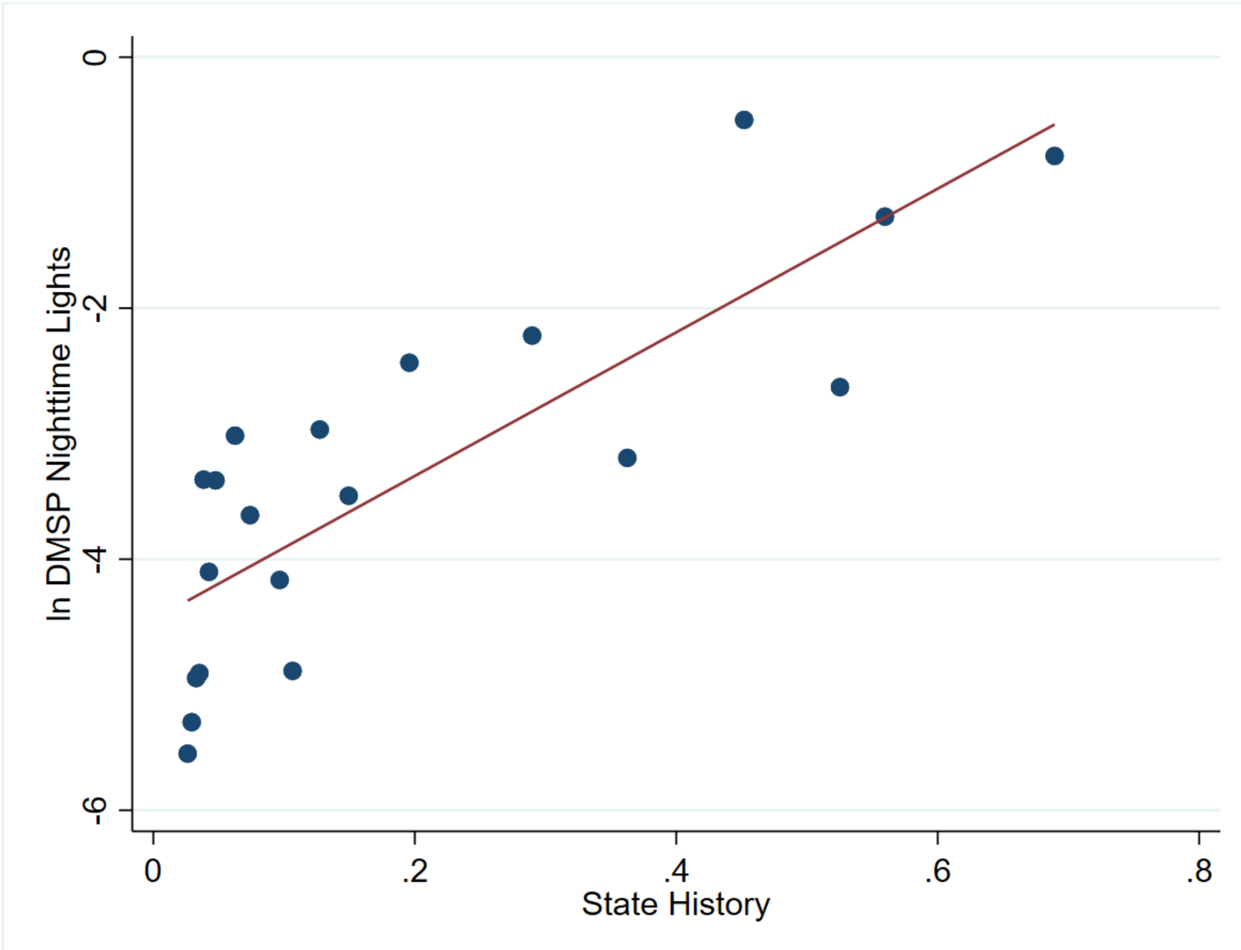



**Figure A.3:** Binscatter: State History and HSSW's ln DMSP Nighttime Lights

**Notes:**
This figure gives the cell-level bivariate binned scatter plot beteen the natural log of DMSP nighttime lights and state history.

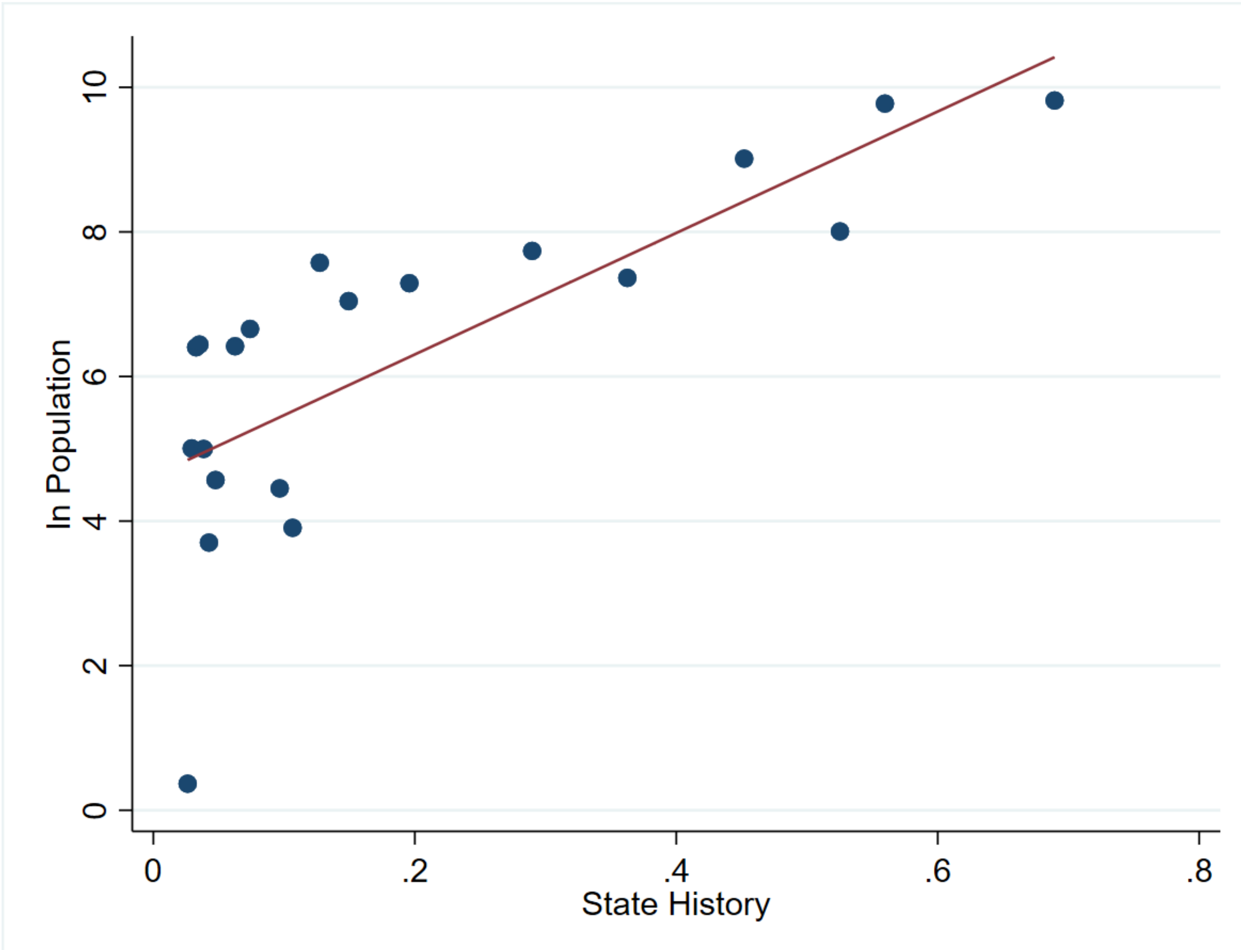



**Figure A.4:** Binscatter: State History and ln Population

**Notes:**
This figure gives the cell-level bivariate binned scatter plot beteen the natural log of SEDAC's population and state history.

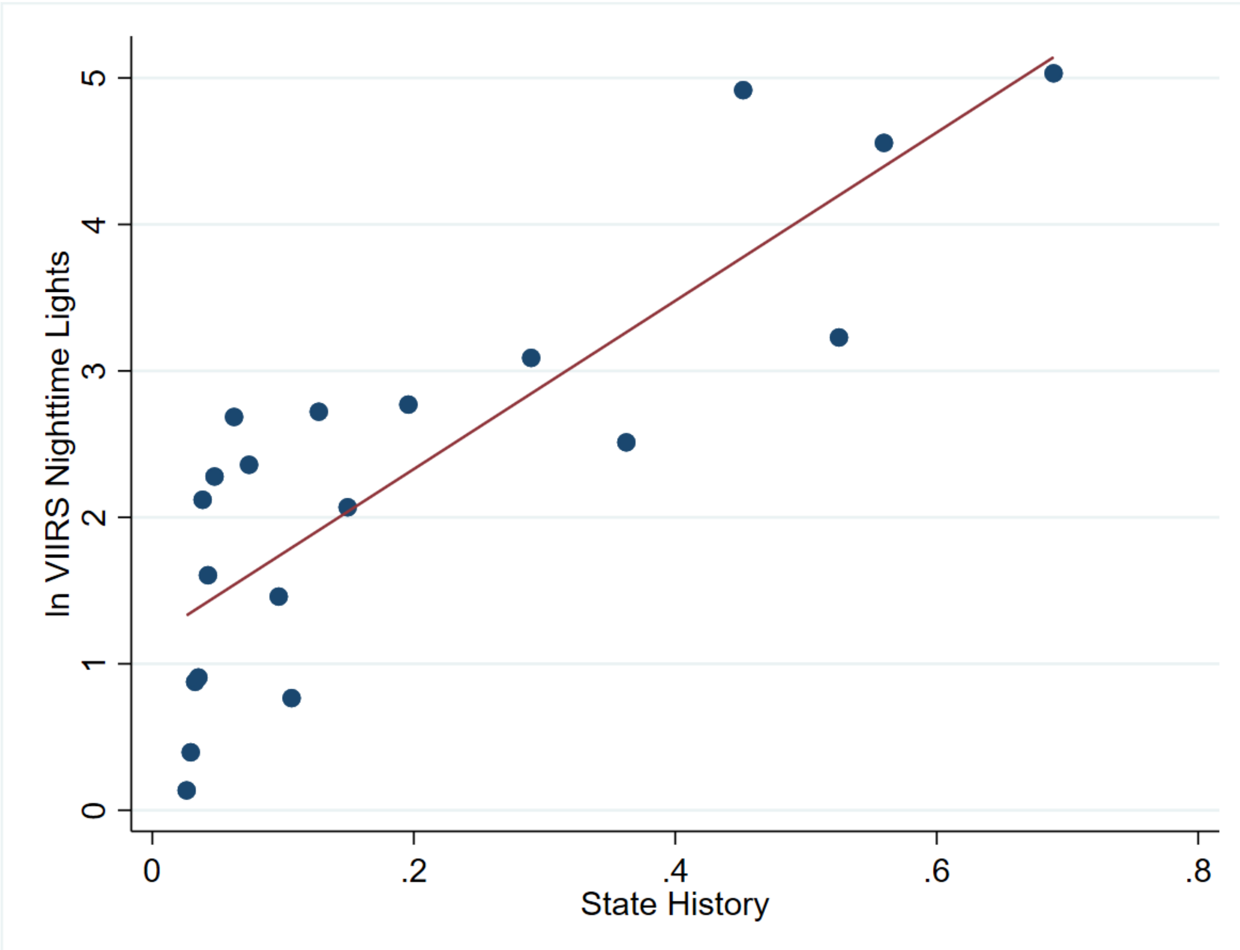



**Figure A.5:** Binscatter: State History and ln VIIRS Nighttime Lights

**Notes:**
This figure gives the cell-level bivariate binned scatter plot beteen the natural log of VIIRS nighttime lights and state history.

# B Agriculture and the Independent Contribution of State History

**Table B.1:** Partial Correlation Matrix (State History Only)

| Variable | Pairwise Correlation Coefficient with State History |
|---|---|
| Base set: | |
| Rugged | 0.1594 |
| Malaria Ecology | -0.2062 |
| Agriculture set: | |
| Biome 1 (tropical moist forest) | -0.0450 |
| Biome 2-3 (tropical dry forest) | 0.0745 |
| Biome 4 (temperate broadleaf) | 0.2101 |
| Biome 5 (temperate conifer) | 0.0224 |
| Biome 6 (boreal forest) | -0.1492 |
| Biome 7-9 (tropical grassland) | -0.2175 |
| Biome 8 (temperate grassland) | 0.0502 |
| Biome 10 (montane grassland) | 0.1289 |
| Biome 11 (tundra) | -0.2211 |
| Biome 12 (Mediterranean forest) | 0.1468 |
| Biome 13 (desert) | 0.1974 |
| Biome 14 (mangroves) | -0.0022 |
| Temperature | 0.0860 |
| Precipitation | -0.1186 |
| Growing Days | -0.0579 |
| Land Suitability | 0.2040 |
| Absolute Latitude | -0.0029 |
| Elevation | 0.1687 |
| Base set: | |
| Coastal Ind. | -0.0433 |
| Dist. to Coast | 0.1917 |
| Harbor w/in 25km | 0.0308 |
| River w/in 25km | 0.0370 |
| Lake w/in 25km | -0.0161 |

**Summary & Notes:** This table provides correlation coefficients for all control variables and state history. All pairwise correlations are statistically significant ($p<0.01$) except for absolute latitude ($p=0.16$) and the Mangrove Biome Indicator ($p=0.27$).

**Table B.2:** Shapley Values

| Outcome: | ln DMSP Nighttime Lights | | SEDAC's ln Population Count | | ln VIIRS Nighttime Lights | |
|---|---|---|---|---|---|---|
| | Shapley Value | Percent | Shapley Value | Percent | Shapley Value | Percent |
| State History | 0.06 | 10.83% | 0.10 | 12.90% | 0.07 | 12.06% |
| Base | 0.01 | 1.45% | 0.03 | 3.61% | 0.01 | 1.22% |
| Agriculture | 0.29 | 49.02% | 0.32 | 41.54% | 0.27 | 49.02% |
| Trade | 0.03 | 4.62% | 0.01 | 1.18% | 0.02 | 3.17% |
| Country FE | 0.20 | 34.09% | 0.32 | 40.77% | 0.19 | 34.53% |
| Total | 0.5855 | 100% | 0.7909 | 100% | 0.5500 | 100% |

**Summary & Notes:** This table provides Shapley values as in Henderson et al. (2018). State history is shown to independently account for roughly 10% of the explained variation (i.e., R squared) for each of the three outcomes; this is behind only the set of country FE and the set of agricultural controls.

**Table B.3:** Lasso Control Ranking:
Top 5 Penalized Coefficients (Standardized; Rank in Parantheses)

| Outcome: | ln DMSP Nighttime Lights | ln Population | ln VIIRS Nighttime Lights |
|---|---|---|---|
| State History | 0.73 (4) | 0.91 (4) | 0.06 (2) |
| Temperature | 1.29 (1) | 1.62 (1) | |
| Australia Indicator | -0.81 (2) | -0.98 (2) | -0.04 (4) |
| Number of Growing Days | 0.79 (3) | 0.97 (3) | 0.08 (1) |
| Land Suitability | 0.69 (5) | 0.65 (5) | |
| Precipitation | | | -0.05 (3) |
| Abs. Latitude | | | -0.04 (5) |

**Summary & Notes:** This table uses a LASSO cross-validation excercise to estimate penalized coefficients for all standardized variables (i.e., mean=0, s.d.=1). The penalized coefficient for (standardized) state history ranges from 2nd to 4th in magnitude out of the 204 variables considered. Additional tests for a LASSO selection of coefficients (double-selection and cross-fit partialing-out) list all used variables as relevant for estimation. This indicates that the state history variable provides valuable variation independent from the set of agricultural controls.

## C Robustness to Unlit/Unpopulated Cells

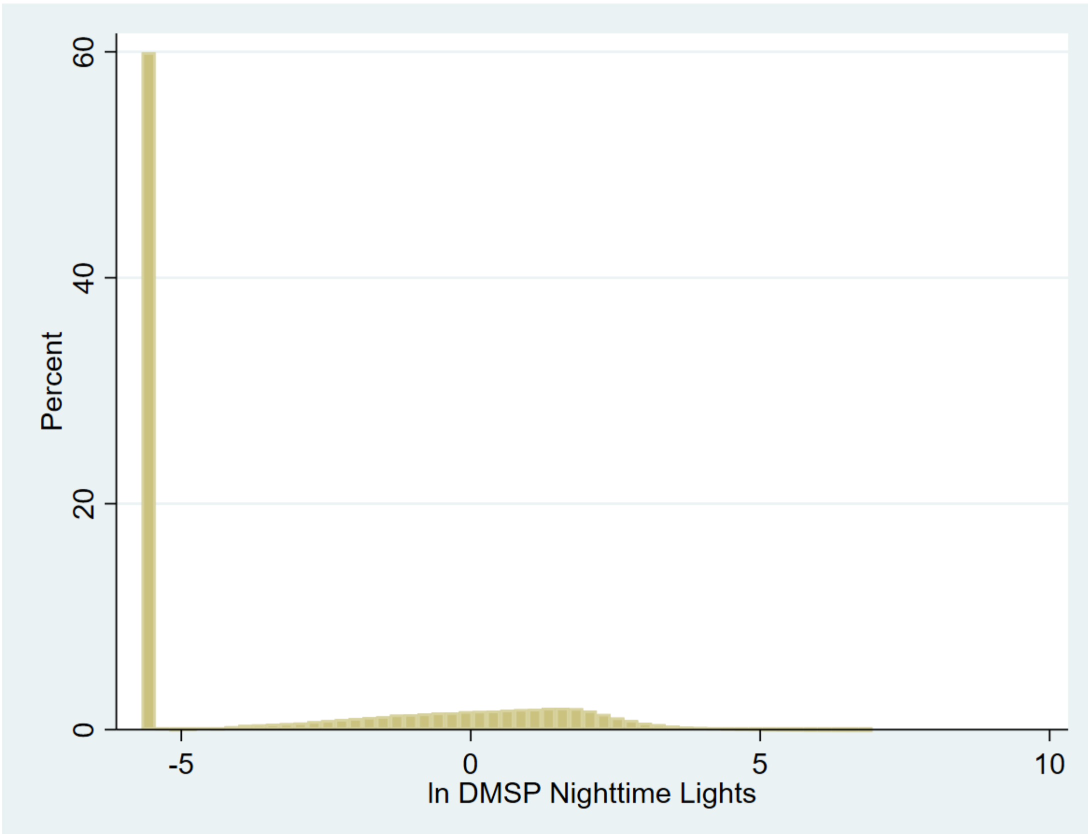



**Figure C.1:** Histogram of Grid DMSP Nighttime Lights

**Notes:**
This figure shows the distribution of DMSP nighttime lights from Henderson et al. (2018). Notice that roughly 60% of values are bottom coded.

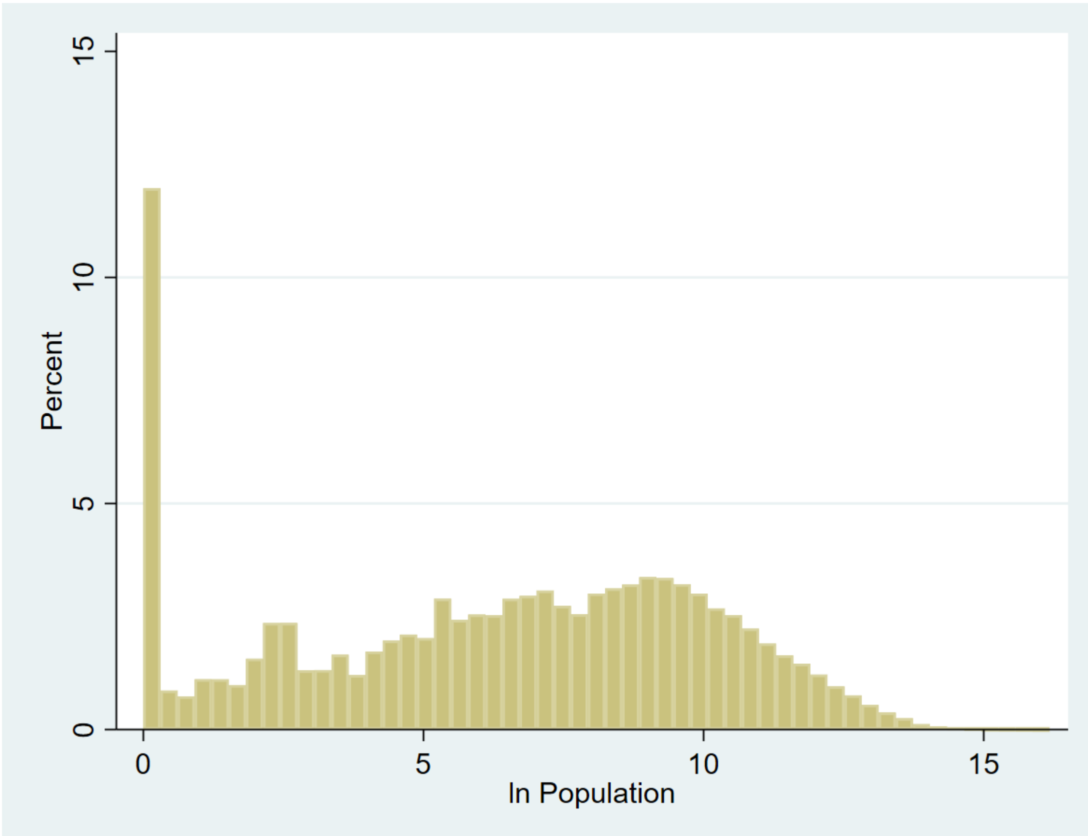



**Figure C.2:** Histogram of Grid Population

**Notes:**
This figure shows the distribution of cell-level population. As with DMSP nighttime lights, there is a large frequency of unpopulated cells.

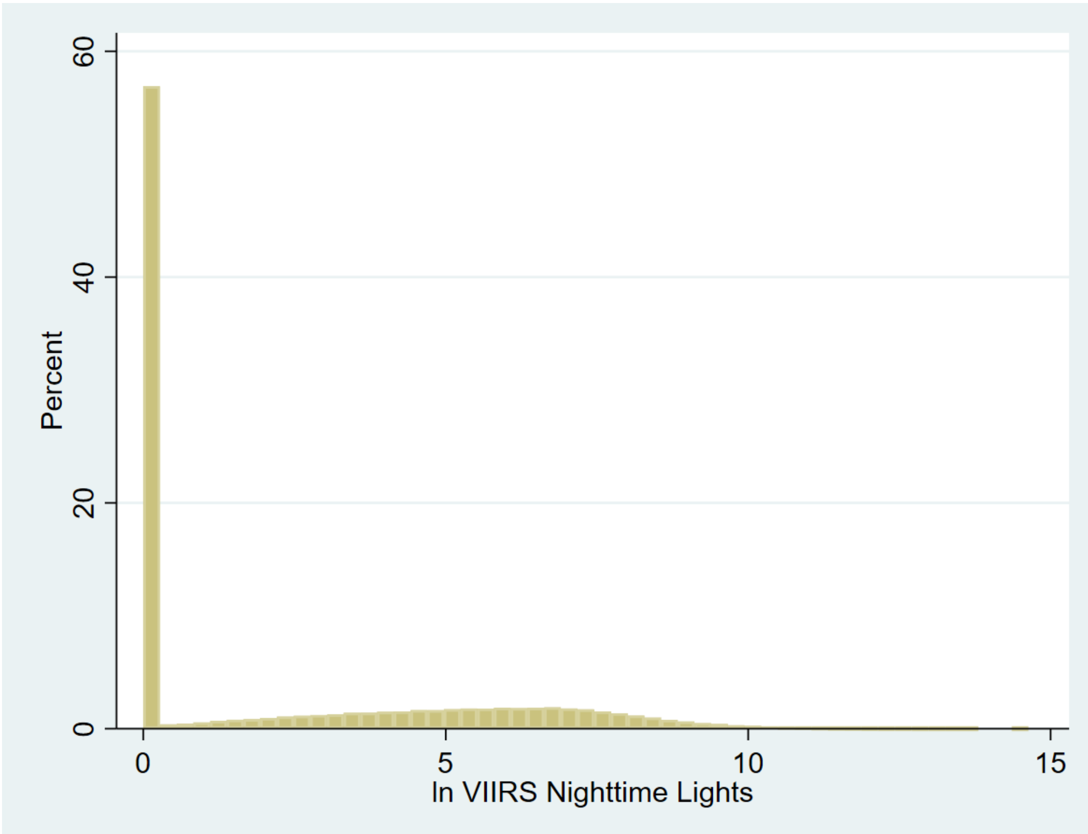



**Figure C.3:** Histogram of Grid VIIRS Nighttime Lights

**Notes:**
This figure shows the distribution of VIIRS nighttime lights. Unlike DMSP nighttime lights and population, the majority of obsersvations are not bottom-coded/zeros.

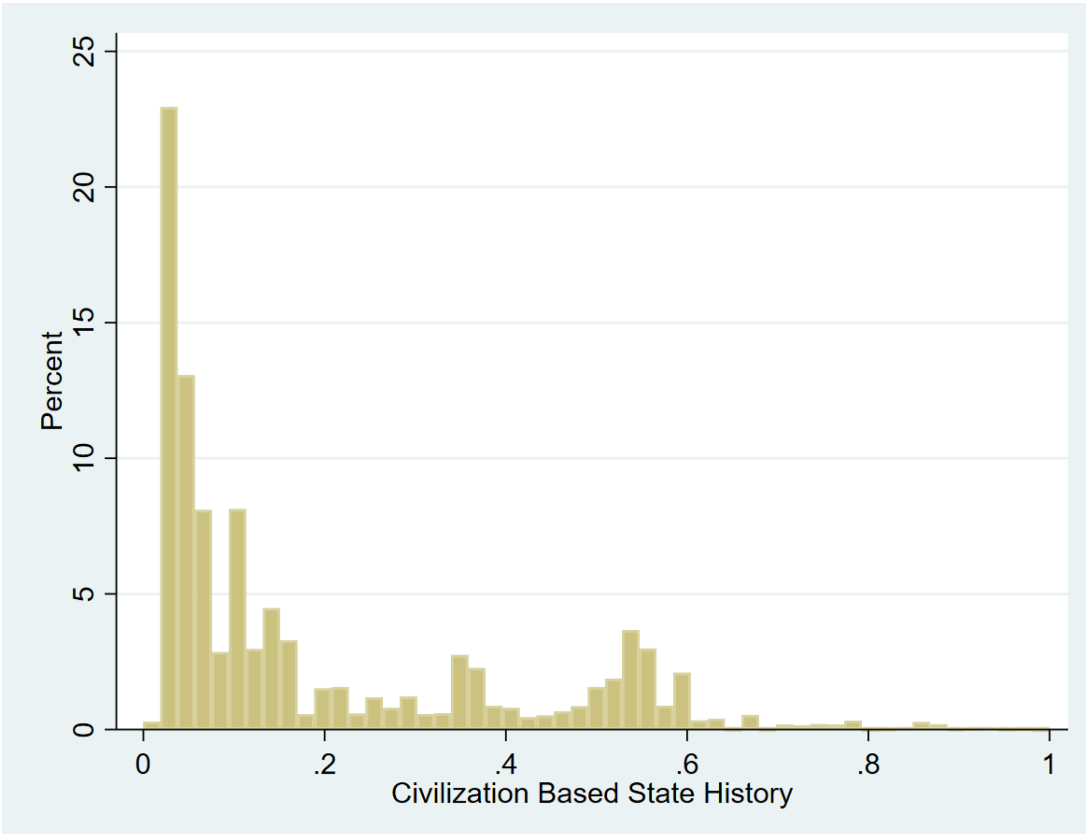



**Figure C.4:** Histogram of Grid State History

**Notes:**
This figure shows the distribution of cell-level state history. Unlike DMSP nighttime lights and population, the majority of obsersvations are not bottom-coded/zeros.

**Table C.1:** Sample Truncation by Outcome Decile

| Decile Omitted from Sample: | Bottom (All "0" cells) | Top | Top & Bottom |
|---|---|---|---|
| | (1) | (2) | (3) |
| | Panel A. ln DMSP Nighttime Lights, 2010 | | |
| State History | 2.3476*** | 2.6803*** | 1.2736*** |
| | (0.5462) | (0.6317) | (0.3958) |
| Observations | 97,181 | 217,966 | 72,963 |
| R Sqr. | 0.3681 | 0.4779 | 0.2232 |
| | Panel B. ln Population, 2010 | | |
| State History | 4.4606*** | 3.7657*** | 3.9422*** |
| | (0.4699) | (0.4239) | (0.4339) |
| Observations | 217,300 | 217,965 | 193,081 |
| R Sqr. | 0.7410 | 0.7585 | 0.7063 |
| | Panel C. ln VIIRS, 2014 | | |
| State History | 2.5828*** | 2.6516*** | 1.4101*** |
| | (0.7599) | (0.5719) | (0.5078) |
| Observations | 104,937 | 217,965 | 80,718 |
| R Sqr. | 0.3254 | 0.4784 | 0.2454 |
| Controls (all panels): | | | |
| Country FE | Y | Y | Y |
| All geoclimatic controls | Y | Y | Y |

**Summary & Notes:** This table omits observations by decile of the outcome. Column (1) omits the bottom decile, which accounts for all "0" observations for both nighttime lights and population measures; column (2) omits the top decile, which is a proxy for urban areas; and column (3) omits bottom and top deciles. Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

# D Additional Sample Adjustments

**Table D.1:** Effect by Large (area) Countries

| Country: | Russia | Canada | USA | China | All other |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | Panel A. ln DMSP Nighttime Lights, 2010 | | | | |
| State History | $0.7040^{***}$ | $24.1765^{***}$ | $20.9960^{***}$ | $1.4513^{***}$ | $4.6311^{***}$ |
| | (0.2716) | (2.2864) | (2.2034) | (0.3707) | (0.7250) |
| Observations | 46,179 | 25,997 | 18,423 | 15,337 | 136,248 |
| R Sqr. | 0.5351 | 0.5789 | 0.6422 | 0.6201 | 0.5871 |
| | Panel B. ln Population, 2010 | | | | |
| State History | $3.0000^{***}$ | $26.0365^{***}$ | $18.3790^{***}$ | $2.2223^{***}$ | $4.4464^{***}$ |
| | (0.2030) | (2.5289) | (2.3118) | (0.4017) | (0.5951) |
| Observations | 46,179 | 25,997 | 18,423 | 15,337 | 136,248 |
| R Sqr. | 0.7665 | 0.6335 | 0.7386 | 0.6419 | 0.7276 |
| | Panel C. ln VIIRS Nighttime Lights, 2014 | | | | |
| State History | $0.9180^{***}$ | $21.3653^{***}$ | $11.8641^{***}$ | $3.2963^{***}$ | $4.6540^{***}$ |
| | (0.2698) | (2.2457) | (2.2315) | (0.3634) | (0.7925) |
| Observations | 46,179 | 25,997 | 18,423 | 15,337 | 136,248 |
| R Sqr. | 0.4940 | 0.4648 | 0.6406 | 0.5431 | 0.5371 |
| Controls: | | | | | |
| Country FE | | | | | Y |
| All geoclimatic controls | Y | Y | Y | Y | Y |
| Standard Errors: | | | | | |
| Couuntry cluster | | | | | Y |
| HSSW cluster | Y | Y | Y | Y | |

**Table D.2:** Effect by Continent
State History Standardized (mean=0, s.d.=1) Within Continent

| | Old World | | | | New World | | | |
|---|---|---|---|---|---|---|---|---|
| Sample: | Africa (1) | Asia (2) | Europe (3) | Old World (4) | South Am. (5) | North Am. (6) | Oceania (7) | New World (8) |
| Panel A. ln Nighttime Lights, 2010 | | | | | | | | |
| State History | 0.3369*** | 0.9699*** | 0.1449 | 0.7208*** | 0.9784*** | 0.2573 | 0.1477*** | 0.5611** |
| | (0.0480) | (0.2156) | (0.1040) | (0.1577) | (0.0837) | (0.1770) | (0.0229) | (0.2769) |
| Observations | 41,346 | 51,258 | 60,050 | 152,654 | 25,369 | 50,941 | 13,220 | 89,530 |
| R Sqr. | 0.4054 | 0.5354 | 0.7136 | 0.6242 | 0.4120 | 0.6658 | 0.3227 | 0.5549 |
| Panel B. ln Population, 2010 | | | | | | | | |
| State History | 0.5258*** | 0.8531*** | 0.3890*** | 0.9478*** | 0.8148*** | 0.1713 | 0.2316*** | 0.4925* |
| | (0.0874) | (0.1033) | (0.0310) | (0.0762) | (0.0826) | (0.1692) | (0.0330) | (0.2584) |
| Observations | 41,346 | 51,258 | 60,050 | 152,654 | 25,369 | 50,941 | 13,220 | 89,530 |
| R Sqr. | 0.6849 | 0.6542 | 0.8342 | 0.7693 | 0.4868 | 0.7923 | 0.6698 | 0.7666 |
| Panel C. ln VIIRS, 2014 | | | | | | | | |
| State History | 0.3544*** | 1.0456*** | 0.1946* | 0.7953*** | 0.9931*** | 0.1804 | 0.1512*** | 0.5134** |
| | (0.0846) | (0.1878) | (0.0987) | (0.1657) | (0.0821) | (0.1299) | (0.0194) | (0.2533) |
| Observations | 41,346 | 51,258 | 60,050 | 152,654 | 25,369 | 50,941 | 13,220 | 89,530 |
| R Sqr. | 0.3943 | 0.4931 | 0.6573 | 0.5816 | 0.4061 | 0.6487 | 0.2269 | 0.5362 |
| Controls: | | | | | | | | |
| Country FE + All geoclimatic controls | Y | Y | Y | Y | Y | Y | Y | Y |

**Summary & Notes:** Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

# E Baseline Results with All Coefficients

**Table E.1:** Table 1, Panel A with Coefficients of Control Variables

| | Dependent Variable: ln Nighttime Lights, 2010 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| State History | 5.7222*** | 8.0480*** | 8.0173*** | 3.4938*** | 7.9298*** | 3.5624*** |
| | (0.9966) | (0.8897) | (0.8573) | (1.0009) | (0.9548) | (0.8074) |
| Ruggedness (000s of index) | | | -0.0690*** | -0.0173*** | -0.0659*** | -0.0201*** |
| | | | (0.0207) | (0.0050) | (0.0215) | (0.0051) |
| index of the stability of malaria transmission, | | | -0.0226 | -0.0509*** | -0.0222 | -0.0440*** |
| Kiszewski et al. (2004) | | | (0.0229) | (0.0123) | (0.0231) | (0.0105) |
| tropical moist forest | | | | -0.2037 | | -0.2392 |
| | | | | (0.3282) | | (0.2829) |
| tropical dry forest | | | | 0.2838 | | 0.2253 |
| | | | | (0.2425) | | (0.2140) |
| temperate broadleaf | | | | 1.3612*** | | 1.1914*** |
| | | | | (0.2998) | | (0.3652) |
| temperate conifer | | | | 0.2986 | | 0.0725 |
| | | | | (0.2799) | | (0.3067) |
| boreal forest | | | | -1.0263 | | -1.2595* |
| | | | | (0.6220) | | (0.6696) |
| tropical grassland | | | | -0.0262 | | -0.0388 |
| | | | | (0.2102) | | (0.1952) |
| temperate grassland | | | | 0.8631*** | | 0.8855*** |
| | | | | (0.2728) | | (0.2781) |
| montane grassland | | | | 0.9553*** | | 0.6692** |
| | | | | (0.2714) | | (0.3015) |
| tundra | | | | -0.9123* | | -1.4013** |
| | | | | (0.5403) | | (0.6802) |
| Mediterranean forest | | | | 1.3232*** | | 1.0489** |
| | | | | (0.4520) | | (0.4544) |
| mangroves | | | | -0.1058 | | -0.4901 |
| | | | | (0.3783) | | (0.3238) |
| Average monthly temperature (1960-1990 mon avg) | | | | 0.0995*** | | 0.0972*** |
| | | | | (0.0295) | | (0.0280) |
| Monthly total precipitation, mm/month (1960-1990 mon avg) | | | | -0.0097*** | | -0.0105*** |
| | | | | (0.0021) | | (0.0019) |
| Length of growing period, days | | | | 0.0085*** | | 0.0082*** |
| | | | | (0.0016) | | (0.0016) |
| Land suitability as prob that cell is cultivated, | | | | 2.1624*** | | 2.1704*** |
| Ramankutty (2002) | | | | (0.3471) | | (0.3222) |
| abs(latitude) | | | | 0.0289 | | 0.0283 |
| | | | | (0.0300) | | (0.0290) |
| Elevation, km above sea level | | | | -0.0992 | | 0.0737 |
| | | | | (0.1709) | | (0.1983) |
| Binary variable for if the coast passes inside the grid square | | | | | -0.0889 | 0.1857 |
| | | | | | (0.1683) | (0.1248) |
| Distance to the nearest coast, 000s km | | | | | -0.4093 | -0.7094** |
| | | | | | (0.6428) | (0.2852) |
| 1(within 25km of natural harbor) | | | | | 1.8356*** | 1.2462*** |
| | | | | | (0.2833) | (0.1780) |
| 1(within 25km of navigable river) | | | | | 0.8461*** | 0.6406*** |
| | | | | | (0.1824) | (0.1442) |
| 1(within 25km of big lake (area¿5000 sq km)) | | | | | 0.9003*** | 0.5475** |
| | | | | | (0.2975) | (0.2206) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.1408 | 0.3965 | 0.4068 | 0.5747 | 0.4196 | 0.5857 |

**Summary & Notes:** Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

## **Table E.2:** Table 1, Panel B with Coefficients of Control Variables

| | Dependent Variable: ln Population, 2010 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| State History | 8.4013*** | 9.4568*** | 9.4860*** | 4.2796*** | 9.4541*** | 4.4186*** |
| | (1.8863) | (1.4901) | (1.4450) | (0.5559) | (1.5534) | (0.4679) |
| Ruggedness (000s of index) | | | -0.0180 | 0.0195 | -0.0200 | 0.0146 |
| | | | (0.0276) | (0.0143) | (0.0288) | (0.0145) |
| index of the stability of malaria transmission, | | | 0.0468 | -0.0064 | 0.0440 | -0.0034 |
| Kiszewski et al. (2004) | | | (0.0287) | (0.0207) | (0.0296) | (0.0188) |
| tropical moist forest | | | | 0.2425 | | 0.1702 |
| | | | | (0.3781) | | (0.3407) |
| tropical dry forest | | | | 0.5735** | | 0.5072** |
| | | | | (0.2498) | | (0.2289) |
| temperate broadleaf | | | | 1.9912*** | | 1.7896*** |
| | | | | (0.3277) | | (0.4042) |
| temperate conifer | | | | 1.0641*** | | 0.8876** |
| | | | | (0.2375) | | (0.3421) |
| boreal forest | | | | -0.3693 | | -0.6934 |
| | | | | (0.9537) | | (1.0663) |
| tropical grassland | | | | 0.7482** | | 0.7288** |
| | | | | (0.3068) | | (0.3029) |
| temperate grassland | | | | 1.3925*** | | 1.3914*** |
| | | | | (0.2153) | | (0.2259) |
| montane grassland | | | | 1.0861*** | | 0.9880*** |
| | | | | (0.2866) | | (0.2972) |
| tundra | | | | -0.4480 | | -0.9766 |
| | | | | (0.7471) | | (0.9660) |
| Mediterranean forest | | | | 1.8362*** | | 1.6433*** |
| | | | | (0.3190) | | (0.3786) |
| mangroves | | | | -0.4485 | | -0.3183 |
| | | | | (0.3413) | | (0.3493) |
| Average monthly temperature (1960-1990 mon avg) | | | | 0.1260*** | | 0.1239*** |
| | | | | (0.0397) | | (0.0377) |
| Monthly total precipitation, mm/month (1960-1990 mon avg) | | | | -0.0090*** | | -0.0091*** |
| | | | | (0.0026) | | (0.0024) |
| Length of growing period, days | | | | 0.0103*** | | 0.0099*** |
| | | | | (0.0019) | | (0.0017) |
| Land suitability as prob that cell is cultivated, | | | | 2.0652*** | | 2.0251*** |
| Ramankutty (2002) | | | | (0.2786) | | (0.2347) |
| abs(latitude) | | | | 0.0128 | | 0.0132 |
| | | | | (0.0316) | | (0.0301) |
| Elevation, km above sea level | | | | 0.2163 | | 0.3109 |
| | | | | (0.2558) | | (0.2579) |
| Binary variable for if the coast passes inside the grid square | | | | | -0.9217*** | -0.5396** |
| | | | | | (0.2303) | (0.2403) |
| Distance to the nearest coast, 000s km | | | | | -0.4207 | -0.8582** |
| | | | | | (0.8834) | (0.3898) |
| 1(within 25km of natural harbor) | | | | | 1.5326*** | 0.7074*** |
| | | | | | (0.3473) | (0.1479) |
| 1(within 25km of navigable river) | | | | | 0.6597*** | 0.4325*** |
| | | | | | (0.2216) | (0.1458) |
| 1(within 25km of big lake (area¿5000 sq km)) | | | | | 0.4746* | 0.0554 |
| | | | | | (0.2819) | (0.1333) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.2100 | 0.6147 | 0.6168 | 0.7749 | 0.6238 | 0.7809 |

**Summary & Notes:** Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

## **Table E.3:** Table 1, Panel C with Coefficients of Control Variables

| | Dependent Variable: ln VIIRS, 2014 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| State History | 5.7467*** | 7.9630*** | 7.9365*** | 3.7159*** | 7.9128*** | 3.8184*** |
| | (1.0050) | (0.9092) | (0.8898) | (1.0285) | (0.9718) | (0.8424) |
| Ruggedness (000s of index) | | | -0.0467** | -0.0043 | -0.0473** | -0.0083 |
| | | | (0.0224) | (0.0062) | (0.0232) | (0.0065) |
| index of the stability of malaria transmission, | | | -0.0235 | -0.0498*** | -0.0253 | -0.0465*** |
| Kiszewski et al. (2004) | | | (0.0219) | (0.0113) | (0.0229) | (0.0106) |
| tropical moist forest | | | | -0.2333 | | -0.2914 |
| | | | | (0.3503) | | (0.3161) |
| tropical dry forest | | | | 0.2324 | | 0.1754 |
| | | | | (0.2153) | | (0.1935) |
| temperate broadleaf | | | | 1.2178*** | | 1.0436** |
| | | | | (0.3404) | | (0.4032) |
| temperate conifer | | | | 0.2069 | | 0.0381 |
| | | | | (0.2362) | | (0.2671) |
| boreal forest | | | | -0.8737* | | -1.1483** |
| | | | | (0.4950) | | (0.5778) |
| tropical grassland | | | | 0.0418 | | 0.0248 |
| | | | | (0.1786) | | (0.1701) |
| temperate grassland | | | | 0.7910*** | | 0.7934*** |
| | | | | (0.2346) | | (0.2690) |
| montane grassland | | | | 0.7705** | | 0.6402** |
| | | | | (0.2987) | | (0.3208) |
| tundra | | | | -0.6602 | | -1.1218* |
| | | | | (0.4175) | | (0.5815) |
| Mediterranean forest | | | | 1.1787** | | 0.9895** |
| | | | | (0.4836) | | (0.4818) |
| mangroves | | | | -1.0503*** | | -1.0317*** |
| | | | | (0.3076) | | (0.2844) |
| Average monthly temperature (1960-1990 mon avg) | | | | 0.1071*** | | 0.1052*** |
| | | | | (0.0221) | | (0.0217) |
| Monthly total precipitation, mm/month (1960-1990 mon avg) | | | | -0.0097*** | | -0.0099*** |
| | | | | (0.0022) | | (0.0019) |
| Length of growing period, days | | | | 0.0079*** | | 0.0076*** |
| | | | | (0.0017) | | (0.0016) |
| Land suitability as prob that cell is cultivated, | | | | 2.1571*** | | 2.1307*** |
| Ramankutty (2002) | | | | (0.3155) | | (0.2938) |
| abs(latitude) | | | | 0.0300 | | 0.0301 |
| | | | | (0.0303) | | (0.0298) |
| Elevation, km above sea level | | | | 0.0839 | | 0.1866 |
| | | | | (0.1618) | | (0.1919) |
| Binary variable for if the coast passes inside the grid square | | | | | -0.6400*** | -0.3541** |
| | | | | | (0.1895) | (0.1602) |
| Distance to the nearest coast, 000s km | | | | | -0.4371 | -0.7425** |
| | | | | | (0.5685) | (0.2953) |
| 1(within 25km of natural harbor) | | | | | 1.3334*** | 0.7915*** |
| | | | | | (0.2225) | (0.1404) |
| 1(within 25km of navigable river) | | | | | 0.7090*** | 0.5463*** |
| | | | | | (0.1707) | (0.1169) |
| 1(within 25km of big lake (area¿5000 sq km)) | | | | | 0.4980*** | 0.1893 |
| | | | | | (0.1848) | (0.1326) |
| Observations | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 | 242,184 |
| R Sqr. | 0.1471 | 0.3843 | 0.3895 | 0.5426 | 0.3977 | 0.5499 |

**Summary & Notes:** Statistical significance is determined by country-clustered standard errors and denoted at the 1, 5, and 10% levels respectively by ***, **, and *.

## F References for Civilization Locations

Maps originally created and compiled by Andrew Tollefson; data taken or derived from:

Canada

• Cordell, Linda S.; Lightfoot, Kent; McManamon, Francis; Milner, George (2008). Archaeology in America: An Encyclopedia, Vol. 1: Northeast and Southeast. ABC-CLIO.
• Parks Canada Directory of Federal Heritage Designations
• Canadian Register of Historic Places
• BC Geographical Names Information System
• Indian treaties and surrenders, from 1680 to 1890. Ottawa: B. Chamberlin (1891).
• Collections of the Maine Historical Society. Series II, vol.1 (1890), pp.179-196.

USA

• National Register of Historic Places
• Relations des Jésuites de la Nouvelle-France
• 1997-2014 Edition of The National Atlas of the United States
• National Park Service: El Camino Real de Tierra Adentro historic map
• Albert, George Dallas, C. M. Busch (1896). The Frontier Forts of Western Pennsylvania, pp.536-586.
• United States Statutes at Large, Volume 7
• Shumway, Burgess M. (1988). California Ranchos: Patented Private Land Grants Listed by County. San Bernardino: The Borgo Press.
• Hayes, Derek (1999). Historical Atlas of the Pacific Northwest: Maps of Exploration and Discovery. Sasquatch Books.
• Hiram Bingham I (1848). A Residence of Twenty-one Years in the Sandwich Islands, pp.40-52.
• Instituto Nacional de Estadistica y Geografia: https://commons.wikimedia.org/wiki/File:Republic_of_the_Rio_Grande_(location_map_scheme).svg
• http://www.westpoint.edu/history/sitepages/our%20atlases.aspx

Antilles

• Mapa Historico de la Trocha Jucaro-Moron Construida per orden del Conde Valmaseda, 1871

Mexico

• Coe, Michael; Snow, Dean; Benson, Elizabeth (1986). Atlas of Ancient America. New York: Facts on File, pp.94-95

• Colman, Arlene (2010). The Construction of Complex A at La Venta, Tabasco, Mexico: A History of Buildings, Burials, Offerings, and Stone Monuments, p.54.

• Script Delivery: New World writing takes disputed turn. Science News December 7th, 2002, Vol.162 #23

• Marcus, Joyce; Flannery, Kent V. (1996). Zapotec Civilization: How Urban Society Evolved in Mexico's Oaxaca Valley.

• Saturno, William A., David Stuart, Boris Beltran (2006). Early Maya Writing at San Bartolo, Guatemala. Science. 311 (5765): 1281-1283.

• Sharer, Robert J., Loa P. Traxler (2006). The Ancient Maya (6th ed.). Stanford University Press.

• Cowgill, George (1997). State and Society at Teotihuacan, Mexico. Annual Review of Anthropology. Palo Alto: Annual Reviews Inc. 26 (1): 129-161.

• Wilkerson, S. Jeffery K. (1987). El Tajin: A Guide for Visitors. ISBN 968-499-293-9.

• George L. Cowgill (2004). Ancient Tollan: Tula and the Toltec heartland. 78 (299). Antiquity: 226-227.

• Pohl, John M.D. (2002). The Legend of Lord Eight Deer: An Epic of Ancient Mexico. Oxford; New York: Oxford University Press.

• María del Carmen Solanes Carraro y Enrique Vela Ramírez (2000). Atlas del México prehispánico, edicion especial numero 5 de Arqueología Mexicana. Mexico.

• Powell, Philip Wayne. Soldiers, Indians, & Silver: The Northward Advance of New Spain, 1550-1600. Berkeley: University of California Press, 1952 (republished 1969).

• Engelhardt, Zephyrin (1916). Missions and Missionaries. San Francisco: The James H. Barry Co.

• Enriqueta García de Miranda and Zaida Falcón de Gyves (1972). Nuevo Atlas Porrúa de la República Mexicana. Editorial Porrúa, S.A.

Guatemala

• Francis Polo Sifontes; Celso A. Lara Figueroa, eds. Título de Alotenango, 1565: Clave para ubicar geograficamente la antigua Itzcuintepec pipil. Antropología e Historia de Guatemala. Guatemala City, Guatemala: Dirección General de Antropología e Historia de Guatemala, Ministerio de Educación. 3, II Epoca: 109–129, pp.61-62.

• Jones, Grant D. (1998). The Conquest of the Last Maya Kingdom, p.4. Stanford University Press.

Honduras

• Chamberlain, Robert Stoner (1953). The Conquest and Colonization of Honduras.

Nicaragua

• Duncan, David Ewing (1995). Hernando de Soto - A Savage Quest in the Americas – Book II: Consolidation. New York: Crown Publishers, Inc.

• E.G. Squier (1894), The States of Central America, pp.634-638.

Panama

• The National Map Small-Scale Collection (USGS)

Colombia

• Revista credencial Historia (1993). Atlas básico de historia de Colombia. Bogotá, Colombia.

Venezuela

• Bartholomew, J.G.; Edinburgh Geographical Institute (1896). The British Guiana frontier. Scottish Geographical Magazine: 12-2, p.88.

Peru

• Moseley, Michael (1992). The Incas and Their Ancestors - The Archaeology of Peru (Rev. Ed.), pp.4,244.

Bolivia

• Hudson, Rex A., Hanratty, Dennis Michael, Weil, Thomas E. (1991). Bolivia: a country study, Washington D.C.: Library of Congress, p.24.

Paraguay

• Colton, G.W. (1865). Argentina Republic, Chile, Uruguay & Paraguay. David Rumsey Historical Map Collection.

Chile

• Galdames, Osvaldo (2005). Atlas de Historia de Chile, p. 48.

Argentina

• Ratto, Silvia (2003). Una Experiencia fronterize exitosa: El negocio pacífico de indios en la provincia de Buenos Aires (1829-1852). Revista de Indias. 13-227, pp.191-222.

• Nellar, Fued (1975). La Expedición al Desierto (1833). Política seguida con el aborigen.

Uruguay

• Barragán, Daniel (2009). Del Sistema de los Pueblos Libres o Protectorado a la derrota de Artigas, p.2.

Brazil

• Sistema Informativo Territoriale Comunale SITGEO

• Ministério das Relações Exteriores mapa da Questão do Pirara

Europe & Periphery

• Overy, Richard (2000). The Times Atlas of World History. Times Books.

• Shepherd, William R. (1911). The Historical Atlas. H. Holt.

• Droysen, G. (1886). Allgemeiner Historischer Handatlas: Mitteleuropa zur Zeit der Staufer. R. Andrée.

Great Britain

• Encyclopædia Britannica. Encyclopedia Britannica Inc.

• Jones, B.; Mattingly, D. (1990). An Atlas of Roman Britain, pp.317-318.

• Frere, Sheppard (1999). Britannia: A History Of Roman Britain: A History Of England.

• Oxford Companion to Scottish History. Oxford University Press, pp.161-162.

• McDonald, R. Andrew (2007). Manx Kingship in its Irish sea Setting, 1187–1229.

France

• Vidal-Lablache (1894). Atlas général d'histoire et de géographie.

• Élegoët, Louis (2000). Limites successives de la Bretagne du VI au IX siècle. CRDP de Bretagne: Bretagne une histoire, p.64.

• The Hundred Years' War after 1400. Dorling Kindersley.

Spain

• Luís Fraga da Silva (2004). Povos Pré-Romanos da Península Ibérica: Esboço de carta etnológica nos finais da 2 Guerra Púnica (circa 200 a.C.). Tavira, Portugal: Associação Campo Arqueológico de Tavira.

• García de Cortázar Ruíz de Aguirre, Fernando (2005). Atlas de historia de España. Editorial Planeta.

• HansenBCN (2008). Mapa de la expansión de la Corona de Aragón en territorio balear y peninsular coloreado según monarcas y fases de conquista y fechado.

• Abreu y Galindo, J. De (1977), Historia de la conquista de las siete islas de Canarias, Tenerife: A. Cioranescu (ed) Goya ediciones.

Germany
• Barford, Paul M. (2001). The Early Slavs, p.356.
• Harck, Ole; Lübke, Christian (1997). Zwischen Reric und Bornhöved: Die Beziehungen zwischen den Dänen und ihren slawischen Nachbarn vom 9. Bis ins 13. Jahrhundert. Beiträge einer internationalen Konferenz, Leipzig, 4.-6. December 1997.
• Droysen, G. (1886). Allgemeiner Historischer Handatlas: Mitteleuropa zur Zeit der Staufer, Plate 26/27. R. Andrée.

Denmark
• Looijenga, Tineke (2004). Texts and Contexts of the Oldest Runic Inscriptions.
• Jansson, Sven B.F. (1987). Runinskrifter i Sverige, pp.177-178,187.

Norway
• Krag, Claus (1995). Vikingtid og rikssamling 800-1130. Aschehougs Norgeshistorie.
• Ekerwald, Carl-Göran (2004). Jämtarnas historia intill 1319, pp.114-115,124.

Finland
• Uino, Pirjo (1997). Ancient Karelia, p. 346. Suomen muinaismuistoyhdistyksen aikakausikirja 104.

Switzerland
• Gasser, Adolf (1932). Die territoriale Entwicklung der Eidgenossenschaft 1291–1797.

Austria
• Daim, Falko; Wolfram, Herwig (1980). Die Völker an der mittleren und unteren Donau im fünften und sechsten Jahrhundert.
• Šavli, Jožko (1995). Sovenija: Podoba evropskega naroda, p.19.

Czech Republic
• Spiesz, Anton; Caplovic, Dusan (2006). Illustrated Slovak History: a Struggle for Sovereignty in Central Europe, map: p.365.
Hungary
• Bartl, Július (2002). Slovak History: Chronology & Lexicon, pp.19-23.
• Fine, J. (1991). The Early Medieval Balkans, A Critical Survey from the Sixth to the Late Twelfth Century, p.78.
Slovakia
• Mapy Uhorska a Slovenska (2007). Nitrianske kniežatstvo r. 800 až 860.

Slovenia

• Štih, Peter (1986). Regnum Carantanum. Zgodovinski časopis 40-30, p.218.

• Štih, Peter (1996). Studien zur Geschichte der Grafen von Görz. Die Ministerialen und Milites der Grafen von Görz in Istrien und Krain. Oldenbourg.

Italy

• The Penguin Historical Atlas of Ancient Rome. Penguin Books.

• Lazenby, John Francis (1996). The First Punic War: A Military History. Stanford, California: Stanford University Press.

• Bury, J.B. (2011). History of the Later Roman Empire, Volume 1: From the Death of Theodosius I to the Death of Justinian. Dover Publications.

• Paulus Diaconus. Historia Langobardorum. Mondadori (1992).

• Benvenuti, G. (1989). Le repubbliche marinare. Amalfi, Pisa, Genova e Venezia. Newton Compton Editori.

Balkan Peninsula

• Hammond, N. G. L. (1986). A History of Greece to 322 B.C. Oxford University Press.

• Hammond; Walbank (1988). A History of Macedonia: 336-167 B.C. Clarendon Press.

• John Fine, Jr. (1991). The Early Medieval Balkans. University of Michigan Press.

• Felczak, Waclaw; Wasilewski, Tadeusz (1985). Historia Jugoslawii. Zaklad Narodowy im. Ossolinskich.

• Nicol, Donald (1984). The Despotate of Epiros 1267-1479: A Contribution to the History of Greece in the Middle Ages. Cambridge University Press.

Romania

• Gudea, Nicolae (1997). Der dakische Limes, Materalien zu seiner Geschichte. Jahrbuch des Römisch-Germanischen Zentralmuseums Mainz 44-2, pp.497-609.

Greece

• Walbank; Astin; Frederiksen; Ogilvie (1984). The Cambridge Ancient History. Cambridge University Press.

• Ostrogorsky, George (1969). History of the Byzantine State. Rutgers University Press.

Turkey

• Hewsen, Robert H. (2001). Armenia: A Historical Atlas. The University of Chicago Press.

• Greatrex, G.; Lieu S.N.C (1993). The Roman Eastern Frontier and the Persian Wars (Part II, 363–630 AD). Routledge.

• Tarih Atlası. Kanaat Yayınları (2010).
Ukraine
• Victorovich, M. (2000). Ethnic components of Chernyakhov culture. Stratum Plus 2000-4, pp.322-330.
• Atlas of History and Civilization for the 11th Grade.
• Zbiorowa, Praca (2008). Wielki atlas historyczny. Warsaw: Demart.
Poland
• Kaliff, Anders (2001). Gothic Connections. Contacts between eastern Scandinavia and the southern Baltic coast 100 BC – 500 AD. Department of Archaeology and Ancient History, Uppsala University.
• Sienkiewicz, Witold; Olczak, Elżbieta (2006). Ilustrowany atlas historii Polski. Tom 1: Od źródeł po schyłek Jagiellonów. Warsaw: Demart.

Russia
• Dubov, I.V. (1985). Cities, majesty shining. Leningrad University Press.
• Alexandrov, B.A.; Tishkov, B.A. (1999). The beginning of Russian history. Institute of Ethnology and Anthropology of the Russian Academy of Sciences: Moscow:
• https://faculty.unlv.edu/pwerth/Russia-Expansion-Asia.jpg
• The Atlas of the Land of the Irregular Forces of 1858: Map of the lands of the Don Cossack Host
• http://www.conflicts.rem33.com/images/abkhazia/KARTEN/kaukaz%2019%20rus.JPG

Georgia
• Fähnrich, Heinz (1993). Putzger historischer Weltatlas: Geschichte Georgiens von den Anfängen bis zur Mongolenherrschaft. Aachen: Shaker.

Levant
• Misr, F. (1949). Farouk of Egypt. Royal Artillery Magazine, November 1949.

Mesopotamia & Iran
• Roaf, M. (1990). Cultural Atlas of Mesopotamia and the ancient Near East, Oxford: Facts on File.
• Busse, Heribert (1975). The Cambridge History of Iran. Cambridge University Press.

Oman
• Oxford Business Group (2007). The Report: Oman 2007. Oxford Business Group.

Yemen

• Fattovich, Rodolfo (2003). "Akkälä Guzay." Encyclopaedia Aethiopica: A-C. Wiesbaden: Otto Harrassowitz KG.

• Alessandro de Maigret (2002). Arabia Felix. London: Stacey International.

Egypt

• Petry, Carl F. (1998). The Cambridge History of Egypt. Cambridge University Press.

Maghreb

• Luscombe, David (2004). The New Cambridge Medieval History. Cambridge University Press.

Morocco

• Lewicki, T. (1988). The Role of the Sahara and Saharians in relationships between north and south: General History of Africa, Africa from the Seventh to the Eleventh Century. UNESCO, p.276-313.

• Duby, G. (1997). L'Atlas Historique Mondial. Larousse.

Sudan

• Shinnie, P.L. (1978). The Cambridge History of Africa. Cambridge University Press.

• O'Fahey, R.S. (2008). The Darfur Sultanate: A History. Londra.

Ethiopia

• Delessa, Tadesse; Alemayehu, Girma. Ethiopian History - From early axumite period to the downfall of Emperor Haile Selassie I. Aster Nega Publishing (2005, revised edition).

Somalia

• Issa-Salwe, Abdisalam M. (1996). The Collapse of the Somali State: The Impact of the Colonial Legacy. London: Haan Associates.

Madagascar

• Ajayi, J. F. Ade (1998). Africa in the Nineteenth Century Until the 1880s. University of California Press.

Angola/Congo

• Lattre, Jean. Atlas Moderne ou Collection de Cartes sur Toutes les Parties du Globe Terrestre. (1775), plate B-30.

• Thornton, John (1998). Warfare in Atlantic Africa. University College of London Press, pp.103,121-122.

Chad

• Collelo, Thomas (1990). Chad: a country study, Washington, D.C.: Library of Congress, pp.5-9.

W. Africa

• Curtin, Philip; Feierman, Steven; Thompson, Leonard; Vansina, Jan. (1995). African History: From Earliest Times to Independence. Pearson, 2nd ed.
• Smith, Robert S. (1989). Warfare & Diplomacy in Pre-Colonial West Africa Second Edition. Madison: University of Wisconsin Press.
• Levtzion, Nehemia (1973). Ancient Ghana and Mali. New York: Methuen & Co Ltd. pp.44,183.

Turkmenistan

• Bosworth, C.E. (1975). The Tāhirids and Saffārids. In "The Cambridge History of Iran, Volume 4." Cambridge University Press, pp.118-120.
• Biran, Michal (2005). The Empire of the Qara Khitai in Eurasian History: Between China and the Islamic World. Cambridge University Press, pp.41-44.

Ferghana

• Tarn, W.W. (1966). The Greeks in Bactria and India. Cambridge University Press.
• Records of the Grand Historian of China. Translated from the Shih Chi of Sima Qian by Burton Watson. Columbia University Press (1961), Volume II.
• Grenet, Frantz (2002). Regional Interaction in Central Asia and North-West India in the Kidarite and Hephtalite Periods. In "Indo-Iranian languages and peoples." Proceedings of the British Academy, 116, pp.220-221.
• Bai, Shouyi; et al. (2003). A History of Chinese Muslims. Beijing: Zhonghua Book Company, p.211.
• Soucek, Svat (2000). A History of Inner Asia. Cambridge University Press, p.136.
• Bosworth, C.E. (1996). The new Islamic dynasties: a chronological and genealogical manual. Columbia University Press.

Afghanistan

• Encyclopædia Britannica- "Shahi Family." Encyclopedia Britannica Inc.

Kashmir

• Chadurah, Haidar Malik (1991). History of Kashmir. Bhavna Prakashan.

Pakistan

• Kenoyer, Jonathan Mark (1998). Ancient Cities of the Indus Valley Civilization. Islamabad: American Institute of Pakistan Studies, p. 65.
• McIntosh, Jane (2007). The Ancient Indus Valley: New Perspectives. ABC-CLIO, pp.xi-xiv.
• Chandra, Satish (2006). Medieval India: From Sultanate to the Mughals-Delhi Sultanate (1206–1526) Part 1. Har-Anand Publication Pvt Ltd, p.18.

India

• Singh, M.R. (1972). Geographical data in the early Puranas: A critical study. Punthi Pustak.
• Schwartzberg, Joseph E. (1992). A Historical Atlas of South Asia, second impression. Oxford University Press.
• http://huntingtonarchive.org/resources/historicalMaps.php
• Quintanilla, Sonya Rhie (2007). History of Early Stone Sculpture at Mathura: Ca. 150 BCE - 100 CE. Leiden, The Netherlands: Koninklijke Brill NV, pp.9-10,14,120.
• Allan, John (1936). A Catalogue of the Indian Coins in the British Museum (Ancient India). London, p.39.22.
• Archaeological Survey of India (1933). Hathigumpha Inscription of Kharavela of Kalinga. Epigraphia Indica, Vol. XX, 86-89.
• Sims-Williams, Nicholas (2012). Bactrian Historical Inscriptions of the Kushan Period. The Silk Road, Vol. 10, pp.77-78.
• Nilakanta Shastri, K.A. (1976). A History of South India: From Prehistoric Times to the Fall of Vijayanagar. OUP India, p.145,168,171-174,177,183-186,191,206-208.
• Kamath, Suryanath U. (2001). A concise history of Karnataka: from pre-historic times to the present. Bangalore: Jupiter books, pp.57,65,75-76.
• Singh, Pritam (2008). Federalism, Nationalism and Development: India and the Punjab Economy. Routledge, pp.25-26.

Nepal

• Schwartzberg, Joseph E. (1978) A Historical Atlas of South Asia: "Gurkha Expansion and the Anglo-Nepali War, 1760-1816." The University Press of Chicago.

China

• http://www.globalsecurity.org/military/world/china/images/map-shang-no-date-2.jpg

• http://www.friesian.com/sangoku.htm#china

• Grousset, Rene (1970). The Empire of the Steppes. Rutgers University Press, pp.27-28,53-54.

• Taylor, Keith Weller (1983). The Birth of Vietnam. University of California Press, pp.23-24.

• Sims-Williams, Nicholas (2012). Bactrian Historical Inscriptions of the Kushan Period. The Silk Road, Vol. 10, pp.77-78.

• http://img.xixik.net/custom/topic/xixik_7a995145fffabca0.png

• Dotson, Brandon (2009). The Old Tibetan Annals: An Annotated Translation of Tibet's First History. VÖAW, Austria.

• Blackmore, M. (1960). The Rise of Nan-Chao in Yunnan. Journal of Southeast Asian History. 1(2): 47-61.

• Zizhi Tongjian

• Mote, Frederick W. (1999). Imperial China: 900–1800. Cambridge: Harvard University Press, pp.47-49.

• Dreyer, Edward L. (1988). Military origins of Ming China. Cambridge University Press.

• Graff, David Andrew; Higham, Robin (2012). A Military History of China. University Press of Kentucky, p.119.

• http://oook.info/anth230/rebelmap.jpg

• http://www.usma.edu/history/SiteAssets/SitePages/Chinese%20Civil%20War/ChineseCivilWar03.gif

• Forbes, Andrew D.W. (1986). Warlords and Muslims in Chinese Central Asia: A political history of Republican Sinkiang 1911-1949. CUP Archive, p.151.

Mongolia

• N. Ishjamts (1994). History of civilizations of Central Asia, Volume II. The development of sedentary and nomadic civilizations: 700 B.C. to A.D. 250. Paris: UNESCO Publishing, pp. 155-156.

• Drompp, Michael R. (2005). Tang China And The Collapse Of The Uighur Empire: A Documentary History. Leiden, The Netherlands: Koninklijke Brill NV, pp.36,38,114.

• Soucek, Svat (2000). A History of Inner Asia. Cambridge University Press, pp.46,77,81-82,105,107,167.

Korea

• Barnes, Gina Lee (2001). State Formation in Korea: Historical and Archaeological Perspectives. Psychology Press, p.10,18,22,24.

• Korea National University of Education, (2004). Atlas Korean History. Four Seasons,

pp.24,26.
• Lee, K.; Yi, K.; Wagner, Edward W.; Schulz, Edward J. (1984). A New History of Korea. Harvard University Asia Center.

Japan

• Turnbull, Stephen (1998). The Samurai Sourcebook. Arms and Armour Press, pp.208-217,221-224,228,230
• Kerr, George H. (2000). Okinawa: The History of an Island People. Tuttle Publishing, p.50,60-61,85.
• http://www.samurai-archives.com/maphome.html
• http://www.maproom.org/00/05/0005.jpg
• Howell, David Luke (1995). Capitalism from Within: Economy, Society, and the State in a Japanese Fishery. University of California Press, pp. xv,33.

Vietnam

• Taylor, Keith Weller (1983). The Birth of Vietnam. University of California Press, pp.17,20,38,41,132,135.
• Whitmore, John K. (1968). The Development of Le Government in 15th Century Vietnam. Cornell University.
• Lafont, Pierre-Bernard (1989). Les Frontieres du Vietnam (Travaux du Centre d'histoire et civilisations de la peninsule indochinoise). L'Harmattan.
• Maspero, G. (2002). The Champa Kingdom. Bangkok: White Lotus Co., Ltd.
• West, Barbara A. (2010). Encyclopedia of the Peoples of Asia and Oceania. Infobase Publishing, p.158.

Cambodia

• Higham, Charles (2001). The Civilization of Angkor. London: Weidenfeld & Nicolson, p.xiii.
• Coedès, George (1968). The Indianized States of Southeast Asia. University of Hawaii Press, pp.136-137,170.

Laos

• Wyatt, David K. (2004). Thailand: A Short History. Chiang Mai: Silkworm Books, 2004, pp.51,75.

Thailand

• Saraya, Dhida (1985). Lavo - A Coastal State of the 11th Century. Muang Boran, vol.11-

2, pp.35-51.

• Coedès, George (1968). The Indianized States of Southeast Asia. University of Hawaii Press, p.195-196,208-209,222

• Jumsai, Manich (1976). Popular History of Thailand: “King Tilokarat (1441-1485)”. pp.54-57.

Burma

• Beckwith, Christopher I. (2002). Medieval Tibeto-Burman languages. Brill, pp.159-161.

• Harvey, G.E. (1925). History of Burma. Laurier Books Ltd.

• Lieberman, Victor B. (2003). Strange Parallels: Southeast Asia in Global Context, c. 800–1830, volume 1, Integration on the Mainland. Cambridge University Press, pp.26,90-91.

• Sandamala Linkara, Ashin (1931). Rakhine Yazawinthit Kyan. Yangon: Tetlan Sarpay.

• Baruah, S L (1993). The Last Days of Ahom Monarchy. Delhi: Munshiram Manoharlal Publishers, p.225.

• Schwartzberg, Joseph E. (1992). A Historical Atlas of South Asia, second impression. Oxford University Press.

Malaysia

• Karim, Safiah (2011). Tatabahasa Dewan Edisi Baharu, page 7. Dewan Bahasa dan Pustaka.

• O’Reilly, Dougald J. W. (2007). Early Civilizations of Southeast Asia. Rowman Altamira. p.54.

• http://www.mykedah2.com/raja_kita/eng/040000.htm

• Syukri, Ibrahim (1985). History of the Malay Kingdom of Patani. Ohio University, Center for International Studies.

• Coedès, George (1968). The Indianized States of Southeast Asia. University of Hawaii Press, p.198,234.

• http://www.royalark.net/Malaysia/perak.htm

• https://web.archive.org/web/20140617035533/http://www.bt.com.bn/2013/03/07/sabah-and-sulu-claims

Brunei

• Sidhu, Jatswan S. (2009). Historical Dictionary of Brunei Darussalam (second ed.). Lanham, Maryland: Scarecrow Press, p.37.

Philippines

• Malkiel-Jirmounsky, Myron (1939). The Study of The Artistic Antiquities of Dutch India. Harvard Journal of Asiatic Studies, 4(1):59–68.
• Adib Majul, Cesar (1999). Muslims in the Philippines. University of the Philippines Press.
• Careri, Giovanni (1728). Giro del mondo: Contenente le cose più ragguardevoli vedute nell'isole Filippine. Coleti, pp.137-138.
• https://web.archive.org/web/20080731121053/http://palawan.ws/index.php/About-Palawan/History.html

Indonesia
• J.Ph.Vogel(1915). The Earliest Sanskrit Inscriptions of Java. ROD:28-35, Plate 27.
• Munoz, Paul Michel (2006). Early Kingdoms of the Indonesian Archipelago and the Malay Peninsula, Singapore, pp.203-204,212.
• Latif, Chalid; Lay, Irwin (1992). Atlas sejarah: Indonesia dan dunia. Pembina Peraga.
• Hobart, Angela; Leemann, Albert; Ramseyer, Urs (2001). The People of Bali. Wiley, p.24.
• Coedès, George (1968). The Indianized States of Southeast Asia. University of Hawaii Press, p.144-147.
• Ricklefs, M.C. (1991). A History of Modern Indonesia since c.1300, 2nd Edition. Stanford University Press, p.15.
• http://www.zum.de/whkmla/region/seasia/xindonesia.html

Australia
• Arrowsmith, John (1842). The South Eastern Portion of Australia; compiled from the Colonial Surveys and from details furnished by Exploratory Expeditions. London.

Various
• Wikipedia: misc. settlement founding dates
• Geohack: coordinates of sites and locations
• http://www.lena-dvorkina.narod.ru/