The Origin of the State: Incentive Compatible Extraction under Environmental Circumscription

David Schönholzer*

April 10, 2017

Abstract

Theories of early state formation posit that the state enabled an emerging elite to extract resources in exchange for protection from outside groups. This paper formalizes and empirically evaluates these forces in a unified framework. The model shows that extraction is closely linked to the idea of environmental circumscription: only if outside options are sufficiently poor for potential extractees, (a) they are willing to accept extraction, and (b) extractors consider extraction capacity a worthwhile investment. In a global dataset of archaeological sites on a grid with 184,523 cells, we then show that circumscription is strongly associated with the location of early states, using the intersection of large rivers through arid regions as an instrument. Our estimates suggest that extraction was more important in the Old World civilizations of Egypt and China while other motives such as protection were more important in the New World.

*Department of Economics, UC Berkeley, `david.s@econ.berkeley.edu`. I thank Ernesto Dal Bó, Brad DeLong, Frederico Finan, Edward Miguel, Gerard Roland, Raúl Sánchez de la Sierra, Joachim Voth and Noam Yuchtman for useful discussions and continuous encouragement as well as seminar participants in the BEHL Economic History Lunch at Berkeley, the Stanford Economic History conference, the Archaeology Brownbag Seminar at Berkeley and the Society for American Archaeology 81st Conference session "Theoretical Approaches in Archaeology". I gratefully acknowledge financial support from the Swiss National Science Foundation, Fellowship number P1SKP1-158639. Word count: 10,044.

1 Introduction

Modern human social organization underwent a major transition from stateless societies to states in the course of the last 6,000 years. This transition marks the origin of taxation, law, organized warfare, bureaucratic rule, property rights, the division of labor, and other novel social phenomena, paving the way for modern economic development (Finer, 1997; Bockstette et al., 2002).

Theories of early state formation are as old as the social sciences themselves, with classic writers such as Marx and Smith speculating about the origin of the state. Research in history, sociology and political science in the last forty years has identified extraction by emerging elites in exchange for protection from predation as a key dynamic underlying the process of state formation (Mann, 1986; Tilly, 1990; Olson, 1993). A burgeoning literature in economics formalizes and empirically assesses these central forces and finds supporting evidence for their role in the creation and development of the state (Dal Bó et al., 2016; Mayshar et al., 2016; Sanchez de la Sierra, 2017).

Despite these advances in our understanding of state formation, two major challenges remain: first, while sedentism was widespread among stateless societies, they were typically mobile and frequently relocated in the face of threats or in search for better land. This is evidenced by migratory episodes of groups throughout the Holocene in Africa (Berniell-Lee et al., 2009), Europe (Hofmanová et al., 2016), Asia (Larson et al., 2010) and the Americas (Benson et al., 2007). Mobility limits an emerging elite's ability to extract resources from groups by imposing an incentive compatibility constraint: life under the yoke of the state can be no worse than evasion through migration. This constraint poses a fundamental theoretical challenge to all models of extraction as drivers of state formation.

Second, it remains unclear whether modern state formation theory can empirically account for the geographic pattern of early state formation. Archaeological evidence of the earliest instances of state formation are concentrated in seven places around the world: Mesoamerica, the Andes, West Africa, Egypt, Mesopotamia, the Indus Valley and China (Trigger, 2003). Whether the geographic environment of these early state nurseries is consistent with predictions from state formation models is unknown. This is true despite recent empirical advances because, while the object of interest is early states that arose thousands of years ago, almost all empirical evidence on early state formation in economics rests on data primarily from the 19$^{\text{th}}$ century, at a time when early state development has been severely polluted or completely disrupted by Europeans.[1] Thus, the potential of archaeological data to advance our understanding of the role of extraction in the creation of the early state is immense.

This paper seeks to address both of these challenges in a unified framework. To this end, we first develop a model of extractive state formation with the possibility of evasion through migration. The model identifies conditions of a region and its surrounding land for which both the extractees accept domination by a state and the extractors consider state formation a worthwhile investment. The key intuition of the model invokes an old idea developed in cultural anthropology (Carneiro, 1970):

[1]For example, Fenske (2013, p. 1366) observes that only eight out of 1,267 societies in Murdock's (1967) *Ethnographic Atlas* are observed before 1500.

the larger the difference between the quality of the land in a region and its surroundings – that is, the more environmentally circumscribed it is – the more likely is the formation of the state. Since evasion through migration is more costly in circumscribed areas, potential extractors could more easily "cage" (Mann, 1986, p. 38-40) potential extractees. While the argument is essentially cross-sectional, the idea is that the high returns to agriculture and sedentism that emerged in the course of the Holocene opened the door to incentive compatible extraction in some regions sooner than others; the development of pristine states in other regions was cut short by the expansion of early states or their descendants.

The structural parameters of the model are then estimated via probit and logit in a global dataset of archaeological excavation sites related to early state formation and global land quality data from the FAO on a grid with cell size ¼ degrees (about 28km at the equator). We find supporting evidence for the role of circumscription in predicting the location of early state sites: a one standard deviation increase in circumscription is associated with a 45% increase in the probability of a state site. This effect is as large or even larger than the corresponding effect of land quality itself.

To study the robustness of the effect, we introduce various sets of control variables related to alternative theories of state formation on the role of ecological diversity, regional climate, and other environmental features, as well as $5^\circ \times 5^\circ$ virtual country fixed effects (about 556km squared at the equator). The effect of circumscription becomes larger and more precise with the introduction of more control variables, speaking for important complementarities between circumscription and alternative theories of the early state. Conceptually, this is intuitive: theories of extractive state formation focus on various regional aspects conducive to state formation, many of which are largely orthogonal to the extent of circumscription. Controlling for outside options in the form of circumscription adds precision to alternative explanations of the rise of the state.

To deal with potential endogeneity of circumscription in the specifications so far, we use the interaction of large rivers and arid regions in an instrumental variable strategy. Desert rivers create large differences in land quality between the river valley and the barren land in its vicinity, which turns out to be an important driver of circumscription: the first stage using the interaction of river flow accumulation and the deserts is statistically powerful and robust to a large set of control variables, including rivers and deserts themselves. IV estimates are substantially larger than OLS and precisely estimated, suggesting the presence of attenuation bias in baseline estimates.

We next turn to the heterogeneity of circumscription across civilizations involved in early state formation. Confirming qualitative descriptions of circumscription, we find that the Andean, Egyptian, Mesopotamian and Aztec regions stand out as heavily circumscribed, while others, such as the Maya, are much less so. However, contrary to these qualitative assessments, we find that archaeological sites associated with early states in China are also tightly circumscribed, offering a new avenue for inquiry in early Shang and Zhou state formation.

Finally, the paper analyzes how the strategic environment across different civilizations may contribute to the heterogeneity across civilizations and thereby highlight regions in which extraction was a particularly important motive. To this end, we compare the extent of circumscription across civilizations to the amount of rugged terrain in the vicinity, offering a strategic advantage to challengers. There is substantial variation in the extent to which early state sites of different civilizations are circumscribed and surrounded by rugged terrain, giving rise to the possibility that extraction may have been much more important in some regions (such as Egypt and China) while other motives such as protection drove state formation in others (such as the Aztec and Maya).

This paper connects to three literatures in economics and the social sciences more broadly. First, it speaks to the literature on early state formation by providing, for the first time, quantitative evidence for the role of circumscription in early state formation. Thereby, it offers the first piece of evidence that modern theories of extractive state formation find empirical support in archaeological data, particularly when considering the outside options of potential state subjects. This finding supports earlier qualitative research by Carneiro (1970) and Allen (1997) on the role of circumscription in state formation, and more broadly on how opportunities to evade state power diminish state formation and integration (Mann, 1986; Scott, 2009). The role of mobility in state formation also features prominently in Olson (1993). In Olson's terms, this paper cautions that for bandits to become stationary, victims need to be sufficiently immobile too.

Within the economics debate on early state formation, this paper sheds light on the role of the surrounding geographic environment for state formation in a given region. In this sense, it complements various other economic theories of state formation and institutional development. Geography has already been found to be an important factor in early state development, such as Fenske's (2014) research on states providing security for trade across ecologically diverse zones in the African context, as originally suggested by Bates (1987). More broadly, the findings here provide an important mechanism through which geography shapes institutions: extractive institutions are more lucrative in highly productive, constrained places, which may explain why the places that experienced a reversal of fortune after the arrival of Europeans (Acemoglu et al., 2002) acquired these fortunes in the first place, and why they already had extractive institutions in place for Europeans to take over (Dell, 2010).

Extraction plays a key role in economic theories of state formation. In the work by Mayshar et al. (2016), the ease with which agricultural surpluses can be extracted by elites is the driving factor of state formation. This paper comes to a similar conclusion but instead of focusing on the type of crops that are grown, it argues that the extent to which groups could escape into similarly productive land generates variation in early state formation. In this sense, the work presented here is complementary to their findings. Sanchez de la Sierra (2017) finds empirical evidence for the role of both extraction and protection in state formation using Eastern Congo as a quasi pre-state environment. He argues compellingly for the difficulty of using archaeological data from early states directly to study state formation quantitatively due to the absence of systematic disaggregated data. This paper argues that, despite these difficulties, direct archaeological evidence can be embedded in topographic, climatic and other environmental data to produce a coherent picture of early state formation. There is no perfect substitute for the pristine pre-state environment of the mid-Holocene to study the rise of the state due to the looming presence of international markets in raw materials and advanced weaponry in modern stateless environments.

While the focus in economics and related fields is typically on extraction, the study of protection and surplus generation as motives for state formation has a long tradition in archaeology (Childe, 1936; Service, 1975; Haas, 1982; Johnson and Earle, 2000). Tilly (1990) also famously argued for the role of military capabilities in the development of the state. This paper finds suggestive evidence for the enhanced role of circumscription in regions with more rugged terrain. Dal Bó et al. (2016) deal with the problem of producing and defending surplus in a complementary model to the one presented here, providing an explanation of how states can invest into productive assets if their strategic environment is sufficiently secure.

This paper is also connected to the literature on state capacity in economics, initiated by Besley and Persson (2009, 2010). Similar to their model, agents choose to invest into state capacity, should the environmental and strategic environment afford it. Incentives to invest in state capacity may also depend on the extent of external threats, as in Gennaioli and Voth (2015).

Historical state capacity and institutions more broadly have been found to be important predictors of long-run development (Bockstette et al., 2002; Acemoglu et al., 2005; Michalopoulos and Papaioannou, 2013). The work presented here provides an explanation of how prehistoric institutions arose. The starting point for these institutions may be found in the spatial configuration of land quality and opportunities for challengers.

The rest of this paper is organized as follows. Section 2 develops a simple game of extractive state formation based on circumscription (Carneiro, 1970). Section 3 describes the environmental and archaeological data and how circumscription is computed. Next, section 4 introduces the baseline estimation strategy derived from the model prediction, presents estimation results and robustness to various alternative theories of state formation. Then, section 5 lays out the instrumental variable strategy to deal with the potential endogeneity of circumscription. Section 6 studies the heterogeneity of circumscription across civilizations and presents suggestive evidence on the relative importance of extraction and other motives such as protection. Finally, section 7 concludes.

2 Model

This section describes a dynamic game of extractive state formation under varying outside options. It is heavily indebted to Carneiro (1970), whose qualitative theory of state formation finds its formal structure here. The purpose of the model is to outline conditions under which both the loser of a conflict is willing to accept political subordination as well as the winner is willing to invest into state capacity, the latter component mirroring ideas first developed in Besley and Persson (2009).

The game is illustrated in figure I. An incumbent group engaged in mixed subsistence farming inhabits a region with agricultural land quality a. The region is surrounded by land of quality $\tilde{a}$. Groups derive a general rate β from any land they work, no matter whether it is their home region or the surrounding region. They additionally derive a home rate γ from their home land, reflecting either a premium:

for instance, for their knowledge of the peculiarities of the land, or for fixed infrastructure they put in place to make the land more amenable to farming; or a penalty: for instance, for overworking the land and exhausting its productivity, or for facing off threats particular to the home region.

[Figure 1 about here.]

Separating the general rate β from the home rate γ allows us to study how state formation depends on the land quality calculus of a region versus its surrounding land: if β is zero, extraction capacity only depends on the quality of land in a given region; if γ is zero, it depends on the relative quality in the home region versus the surrounding region; and in the extreme, if γ is negative, it may depend only on the quality of the surrounding land.

As populations grow in the course of the favorable climate of the mid-Holocene, a challenger (internal or external) arrives to compete over arable land, acquiring the skills to benefit from the home premium in the process. When the two groups clash to determine who gets to farm the best land, nature determines the winner and the loser of the conflict. The victory in conflict elevates the winning group to the status of an emerging elite with the potential for extraction.

The winner then chooses between displacing the losing group or attempting to dominate the loser, exacting a tribute of chosen magnitude $\tau \geq 0$ and paying a fixed cost c plus a stochastic component ε to build and maintain an infrastructure to monitor and enforce the payment of the tribute. The stochastic component is mean zero and distributed according to the distribution function $F(\cdot)$. It is known to the winner when choosing whether to displace or dominate; it captures idiosyncratic aspects of domination left outside the model and will be discussed further in the next section. The infrastructure for tribute extraction – consisting of capacities such as keeping track of tribute paid through a specialized bureaucracy or maintaining a monopoly of violence over the losing group – constitute the origins of the extractive state. This is similar to Besley and Persson (2009, 2010) whose interest is the state's taxation capacity.

If the winning group attempts to dominate the losing group, the latter can either accept domination, staying in the winner's home region and sharing in the home premium γ with the winner, but paying the tribute τ to the winner. Alternatively, the group can flee from the area to the surrounding area with quality $\tilde{a}$, which amounts to the same outcome as being displaced.

We solve for the subgame-perfect equilibrium of the game by backward induction. For a given $\tau \geq 0$, the loser will accept domination if and only if $(\beta + \gamma)a - \tau \geq \beta\tilde{a}$. Otherwise, it will flee to the surrounding region. This is an incentive compatibility constraint: since there is no way to physically constrain the losing group, no more than the surplus from the home region relative to the surrounding region can be extracted for the losing group to stay. If this incentive compatibility condition holds, it determines the optimal tribute chosen by the winner in case of domination by making the loser indifferent between domination and flight: $\tau^* = (a - \tilde{a})\beta + a\gamma$. Given this optimal tribute τ^*, the winner will attempt domination if and only if the return from tribute extraction is greater than the investment cost into state capacity: $(\beta + \gamma)a + \tau^* - (c + \varepsilon) \geq (\beta + \gamma)a$, which yields a simple participation condition for the winner to form the extractive state:

$$(a - \tilde{a})\beta + a\gamma \geq c + \varepsilon \tag{1}$$

We have thus proven the following proposition:

Proposition 1 (Extraction under Circumscription)**.** Assume the setup of the game along the lines of the game in figure I, and let Y be a binary random variable evaluating at unity if the extractive state is formed. There are three subgame-perfect equilibria:

1. {Dominate with $\tau^* = (a - \tilde{a})\beta + a\gamma$, Accept} if $c + \varepsilon \leq (a - \tilde{a})\beta + a\gamma$,

2. {Displace, Accept} if $c + \varepsilon > (a - \tilde{a})\beta + a\gamma \geq 0$,

3. {Displace, Flee} if $(a - \tilde{a})\beta + a\gamma < 0$.

The probability of state formation is $\Pr(Y = 1|a, \tilde{a}) = F((a - \tilde{a})\beta + a\gamma - c)$. The size of the extractive state τ^* and the probability of state formation increase with higher land quality gradient between home region and surrounding region $a - \tilde{a}$ and higher land quality a in the home region; and the state formation probability falls with higher investment costs of extractive state capacity c.

The core intuition of the model is that a high gradient in agricultural land quality (i.e. the region is "circumscribed" by lower quality land) means that the losing group faces a relatively worse outside option from fleeing. It will thus be more willing to accept domination through a state if it resides in a place with high quality land if it is surrounded by low quality land. This captures the essence of Carneiro's (1970) circumscription theory.[2]

While the basic intuition of the model is fairly straightforward, the model clarifies the relationship between environmental circumscription – which is a function of land quality – and land quality itself: controlling for land quality, circumscription captures the extent to which evasion through migration matters for early state formation. In addition to the gradient in land quality affecting the probability of state creation, land quality also has a direct effect through the size of the home rate γa. More states should form if groups are more strongly attached to their land due to specific knowledge or infrastructure. On the other hand, if regions with high land quality face frequent challenges from outsiders (meaning that γ is negative), less states may form there.

While state formation is modeled as a binary outcome, the underlying latent magnitude of extraction τ^* quantifies the extent to which extraction is a motive for state formation. In section 6, we discuss the intensity of extraction across civilizations in more detail.

The focus on competition over arable land of semi-nomadic groups engaged in some agricultural production in the model deserves further discussion. The groups in the model conduct some small-scale agriculture, but they may still be engaged in hunting and gathering, either as part of their nutritional intake throughout the year

[2] Interestingly, the model suggests that for some parameter values (the second equilibrium in the list), the loser would be better off if the state were formed; however, the surplus produced in the home region is not large enough to pay the fixed cost of state formation, and so the winner prefers to displace the loser. It seems this possibility has been overlooked so far in the literature on circumscription.

or for part of the season. The transition to full-scale agriculture took several thousand years in some regions (Arranz-Otaegui et al., 2016), during which groups remained semi-nomadic; they often brought their domesticated crops with them across large migratory distances (Erickson et al., 2005; Crowther et al., 2016). After the initial domestication period, agriculture spread and became available almost everywhere on the globe (Diamond and Bellwood, 2003), often displacing pure hunter-gatherer groups (Hofmanová et al., 2016).

3 Data

To empirically assess the importance of environmental circumscription for early state formation, we combine data on agricultural land quality, archaeological data on excavation sites associated with early state formation, as well as a number of other data describing the climatic and topographic environment. We describe each of these in turn.

As a proxy for agricultural land quality, we use data on the maximal potential production capacity in t/ha over seventeen crops from the FAO's Global Agroecological Zones (GAEZ) database (Fischer et al., 2008), scaled by historical calories per ton for each crop by the FAO (Chatfield, 1953) and corrected for the Columbian Exchange (Nunn and Qian, 2010). This proxy is very similar to the Caloric Suitability Index developed by Galor and Özak (2016), with one important difference: since we are interested in identifying steep gradients in land quality, we want to allow for alluvial agriculture based on rivers in addition to rainfed agriculture. The GAEZ database and covers total crop production value for all important crops across the globe on a 5-minute resolution grid, totaling 9,331,200 cells. Figure II shows a map of the resulting agricultural land quality data; details about data construction are described in the appendix.

[Figure 2 about here.]

To compute the intensity of circumscription between a location and its surrounding land, we compute the difference between the the quality in a given cell and the average quality of all surrounding cells at various radii (see section 4). In contrast to land quality per se, circumscription intensity may be relatively low in large fertile expanses such as the American Midwest or Western Russia since most cells are surrounded by equally productive cells; on the other hand, circumscription intensity may be high in areas that are agriculturally productive but confined in otherwise barren land or surrounded by oceans, such as the river valleys of the Nile or the Indus, or the islands of Cuba or Sicily. More details about the construction of circumscription intensity are discussed in the appendix.

The data on the location of early states comes from the archaeological record presented in Bogucki (1999) and shown in figure III. States are defined as "powerful, complex, institutionalized hierarchies of public decision-making and control" based on (Brumfiel, 1994). In his book, several regional maps show key sites involved in the creation of the first states across Africa, Asia and the Americas. This selection of early state sites captures the broad consensus of archaeological sites in what are traditionally considered the geographic centers of early state creation: Mesopotamia, Mesoamerica, the Andes, Egypt, the Indus valley, and China; this collection thus includes both groups of city states and territorial states (Trigger, 2003). He also includes sites in West Africa (Yoruba) and Great Zimbabwe. The civilizations in these regions are frequently referred to as single territorial states, but archaeological and historical evidence points towards systems of city states or small territorial states, at least at the beginning of their development (Finer, 1997; Trigger, 2003). While this selection of early state sites is by no means exhaustive and some of these archaeological sites play a larger role in each regions formation of early states than others, they are the most prominent sites to provide key evidence of public architecture, kingship, or urban elites typically associated with early states, and as such each of these sites locations are indicative of the environment in which early states formed.

[Figure 3 about here.]

In addition to agricultural land quality, a number of other climatic, topographic and environmental factors may be more or less conducive to state formation. Included are annual mean and standard deviation of temperature from the from the WorldClim global climate database (Hijmans et al., 2005); river flow freshwater accumulation from the HydroSHEDS database (Lehner et al., 2008); potential vegetation from SAGE (Ramankutty and Foley, 1999); and topographic slope and ruggedness (Riley et al., 1999; Burchfield et al., 2006; Nunn and Puga, 2012). We address in section 4 how each of these relate to alternative theories of early state formation.

[Table 1 about here.]

Table I summarizes descriptive statistics of all the data. First, it can be seen that given the rarity of early state formation, only very few cells (0.03%, a total of 60 cells) have a state site in them. Noteworthy is also that the median cell is negatively circumscribed, meaning it is surrounded by land on average of slightly better quality than itself; but mean circumscription is positive with a magnitude of about 9% of land quality. This skew is due to the fact that most cells have at least some cell in their vicinity with some positive land quality, although the majority of cells has land quality zero. Naturally, the correlation between land quality and circumscription is high at 0.74.

4 Baseline Estimation

In this section, we take proposition 1 to the data. To this end, we divide the entire landmass of the globe into cells with size ¼ degrees (28km at the equator) and index them by $i = 1, ..., N$, which results in 194,102 cells. The outcome variable is an indicator Y_i taking on unity if cell i includes a site from Bogucki's sample of early state locations. From the FAO data, we know agricultural land quality a_i and circumscription $a_i - \tilde{a}_{ir}$, where

$$\tilde{a}_{ir} = \frac{1}{N_r} \left[\sum_{j \neq i : \text{dist}(i,j) \leq r} a_j \right]$$

for radius r in cell units. N_r is the number of cells (minus the cell at i) in the circle with radius r, as explained in the data section. When we write $\tilde{a}_i$ without a radius subscript, we are using our default radius at $r = 10$, resulting in a radius of about 280km at cell size ¼.

If we assume that the stochastic component ε is distributed as $\varepsilon \sim \mathcal{N}(0, 1)$, i.e. standard normal, or that $\varepsilon \sim \text{logistic}(0,1)$, then we can directly estimate the structural parameters (β, γ, c) from the model. That is, we run

$$\Pr(Y_i = 1 | a_i, \tilde{a}_i) = F(-c + (a_i - \tilde{a}_i)\beta + a_i\gamma) \tag{2}$$

with $F(\cdot) = \Phi(\cdot)$ or $F(\cdot) = \Lambda(\cdot) \equiv \exp(\cdot)/(1 + \exp(\cdot))$. This is the key empirical prediction from proposition 1. Results are shown in table II. We run three variations of the model with both probit and logit: in column (1) and (4), we force the home rate γ to be zero, studying the role of circumscription in isolation; in column (2) and (5), we force the general rate β to be zero, focusing only on the land quality of the home cell; in column (3) and (6) we estimate both of them simultaneously.

[Table 2 about here.]

In terms of the model structure, the estimates in table II show that both the general rate $\hat{\beta}$ and the home rate $\hat{\gamma}$ are strongly associated with the presence of a state site. The fact that the general rate $\hat{\beta}$ is significantly different from zero provides evidence for the relevance of circumscription; and the significance of $\hat{\gamma}$ confirms the importance of agricultural land quality for the formation of early states, as proposed in much of the early state formation literature. Interestingly, including both of them in columns (3) and (6) suggests that circumscription may actually be more important than land quality; or, in terms of the model, the general yield rate is more important than a home premium. However, the main point here is that, according to these estimates, the difference in the quality of the home land and the land surrounding a particular location may be just as important as the land at the location itself (and in fact, we cannot reject that $\hat{\beta} = \hat{\gamma}$).

It is also useful to recall that since $(a - \tilde{a})\beta + a\gamma = (\beta + \gamma)a - \beta\tilde{a}$, the sum of the coefficients on circumscription and land quality give the gross effect of land quality and $-\beta$ gives the effect of a change in the quality of surrounding land. Interpreting the coefficients in this way speaks for a substantial effect of the quality of surrounding land, almost as large as the effect of the own land.

The magnitudes of the coefficients can also be interpreted in terms of an underlying latent index of state formation. The rarity of state formation is reflected in a very high investment cost estimate $\hat{c}$, relative to the stochastic error distribution. For example, in case of the probit models, the standard deviation of the error is unity, and in the unconstrained model (3) we have that $\hat{c} = 3.51$, setting a very high bar for any cell to feature a state site. The magnitude of $\hat{\beta}$ of around 0.12 suggests that a one standard deviation unit increase in circumscription increases the latent index value by 12% relative to the distribution of the error. Of course, this magnitude will be different in the logit models, given that the standard deviation of the error is assumed to be higher (if ε is logistic$(0, 1)$, then $\mathrm{Var}(\varepsilon) = \pi^2/3$, meaning the standard deviation of the error is about 1.81).

These magnitudes translate into economically relevant changes in the probability of a state site in a given cell. As shown in the model summary section of the table, starting from a base of 0.033%, the marginal effect of circumscription is 0.015%, which represents a 45% increase in the probability of a state. Assuringly, marginal effects are almost the same across probit and logit models.

Having established the basic relationship between state formation, circumscription, and land quality, let us turn to the robustness of this relationship to alternative theories of state formation and early development. To this end, we allow the investment cost into extraction capacity to vary from cell to cell according to $c = -\mathbf{X}_i'\eta$, where $\mathbf{X}_i$ includes a constant and groups of control variables accounting for these alternative theories. First, measures of ecological diversity are introduced, specifically, the standard deviation of land quality of cells within the radius, the ecological diversity index and the ecological polarization index (Fenske, 2014) using potential vegetation. This group of controls addresses the significance of trade across ecological boundaries as proposed by Bates (1987).

Second, we include climatic conditions: the annual mean and standard deviation temperature, and absolute latitude (which is a measure of seasonal daylight fluctuations). This group of controls mainly speaks to the extent to which an early Neolithic transition, driven by a favorable climatic environment, may have given some regions a head start in the formation of a state (Olsson and Hibbs Jr, 2005; Putterman, 2008; Ashraf and Michalopoulos, 2015; Matranga, 2017).

Third, topographic, biogeographic and other environmental variables are included: average slope, ruggedness, freshwater accumulation and fixed effects for potential vegetation. These controls address theories about early development related to the ease of control and extraction by state institutions (Nunn and Puga, 2012; Michalopoulos and Papaioannou, 2014). Finally, including virtual country fixed effects takes into account various regional unobservables that may be important for early state formation. Each virtual country is $5° \times 5°$ large, including up to 400 cells ($5/(1/4)$ squared). Differences in unobserved geographic features, population density, and regional culture may all have contributed to state formation, and virtual country fixed effects helps to account for them.

With these groups of control variables, we can now test the robustness of the results presented in table (II). To this end, we run the following linear probability model (LPM):

$$Y_i = (a_i - \tilde{a}_i)\beta + a_i\gamma + \mathbf{X}_i'\eta + u_i \tag{3}$$

where u_i is the CEF error (i.e. $u_i \equiv Y_i - E[Y_i|a_i, \tilde{a}_i, \mathbf{X}_i]$). While the LPM no longer allows for a structural interpretation of the parameters in the model (Horrace and Oaxaca, 2006), it approximates the CEF and thereby provides reasonable approximations of the marginal effects (Angrist and Pischke, 2008). Similar estimates using logit can be found in the appendix.

[Table 3 about here.]

Table III shows eight increasingly demanding specifications of the association of early state sites with circumscription and land quality. In addition to introducing groups of control variables, these estimates also show three different types of standard errors for each point estimate: robust standard errors in parentheses, as in table II; standard errors clustered on the virtual country level, forming 712 clusters, in brackets; and standard errors taking into account the spatial correlation of the error structure (Conley, 1999) in curly brackets.

The first three models (1)-(3) replicate the results from table II in an LPM framework and find similar results. Again we first see both circumscription and land quality having a significant and positive effect on the probability of a state site. Due to the LPM, marginal effects are about 50% higher in model (3), meaning that controlling for land quality, a one standard deviation increase in circumscription leads to an increase in state site probability of around 70%.

The introduction of ecological diversity controls in model (4) substantially changes these conclusions: first, we see that the coefficient on circumscription more than doubles to 0.052 percentage points, implying a one standard deviation increase in circumscription being associated with a 157% increase in the probability of finding an early state site in the cell. Interestingly, we now see the coefficient on land quality turning significantly negative, with a fairly large magnitude. The implication is that, controlling for circumscription, a one standard deviation increase in land quality actually leads to a decrease in the probability of a state site of around 109%.

Given the negative point estimate of $\hat{\gamma}$, it is useful to again recall that $(a-\tilde{a})\beta + a\gamma = (\beta+\gamma)a - \beta\tilde{a}$. That is, we can also interpret the coefficients in terms of a gross land quality effect and a surrounding-land effect. The estimates in model (4) suggest that a one standard deviation increase in gross land quality $(\hat{\beta}+\hat{\gamma})$ increases the probability of a state site by about 0.016 percentage point, or 48%, and an increase in the quality of the surrounding land by one standard deviation *decreases* the probability by -109%. This interpretation implies the quality of surrounding land is more important than the quality of a land in a given place.

Moving on to models (5)-(8), estimates are largely stable and lend themselves to a similar interpretation as model (4). This suggests that ecological diversity and the trade theory of state formation typically associated with it has merit in the data, but is largely complementary to circumscription. Adding further explanatory variables from other theories hardly moves the needle. Most estimates across all models are significant at a one percent level using all three types of standard errors, except land quality in models (3), (7) and (8) (significant at five and ten percent level, respectively) and circumscription in model (8) (significant at five percent level).

[Figure 4 about here.]

So far, all results presented were using a grid with cell size ¼ and circumscription radius of $r = 10$. Figure IV presents results varying the cell size between ⅕ and ⅓ degrees as well as radii between 4 and 20 cell units. Each point in the graph represents an estimate from a separate regression, with corresponding confidence intervals extending vertically. We use the most demanding specification from table III (model (8), with robust standard errors). Since this is a very demanding specification, we also show results for model (7) from table III in the appendix.

Looking across radii, we see that the point estimates show an inverted-U shape as the radius increases: coefficients are small and insignificant at a small radius, then they rise at intermediate ranges, and finally they drop off slightly. This suggests there may be a distance at which circumscription is most relevant, for a given cell size. At ¼, the "Goldilocks" zone for circumscription seems to be at a radius of around 400km, showing the strongest association with early state formation.

Comparing across cell sizes, we can see that there is a level effect as cell size decreases from ⅓ to ⅕. This is to be expected due to the fact that the mean of the dependent variable falls with with cell size as well (since a smaller share of cells has any state site in them at smaller cell sizes). In fact, looking at the marginal effect of state creation for a one standard deviation unit increase in circumscription on $r = 10$, it is 119% for ⅓, 166% for ¼ and 144% for ⅕. We can also see that the arcs formed across larger radii are shorter for smaller cell sizes. This implies that the relevant magnitude is the number of cell units for a given cell size. In this sense, circumscription seems most relevant within a radius of about twelve cell size units from a given location.

5 Instrumental Variable Estimation

So far, we assumed that the degree of circumscription is exogenous with respect to the process of state formation. This assumption relies on the fact that many features determining the distribution of land quality are immutable: daylight time and topographic features are unaffected by human activity; and things like regional climate, average cloud cover, potential vegetation or freshwater accumulation change only slowly over time. However, there remain two sources of bias that may affect a causal interpretation of the estimation. First, the land quality measure may be subject to measurement error, for example due to assuming domesticated crops on either side of the Atlantic were available in all places on their respective continents. Second, it is possible that early states induced a very particular form of land degradation that may have produced a pattern of circumscription: peripheral erosion and preservation at the center. Although it is unlikely that degradation would be systematically higher outside the core region of a state than inside, this type of reverse causation would lead to biased estimates of circumscription.

To deal with these potential concerns, we use an instrumental variable strategy. Specifically, we instrument circumscription with the interaction of two geographic features that, together, are an important source of steep land quality gradients but are unlikely to directly affect the cost of state formation: large rivers intersecting arid land. We employ this instrument while controlling for the direct effects of rivers and potential vegetation. This means any threat to identification has to rely on an explanation of how the simultaneous presence of rivers and deserts affects state formation other than through circumscription; the identification strategy controls for any explanation involving rivers or deserts in isolation.

[Figure 5 about here.]

Figure V shows a map of the top one percent of river flow accumulation and areas whose potential vegetation has been classified as desert by SAGE. This highlights that particularly four regions would exhibit circumscription driven by the intersection of these two geographic phenomena: Egypt, Mesopotamia, the Indus Valley and Western China.

The instrumental variable strategy is formalized in the following first stage regression equation:

$$a_i - \tilde{a}_i = Z_i\psi + a_i\xi + \mathbf{X}_i'\delta + e_i \tag{4}$$

where $Z_i = \log(\text{RiverAccum}_i) \times \mathbf{1}[\text{PotVeg}_i = \text{Desert}]$ and e_i is assumed to be an uncorrelated error term. Both $\log(\text{RiverAccum}_i)$ and $\mathbf{1}[\text{PotVeg}_i = \text{Desert}]$ are always included in the set of control variables $\mathbf{X}_i$. Similar to table III, we introduce consecutively larger sets of additional control variables controlling for alternative theories of state formation and early development. Equation (3) forms the second stage.

[Table 4 about here.]

Results are presented in table IV. Panel A shows the reduced form and the first stage of the regression. Across all specifications, we see very stable coefficients, except when virtual country fixed effects are introduced. The first stage coefficient on the instrument is highly significant, yielding partial F statistics substantially above common thresholds for weak instruments. In columns (1) and (2), since river flow accumulation is in logs, the coefficient is interpreted as a ten percent increase in river flow accumulation in deserts leads to an increase of circumscription of 0.161 standard deviation units. This coefficient drops by about a factor of three after controlling for virtual country fixed effects: now that we are only comparing within virtual countries, the fact that only a minority of virtual countries have desert rivers sucks up a lot of the variation. However, the coefficient is still highly significant.

In the reduced form models, a one percent increase in desert river flow accumulation has a large (considering the small base) effect of around 0.1% on the probability of a state site in a cell in all specifications. This is around twice as large as the effect of a standard deviation increase in circumscription reported in table III, confirming qualitative assessments about the importance of desert rivers such as the Nile in accelerating the formation of early states (Allen, 1997). Interestingly, the coefficient is virtually unaffected by the introduction of virtual country fixed effects, suggesting that the result is just as strong comparing cells only within the virtual countries that have state sites.

Panel B shows the instrumental variable estimates. We find large, precise and stable coefficients across all specifications except when introducing virtual country fixed effects in model (6), which, while maintaining precision, lead to a four-fold increase in the coefficient. The magnitude in models (1)-(5) implies a one standard deviation increase in circumscription leads to an increase in probability of finding a state site in a cell of around 5.7%. Off of the minuscule base of 0.03%, this is obviously an enormous effect that needs to be interpreted with caution. However, the stability and precision of the results speaks for the possibility of substantial attenuation bias due to measurement error in the OLS specifications, and for the importance of circumscription as a driver of early state formation.

6 The Intensity of Extraction across Civilizations

[Figure 6 about here.]

The extent to which early state formation was driven by circumscription may differ from one civilization to the next. To assess heterogeneity in a straightforward way, we simply plot average circumscription by civilization in figure VI. The skew in the distribution of circumscription is due to the fact that most cells have at least one cell in their vicinity with positive land quality, even if their own land quality is zero.

We see that circumscription indeed differs substantially from one civilization to the next, in ways that are largely consistent with the qualitative literature on circumscription: the original examination by Carneiro (1970) begins with contrasting the extent of circumscription of agricultural groups in the Andes with those in the Amazon basin. The evidence in figure VI confirms his assessment of the Andes being a strongly circumscribed region, with average circumscription of Andean state sites being more than two standard deviations higher than the mean. Egypt is the other classic example frequently employed (Allen, 1997), and it also shows up as heavily circumscribed. On the other hand, the formation of states in China is much less closely associated with circumscription, although the data here speaks for circumscription being an important source of state formation in China as well.

Not all civilizations have circumscribed early state sites. At the bottom of the list rank the Maya, whose sites are actually on average negatively circumscribed, meaning they are on average surrounded by better land than their own. The fact that the Maya do not adhere to the pattern of circumscription is consistent with the detailed archaeological assessment by Trigger (2003), who concludes that the Mayan region is not circumscribed. McAnany (2014) finds that the Maya could have expanded further South into Central America, offering an escape route to state evaders. The Indus Valley civilization, while about 0.4 standard deviations above the mean, also ranks low in terms of circumscription. This stands contrary to the qualitative archaeological assessment. However, the aridification of the Indus region due to the weakening of the Monsoon is well documented in the archaeological and paleoclimatic record (Giosan et al., 2012), which may explain why it looks much less circumscribed today.

[Figure 7 about here.]

One possible source of this heterogeneity across civilizations may arise from the extent to which other drivers of early state formation were present. The strategic environment for outside groups to stage a challenge to incumbents is a particularly relevant alternative, as it directly connects to the model components of conflict, displacement and domination. The ruggedness of the surrounding terrain has been found to present a favorable strategic environment for conflict in studies of insurgency (Fearon and Laitin, 2003) and slave capture (Nunn and Puga, 2012). Figure VII plots average circumscription and average surrounding ruggedness by civilization. We interpret higher circumscription relative to surrounding ruggedness primarily as evidence for extraction, while the reverse points towards other motives such as protection.

It is interesting to note that the three civilizations that eventually formed the first territorial early states – China, Egypt and the Andes – all show relatively higher circumscription, while civilizations that maintained city-state systems for a long time, such as the Aztec, the Maya and Mesopotamia, are relatively more surrounded by rugged terrain. Also noteworthy is that the three civilizations that brought about early territorial states were also arguably the most extractive, with substantial agricultural shares and corvée labor collected by the state (Finer, 1997).

The Andean civilization stands testament to the massive extent of the extractive state that was possible in areas that were both circumscribed and surrounded by rugged terrain. The conquistador Pedro Cieza de León noted that the population in today's Peru was easy to control because of the lack of refuge for dissidents (Salomon, 1986). The extractive institution of the *Mita* maintained by the Spanish in Peru until 1812 was a remnant of the Inca state designed as a means to extract agricultural labor (see Dell (2010) and references therein). This example illustrates the long-lasting impact of the geographic environment on economic development through the extractive institutions of early states.

7 Conclusion

This paper shows how the lack of outside options for mobile groups enhances the capacity for extraction through states. Based on this argument, it then provides evidence that the quality of land surrounding an area is an important cause of state formation, as important or even more important than the quality of land of the area itself. Thereby, it introduces the notion of circumscription as a driving force of political economic dynamics.

The strong relationship of archaeological excavation sites related to state formation with circumscription masks substantial heterogeneity across civilizations. We explore the possibility of ruggedness enhancing the role of circumscription by offering strategic opportunities for challengers: as more conflicts take place due to the favorable strategic environment offered by mountains, more trials of state creation are undergone, leading to the formation of more early states. Another possible explanation for the heterogeneity is that the migratory capacity of groups may have depended on other factors besides the configuration of land quality. For example, Diamond (1998) suggests that it is harder to transport crops and agricultural techniques along a meridian (e.g. from North to South) than along circles of latitude (e.g. East to West) due to the relatively higher ecological similarity along the East-West axis. This would make an escape out of the narrow, ecologically diverse isthmus connecting the Americas harder than other in other environments.

It is also noteworthy that circumscription may be a necessary condition for early state formation, but it is not sufficient: there are several regions across the globe that show substantial circumscription but apparently never brought about early states. This is particularly true for coastal regions of Argentina and Uruguay. While there is a lively about the role of population density and more complex societies (Powell et al., 2009; Bettinger, 2016), one explanation for the absence of states in this region is that there was simply not enough time for human societies to grow large enough in this region for the dynamics of circumscription to kick in. This possibility underlines the issue that circumscription may favor the development of different societies at different scales of circumscribed areas, which may be depend on the overall size of human societies in a given region.

The debate in economics about the origins of states is centered on the theme of extraction, with some research on the role of protection. These are the classic motives identified in political science and economics and to which this paper contributes. However, a more comprehensive organization may include the distinction from evolutionary archaeology into voluntaristic and coercive theories. The former focus on the state's ability to generate a surplus, while the latter focus on the state's ability to extract resources from its subjects. Putting together both traditions yields a conceptual triad of protection, extraction and surplus production that may prove useful in future research.

More generally, this paper adds to the growing economics literature highlighting the role of spatial proximity for social phenomena in a given place. This notion has been fruitfully exploited in such diverse fields as international trade (Allen and Arkolakis, 2014) or real estate and crime Linden and Rockoff (2008). In this sense, we contribute to the intellectual development of what has been called geography's “second law”: ”the phenomenon external to an area of interest affects what goes on in the inside” (Tobler, 1999).

Finally, the results presented here also point towards the dearth of globally consistent archaeological data currently available to researchers to study the rise of complex societies, and state societies in particular. Boguckis collection of sites satisfies some basic conditions of a global, geocoded dataset of archaeological sites relevant to state creation, but larger and more ambitious data collection and standardization exercises may bring about much more precise insights into the mechanics of early state formation. In this sense, efforts such as Wrights Atlas of Chiefdoms and Early States Wright (2006) promise to open up new approaches to study this question through the collaboration of economists and archaeologists.

A Appendix

A.1 Data Sources and Preparation

All datasets except the archaeological data from Bogucki (1999) are raster data (i.e. grid-cell level data), usually at resolution of 1⁄12 or higher. These datasets have been down-sampled to 1⁄4 (or another cell size) using nearest-neighbor interpolation.

A.1.1 Archaeological Sites Related to State Formation

We use the maps in Bogucki's (1999) chapter 8 ("Early States and Chiefdoms in the Shadow of States"). Since the maps are no longer digitally available, all relevant sites were geocoded by hand using google maps, relying on satellite evidence of the site whenever possible. We exclude chiefdom sites. We also exclude state sites that have been deemed not pristine in the text. This leaves us with 68 sites spread across Mesoamerica, South America, Africa, the Near East, South Asia and East Asia.

A.1.2 FAO Agricultural Data and Derivates

Land Quality Proxy. Agricultural productivity is the maximum potential production capacity in tons per hectare over the seventeen crops (buckwheat, barley, chickpea, foxtail millet, groundnut, maize, oat, pearl millet, wetland rice, rape, rye, sunflower, soybean, sweet potato, sorghum, wheat and white potato) after correcting for the Columbian Exchange (e.g. no wheat in the Americas, no potatoes in Eurasia). See http://gaez.fao.org/Main.html for the database.

We use the earliest data available (baseline period 1961-1990), with intermediate input level and irrigated water supply. While the data is also available for rainfall water supply, it renders areas that were very productive but relied on river flooding and alluvial agriculture completely unproductive if they are in arid environments. Note that most early civilizations substantially relied on irrigation based on freshwater delivered via rivers. As such, the data with rainfall water supply does not represent a good proxy for agricultural land quality when other water sources are available. For example, according to the FAO GAEZ data on potential production capacity using low input level and rainfed agriculture (and data derived from it, such as Galor and Özak's (2016) Caloric Suitability Index), Egypt, including all of the Nile valley and most of the Delta, are completely unproductive, despite thousands of years of highly productive agriculture. Concerning the input level, the intermediate level is the lowest for irrigated water supply.

Next, we apply restrictions from the Columbian Exchange on which crops are available to a given region. To this end, we divide the crops into Old World crops, applied to Africa, Asia and Europe, and New World Crops, applied to the Americas. New World crops are maize, sweet potato, white potato and sunflower; the others are Old World crops.

In the next step, each crop is scaled using the historical Food Composition Tables by the FAO (Chatfield, 1953), using the following calorie values per 100 grams for each crop:

- buckwheat: 330
- barley: 332

- chickpea: 345
- foxtail millet: 343
- groundnut: 388
- maize: 356
- oat: 385
- pearl millet: 348
- wetland rice: 357
- rape: 26
- rye: 319
- sunflower: 284
- soybean: 335
- sweet potato: 97
- sorghum: 343
- wheat: 334
- white potato: 70

Finally, we take the maximal calorie amount across all crops available in a region as the land quality proxy.

Circumscription. To compute circumscription, we subtract average land quality in all cells within radius r (without the cell of interest) from the value at the cell of interest, as described in the main text. This is illustrated in figure VIII.

[Figure 8 about here.]

Standard Deviation of Land Quality. The standard deviation in land quality is computed for each cell i as

$$\text{StdDev}(a_i) = \sqrt{\frac{1}{N_r - 1} \sum_{j:\text{dist}(i,j)\leq r} (a_j - \bar{a}_{ir})^2}$$

where N_r is the number of cells within r cell units; and $\bar{a}_{ir}$ is the average land quality wihthin a radius of r. As usual, we use a default radius of $r = 10$.

A.1.3 Potential Vegetation and Derivates

SAGE Potential Vegetation. Global potential vegetation data is from the Center for Sustainability and the Global Environment (SAGE) in the Nelson Institute at the University of Wisconsin-Madison. The data consists of a global map of natural vegetation at a 5 min resolution classified into 15 vegetation types. These are: tropical evergreen forest / woodland, tropical deciduous forest / woodland, temperate broadleaf evergreen forest / woodland, temperate needleleaf evergreen forest / woodland, temperate deciduous forest / woodland, boreal evergreen forest / woodland, boreal deciduous forest / woodland, evergreen / deciduous mixed forest / woodland, savanna, grassland / steppe, dense shrubland, open shrubland, tundra, desert, and polar desert / rock / ice. The data is representative of the world's "potential" vegetation (i.e., vegetation that would most likely exist now in the absence of human activities). See Ramankutty and Foley (1999) for further details. The data is available at https://nelson.wisc.edu/sage/data-and-models/global-potential-vegetation/index.php.

The SAGE data is also being used in section 5 to identify desert cells. Note that given that it is based on potential vegetation cells, the desert measure should be unaffected by human activity that may have led to desertification.

Ecological Diversity and Polarization Indices. As described in Fenske (2014), the ecological diversity index is the Herfindahl index constructed from shares s^t of various ecological types $t = 1, ..., T$. We use SAGE's potential vegetation as the ecological types. In the case of our grid cell data set, we compute for each cell i:

$$\text{EcoDiv}_{ir} = 1 - \sum_{t=1}^{T} \left(s_{ir}^{t}\right)^2$$

where

$$s_{ir}^{t} = \frac{1}{N_r} \sum_{j:\text{dist}(i,j) \leq r} n_j^t$$

with n_i^t being an indicator for cell i having ecological type t and N_r as the sum of cells within r as before. As usual, we use $r = 10$ as our default radius. Similarly, ecological polarization is

$$\text{EcoDiv}_{ir} = 1 - \sum_{t=1}^{T} \left(1 - 2s_{ir}^{t}\right)^2 s_{ir}^{t}.$$

A.1.4 Mean and Standard Deviation Temperature

Data on the annual mean and standard deviation temperature are from the WorldClim global climate database. We use current condition (1960-1990) 5 minute resolution data for mean monthly temperature. This is averaged over the year to get annual mean temperature for each cell. The standard deviation is taken across months for each grid cell. See http://www.worldclim.org/ for the data source.

A.1.5 Ruggedness and Slope

The Terrain Ruggedness Index was developed in Riley et al. (1999). The data on ruggedness and slope used here the grid-cell level data prepared for Nunn and Puga (2012). The data is described in detail on their data repository at:

http://diegopuga.org/data/rugged/.

A.1.6 River Flow Accumulation

River flow accumulation data is from the Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) project, which offers hydrographic information in a consistent and comprehensive format on a global scale. River flow accumulation is derived primarily from elevation data of the Shuttle Radar Topography Mission (SRTM) at 3 arc-second resolution and measures the number of cells of drainage accumulation due to the elevation data. See http://www.hydrosheds.org/.

A.2 Further Tables and Figures

[Table 5 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

[Figure 11 about here.]

[Figure 12 about here.]

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2002) "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*," *The Quarterly Journal of Economics*, Vol. 117, No. 4, p. 1231.

——— (2005) "Institutions as a Fundamental Cause of Long-Run Growth," *Handbook of Economic Growth*, Vol. 1, Part A, pp. 385 – 472.

Allen, Robert C (1997) "Agriculture and the Origins of the State in Ancient Egypt," *Explorations in Economic History*, Vol. 34, No. 2, pp. 135–154.

Allen, Treb and Costas Arkolakis (2014) "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, Vol. 129, No. 3, pp. 1085–1140.

Angrist, Joshua D and Jörn-Steffen Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*: Princeton university press.

Arranz-Otaegui, Amaia, Sue Colledge, Lydia Zapata, Luis Cesar Teira-Mayolini, and Juan José Ibáñez (2016) "Regional diversity on the timing for the initial appearance of cereal cultivation and domestication in southwest Asia," *Proceedings of the National Academy of Sciences*, Vol. 113, No. 49, pp. 14001–14006.

Ashraf, Quamrul and Stelios Michalopoulos (2015) "Climatic Fluctuations and the Diffusion of Agriculture," *Review of Economics and Statistics*, Vol. 97, No. 3, pp. 589–609.

Bates, Robert H (1987) *Essays on the Political Economy of Rural Africa*, Vol. 38: Univ of California Press.

Benson, Larry, Kenneth Petersen, and John Stein (2007) "Anasazi (Pre-Columbian Native-American) Migrations During The Middle-12Th and Late-13th Centuries – Were they Drought Induced?" *Climatic Change*, Vol. 83, No. 1, pp. 187–213.

Berniell-Lee, Gemma, Francesc Calafell, Elena Bosch, Evelyne Heyer, Lucas Sica, Patrick Mouguiama-Daouda, Lolke van der Veen, Jean-Marie Hombert, Lluis Quintana-Murci, and David Comas (2009) "Genetic and Demographic Implications of the Bantu Expansion: Insights from Human Paternal Lineages," *Molecular Biology and Evolution*, Vol. 26, No. 7, p. 1581.

Besley, Timothy and Torsten Persson (2009) "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, Vol. 99, No. 4, pp. 1218–44.

——— (2010) "State Capacity, Conflict, and Development," *Econometrica*, Vol. 78, No. 1, pp. 1–34.

Bettinger, Robert L. (2016) "Prehistoric hunter–gatherer population growth rates rival those of agriculturalists," *Proceedings of the National Academy of Sciences*, Vol. 113, No. 4, pp. 812–814.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman (2002) "States and Markets: The Advantage of an Early Start," *Journal of Economic Growth*, Vol. 7, No. 4, pp. pp. 347–369.

Bogucki, Peter (1999) *The Origins of Human Society*: Blackwell Publishers.

Brumfiel, Elizabeth M (1994) *The Economic Anthropology of the State*, No. 11: Lanham: University Press of America.

Burchfield, Marcy, Henry G Overman, Diego Puga, and Matthew A Turner (2006) "Causes of Sprawl: A Portrait from Space," *The Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 587–633.

Carneiro, Robert L. (1970) "A Theory of the Origin of the State," *Science*, Vol. 169, No. 3947, pp. 733–738.

Chatfield, Charlotte (1953) "Food Composition Tables for International Use," in *FAO Nutritional Studies No. 3.*: Nutrition Division, FAO UN, Washington, DC, p. 56 pp.

Childe, V Gordon (1936) *Man Makes Himself*: London, Watts & Co.

Conley, T.G. (1999) "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, Vol. 92, No. 1, pp. 1 – 45.

Crowther, Alison, Leilani Lucas, Richard Helm, Mark Horton, Ceri Shipton, Henry T. Wright, Sarah Walshaw, Matthew Pawlowicz, Chantal Radimilahy, Katerina Douka, Llorenç Picornell-Gelabert, Dorian Q. Fuller, and Nicole L. Boivin (2016) "Ancient crops provide first archaeological signature of the westward Austronesian expansion," *Proceedings of the National Academy of Sciences*, Vol. 113, No. 24, pp. 6635–6640.

Dal Bó, Ernesto, Pablo Hernández, and Sebastián Mazzuca (2016) "The Paradox of Civilization: Pre-Institutional Sources of Security and Prosperity," *Working Paper*.

Dell, Melissa (2010) "The Persistent Effects of Peru's Mining "Mita"," *Econometrica*, Vol. 78, No. 6, pp. pp. 1863–1903.

Diamond, Jared and Peter Bellwood (2003) "Farmers and Their Languages: The First Expansions," *Science*, Vol. 300, No. 5619, pp. 597–603.

Diamond, Jared M (1998) *Guns, Germs, and Steel: A Short History of Everybody for the Last 13,000 Years*: Random House.

Erickson, David L., Bruce D. Smith, Andrew C. Clarke, Daniel H. Sandweiss, and Noreen Tuross (2005) "An Asian origin for a 10,000-year-old domesticated plant in the Americas," *Proceedings of the National Academy of Sciences*, Vol. 102, No. 51, pp. 18315–18320.

Fearon, James D. and David D. Laitin (2003) "Ethnicity, Insurgency, and Civil War," *American Political Science Review*, Vol. 97, No. 1, pp. 75–90.

Fenske, James (2013) "Does Land Abundance Explain African Institutions?" *The Economic Journal*, Vol. 123, No. 573, pp. 1363–1390.

——— (2014) "Ecology, Trade, and States in Pre-Colonial Africa," *Journal of the European Economic Association*, pp. n/a–n/a.

Finer, Samuel Edward (1997) *The History of Government from the Earliest Times: Ancient Monarchies and Empires*, Vol. 1: Oxford University Press.

Fischer, Guenther, Freddy Nachtergaele, Sylvia Prieler, HT Van Velthuizen, Luc Verelst, and David Wiberg (2008) "Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008),"Technical report, IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Galor, Oded and Ömer Özak (2016) "The Agricultural Origins of Time Preference," *American Economic Review*, Vol. 106, No. 10, pp. 3064–3103.

Gennaioli, Nicola and Hans-Joachim Voth (2015) "State Capacity and Military Conflict," *The Review of Economic Studies*, Vol. 82, No. 4, pp. 1409–1448.

Giosan, Liviu, Peter D. Clift, Mark G. Macklin, Dorian Q. Fuller, Stefan Constantinescu, Julie A. Durcan, Thomas Stevens, Geoff A. T. Duller, Ali R. Tabrez, Kavita Gangal, Ronojoy Adhikari, Anwar Alizai, Florin Filip, Sam VanLaningham, and James P. M. Syvitski (2012) "Fluvial Landscapes of the Harappan Civilization," *Proceedings of the National Academy of Sciences*, Vol. 109, No. 26, pp. E1688–E1694.

Haas, Jonathan (1982) *The Evolution of the Prehistoric State*: Columbia University Press.

Hijmans, Robert J., Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis (2005) "Very high resolution interpolated climate surfaces for global land areas," *International Journal of Climatology*, Vol. 25, No. 15, pp. 1965–1978.

Hofmanová, Zuzana, Susanne Kreutzer, Garrett Hellenthal, Christian Sell, Yoan Diekmann, David Díez-del Molino, Lucy van Dorp, Saioa López, Athanasios

Kousathanas, Vivian Link, Karola Kirsanow, Lara M. Cassidy, Rui Martiniano, Melanie Strobel, Amelie Scheu, Kostas Kotsakis, Paul Halstead, Sevi Triantaphyllou, Nina Kyparissi-Apostolika, Dushka Urem-Kotsou, Christina Ziota, Fotini Adaktylou, Shyamalika Gopalan, Dean M. Bobo, Laura Winkelbach, Jens Blöcher, Martina Unterländer, Christoph Leuenberger, Çiler Çilingiroğlu, Barbara Horejs, Fokke Gerritsen, Stephen J. Shennan, Daniel G. Bradley, Mathias Currat, Krishna R. Veeramah, Daniel Wegmann, Mark G. Thomas, Christina Papageorgopoulou, and Joachim Burger (2016) "Early farmers from across Europe directly descended from Neolithic Aegeans," *Proceedings of the National Academy of Sciences*, Vol. 113, No. 25, pp. 6886–6891.

Horrace, William C. and Ronald L. Oaxaca (2006) "Results on the bias and inconsistency of ordinary least squares for the linear probability model," *Economics Letters*, Vol. 90, No. 3, pp. 321 – 327.

Johnson, Allen W and Timothy K Earle (2000) *The Evolution of Human Societies: From Foraging Group to Agrarian State*: Stanford University Press.

Kleibergen, Frank and Richard Paap (2006) "Generalized Reduced Rank Tests using the Singular Value Decomposition," *Journal of Econometrics*, Vol. 133, No. 1, pp. 97 – 126.

Larson, Greger, Ranran Liu, Xingbo Zhao, Jing Yuan, Dorian Fuller, Loukas Barton, Keith Dobney, Qipeng Fan, Zhiliang Gu, Xiao-Hui Liu, Yunbing Luo, Peng Lv, Leif Andersson, and Ning Li (2010) "Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA," *Proceedings of the National Academy of Sciences*, Vol. 107, No. 17, pp. 7686–7691.

Lehner, B., K. Verdin, and A. Jarvis (2008) "New Global Hydrography Derived From Spaceborne Elevation Data," *Eos, Transactions American Geophysical Union*, Vol. 89, No. 10.

Linden, Leigh and Jonah E. Rockoff (2008) "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws," *American Economic Review*, Vol. 98, No. 3, pp. 1103–27.

Mann, Michael (1986) "The Sources of Social Power: A History of Power from the Beginning to AD 1760, volume I."

Matranga, Andrea (2017) "The Ant and the Grasshopper: Seasonality and the Invention of Agriculture," *Working Paper*.

Mayshar, Joram, Omer Moav, Zvika Neeman, and Luigi Pascali (2016) "Cereals, Appropriability and Hierarchy," *Warwick Economics Research Papers*.

McAnany, Patricia A (2014) *Living with the Ancestors: Kinship and Kingship in Ancient Maya Society*: Cambridge University Press.

Michalopoulos, Stelios and Elias Papaioannou (2013) "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, Vol. 81, No. 1, pp. 113–152.

——— (2014) "National Institutions and Subnational Development in Africa," *The Quarterly Journal of Economics*, Vol. 129, No. 1, pp. 151–213.

Murdock, George Peter (1967) *Ethnographic Atlas*: Pittsburgh, PA: University of Pittsburgh Press.

Nunn, Nathan and Diego Puga (2012) "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, Vol. 94, No. 1, pp. 20–36.

Nunn, Nathan and Nancy Qian (2010) "The Columbian Exchange: A History of Disease, Food, and Ideas," *Journal of Economic Perspectives*, Vol. 24, No. 2, pp. 163–88.

Olson, Mancur (1993) "Dictatorship, Democracy, and Development," *American Political Science Review*, pp. 567–576.

Olsson, Ola and Douglas A Hibbs Jr (2005) "Biogeography and Long-run Economic Development," *European Economic Review*, Vol. 49, No. 4, pp. 909–938.

Powell, Adam, Stephen Shennan, and Mark G. Thomas (2009) "Late Pleistocene Demography and the Appearance of Modern Human Behavior," *Science*, Vol. 324, No. 5932, pp. 1298–1301.

Putterman, Louis (2008) "Agriculture, Diffusion and Development: Ripple Effects of the Neolithic Revolution," *Economica*, Vol. 75, No. 300, pp. pp. 729–748.

Ramankutty, Navin and Jonathan A. Foley (1999) "Estimating historical changes in global land cover: Croplands from 1700 to 1992," *Global Biogeochemical Cycles*, Vol. 13, No. 4, pp. 997–1027.

Riley, Shawn J, Stephen J DeGloria, and Robert Elliot (1999) "Index That Quantifies Topographic Heterogeneity," *Intermountain Journal of Sciences*, Vol. 5, No. 1–4, pp. 23–27.

Salomon, Frank (1986) *Native Lords of Quito in the Age of the Incas: The Political Economy of North Andean Chiefdoms*: Cambridge University Press Cambridge.

Scott, James C (2009) *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*: Yale University Press.

Service, Elman R (1975) *Origins of the State and Civilization*: New York. Norton.

Sanchez de la Sierra, Raul (2017) "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo," *Working Paper*.

Tilly, Charles (1990) *Capital, Coercion and European States*: Cambridge, MA: Basil Blackwell.

Tobler, Waldo (1999) "Linear Pycnophylactic Reallocation Comment on a Paper by D. Martin," *International Journal of Geographical Information Science*, Vol. 13, No. 1, pp. 85–90.

Trigger, Bruce G (2003) *Understanding Early Civilizations: A Comparative Study*: Cambridge University Press.

Wright, Henry T (2006) "Atlas of Chiefdoms and Early States," *Structure and Dynamics*, Vol. 1, No. 4.

List of Tables

Table I: Summary statistics for global grid dataset.

Variable	Mean	Median	Std.dev.	Min	Max
State site indicator	0.00033	0	0.018	0	1
Circumscription in MCal/ha	370.2	-120.6	5348.5	-31280.6	32676.2
Land quality in MCal/ha	4034.4	0	7818.9	0	40722.8
S.d. land quality in MCal/ha	4155.7	3870.0	3489.7	0	16668.8
Ecological diversity index	0.29	0.20	0.28	0	0.87
Ecological polarization index	0.37	0.36	0.31	0	0.99
Mean of annual temperature in C°	14.5	17.0	10.7	-15.4	31.4
Std.dev. of annual temperature in C°	6.95	6.45	4.69	0.12	22.1
abs(latitude)	31.1	30.9	16.7	0.12	62.1
Slope in percent	2.89	0.86	4.95	0.0090	92.8
Ruggedness in meters	98.0	29.9	166.0	1	2660.0
log(Riverflow accumulation in km^2 cells)	4.10	3.80	2.07	-15.4	12.9
Desert indicator	0.12	0	0.33	0	1
N	177,299				

Note: State site indicator is one if there is at least one archaeological site relevant to state formation in it. Land quality is measured as maximal megacalories (MCal) per hectare (ha) across the seventeen crops in the FAO database, using historical calorie tables to scale them and correcting for the Columbian Exchange. See text on how circumscription is computed. The radius of $r = 10$ applies to circumscription, std.dev. land quality, the ecological diversity index, and the ecological polarization index. The number of observations is lower than the number of land cells with land quailty data (194,102) because riverflow accumulation does not cover some parts north of 60°. More details in the appendix.

Table II: Probit/Logit estimation of structural parameters (β, γ, c).

	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription: $\hat{\beta}$	0.16***		0.13***	0.56***		0.46***
	(0.03)		(0.05)	(0.08)		(0.16)
Land quality: $\hat{\gamma}$		0.14***	0.04		0.49***	0.12
		(0.02)	(0.04)		(0.06)	(0.13)
investment cost: $-\hat{c}$	-3.48***	-3.53***	-3.50***	-8.28***	-8.44***	-8.33***
	(0.04)	(0.04)	(0.04)	(0.15)	(0.15)	(0.15)
Estimator	Probit	Probit	Probit	Logit	Logit	Logit
Mean dependent variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Marginal effect of circumscription	0.00019		0.00015	0.00018		0.00015
Marginal effect of land quality		0.00017	0.00005		0.00016	0.00004
Test $\hat{\beta} = \hat{\gamma}$			0.28			0.22
Log-Likelihood	-525.08	-528.85	-524.73	-525.16	-529.43	-524.94
N	184,523	184,523	184,523	184,523	184,523	184,523
Pseudo-R^2	0.031	0.024	0.032	0.031	0.023	0.031

Note: Logit models in columns (1) to (3), and Probit models in (4) to (6). Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table III: Robustness estimates using OLS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Circumscription: $\hat{\beta}$	0.00028		0.00023	0.00052	0.00051	0.00051	0.00047	0.00054
	(0.00007)		(0.00009)	(0.00010)	(0.00010)	(0.00010)	(0.00010)	(0.00014)
	[0.00009]		[0.00011]	[0.00015]	[0.00015]	[0.00015]	[0.00015]	[0.00021]
	{0.00011}		{0.00013}	{0.00017}	{0.00017}	{0.00017}	{0.00016}	{0.00021}
Land quality: $\hat{\gamma}$		0.00025	0.00008	-0.00036	-0.00037	-0.00035	-0.00027	-0.00038
		(0.00005)	(0.00005)	(0.00007)	(0.00008)	(0.00008)	(0.00008)	(0.00014)
		[0.00008]	[0.00009]	[0.00012]	[0.00013]	[0.00013]	[0.00012]	[0.00020]
		{0.00009}	{0.00009}	{0.00012}	{0.00013}	{0.00013}	{0.00012}	{0.00019}
Mean dep. variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Ecol. div. controls				X	X	X	X	X
Climate controls					X	X	X	X
Topog. & env. controls						X	X	X
Pot. vegetation FEs							X	X
Virtual country FEs								X
N	184,523	184,523	184,523	184,523	184,523	177,299	177,299	177,289
Adjusted R^2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.007

Note: Linear probability models (via OLS). Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units. Robust standard errors in parentheses; standard errors clustered by $5° \times 5°$ virtual countries (712 clusters) in brackets; spatially robust (Conley) standard errors with radius 278km (standard cell size of 1/4 (28km at equator) times standard radius $r = 10$ in cell units) in curly brackets. Significance stars omitted. Ecological diversity controls includes the standard deviation of land quality, the ecological diversity index and the ecological polarization index; climate controls includes annual mean and standard deviation temperature and absolute latitude; topographic and environmental controls includes slope, ruggedness and river flow accumulation.

Table IV: Instrumenting using desert rivers.

Panel A:	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
RiverAccum × Desert	0.0161***	0.0160***	0.0046***	0.0009***	0.0009***	0.0010***
	(0.0012)	(0.0012)	(0.0005)	(0.0003)	(0.0003)	(0.0003)
Control variables	X	X	X	X	X	X
Pot. Vegetation FEs		X	X		X	X
Virtual Country FEs			X			X
N	177,299	177,299	177,289	177,299	177,299	177,289
R^2	0.764	0.771	0.952	0.001	0.002	0.012

Panel B:	IV estimates					
	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription	0.059***	0.057***	0.056***	0.055***	0.056***	0.210***
	(0.022)	(0.020)	(0.019)	(0.019)	(0.020)	(0.073)
Ecol. div. controls		X	X	X	X	X
Climate controls			X	X	X	X
Topo. & env. controls				X	X	X
Pot. vegetation FEs					X	X
Virtual country FEs						X
Partial F of instr.	53.2	162.7	175.1	176.5	180.0	93.0
N	177,299	177,299	177,299	177,299	177,299	177,289
Within R^2	0.008	0.095	0.110	0.112	0.114	0.001

Note: IV/2SLS estimates. Panel A: first stage (1)-(3) and reduced form (4)-(6). Panel B: IV estimates. In the first stage, the dependent variable is circumscription; in the reduced form and the IV estimates, it is an indicator for whether the cell has a state site. All regressions include land quality, river flow accumulation and a desert indicator as controls. Kleinbergen-Paap rk statistic (2006) as Partial F for instrument.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table V: Robustness using Probit, similar to table III.

	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription: $\hat{\beta}$	0.160		0.126	0.483	0.472	0.442
	(0.026)		(0.046)	(0.132)	(0.145)	(0.138)
	[0.030]		[0.064]	[0.240]	[0.255]	[0.239]
Land quality: $\hat{\gamma}$		0.142	0.041	-0.534	-0.526	-0.475
		(0.019)	(0.035)	(0.187)	(0.206)	(0.195)
		[0.027]	[0.061]	[0.346]	[0.368]	[0.346]
Mean dependent variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Ecological diversity controls				X	X	X
Climate controls					X	X
Topog. and environm. controls						X
N	184,523	184,523	184,523	184,523	184,523	177,299
Adjusted R^2						

Note: Probit model. Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units. Robust standard errors in parentheses; standard errors clustered by $5° \times 5°$ virtual countries (712 clusters) in brackets. Significance stars omitted. Ecological diversity controls includes the standard deviation of land quality, the ecological diversity index and the ecological polarization index; climate controls includes annual mean and standard deviation temperature and absolute latitude; topographic and environmental controls includes slope, ruggedness and river flow accumulation.

List of Figures

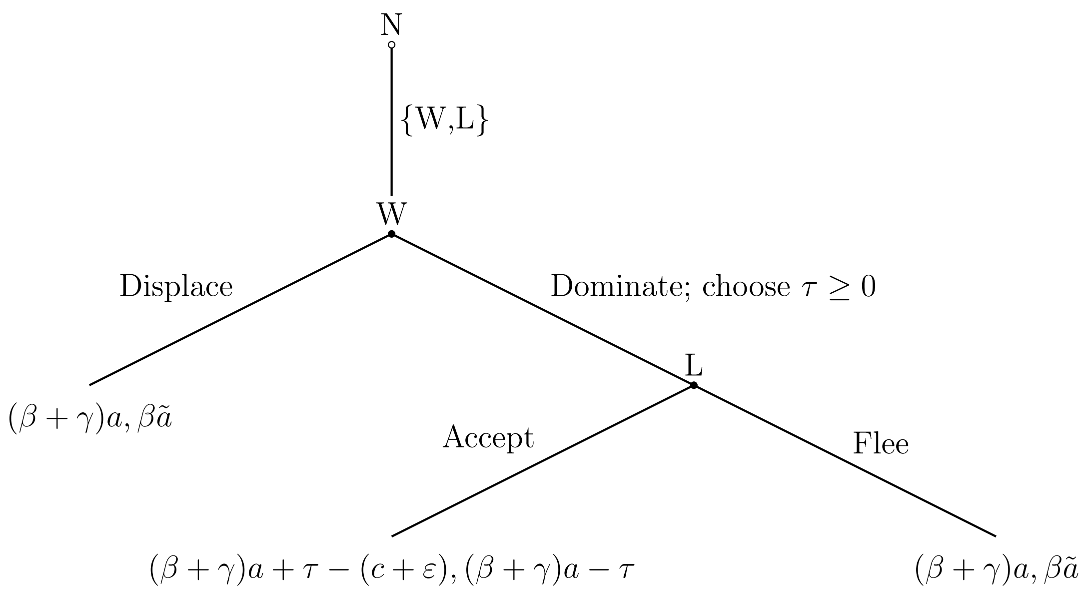

Figure I: A dynamic state formation game based on Carneiro (1970).

Dominate the loser of conflict (that is, construct the extractive state) if the agricultural gradient between the shared land and the outside option is large enough. The payoffs listed first are the winner's.

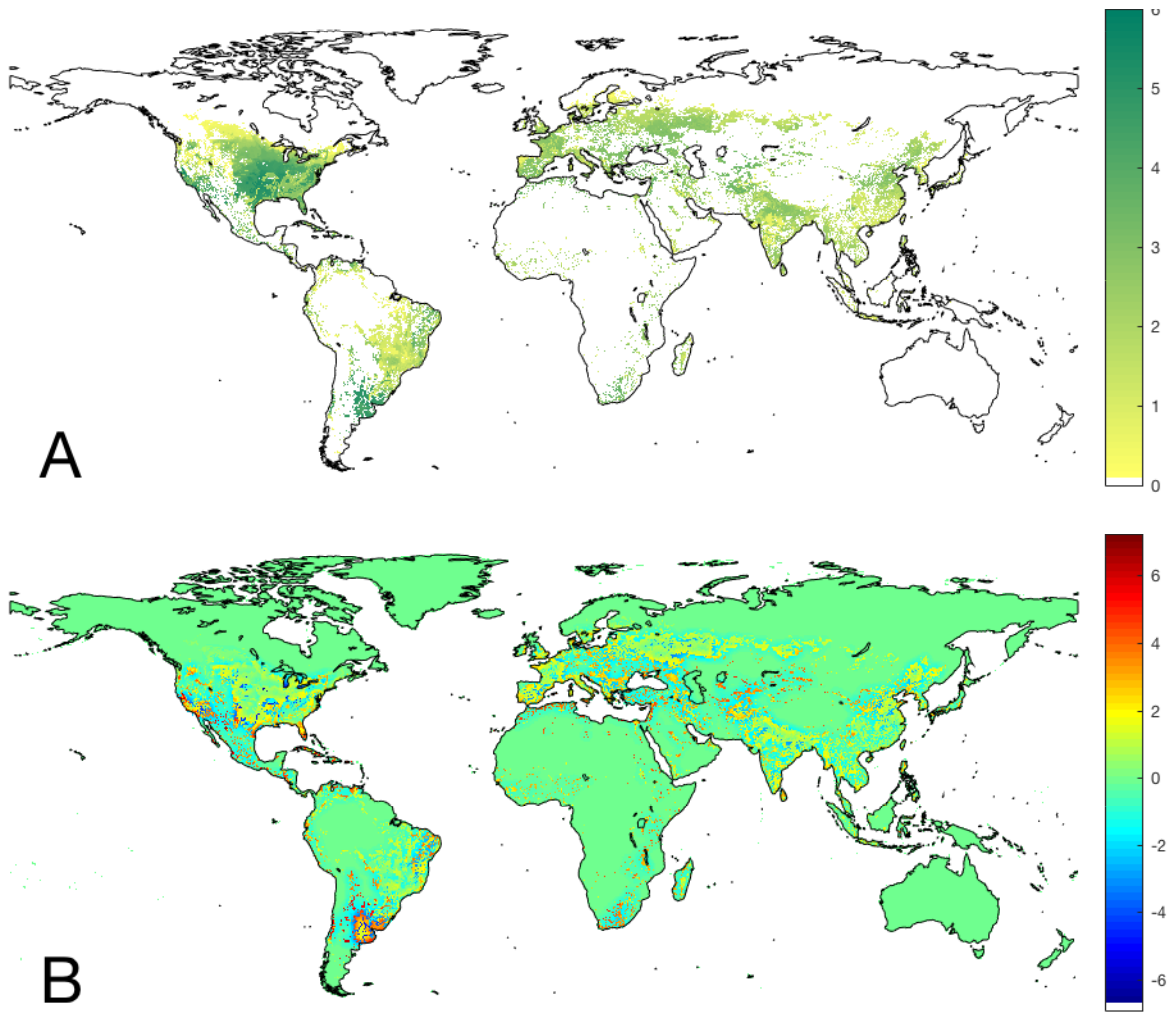

Figure II: Map of (A) land quality and (B) circumscription.

Panel A: Land quality proxy: maximal potential production capacity (FAO GAEZ database) over the 18 most important crops, scaled by calories of each crop and corrected for the Columbian Exchange, in standard deviation units, with cell size ¼. Panel B: Circumscription with cell size ¼ and $r = 10$.

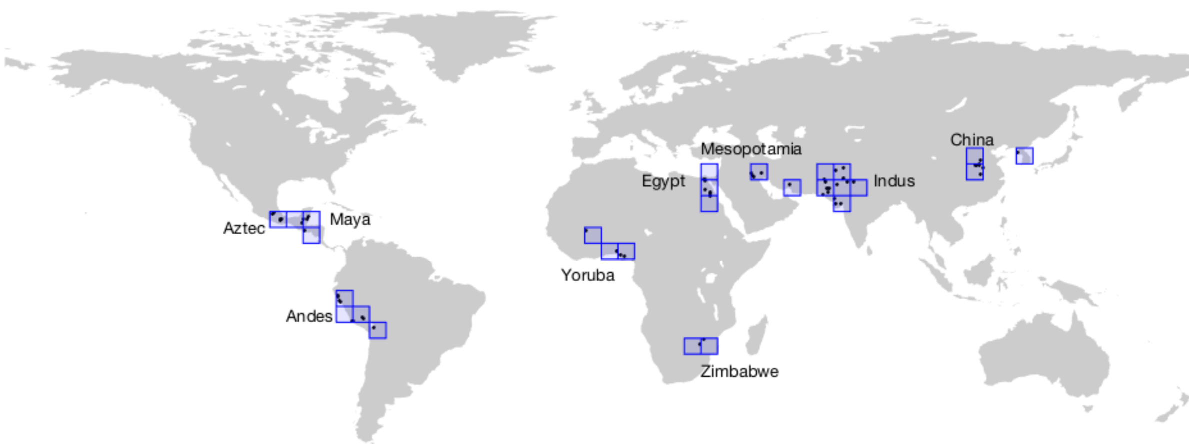

Figure III: Map of archaeological sites from (Bogucki, 1999).

Archaeological sites of early states (Bogucki, 1999) are shown as black dots. The blue rectangles show the twenty-seven $5^\circ \times 5^\circ$ virtual countries which have at least one early state site. The names of civilizations are common shorthands used for their respective regions.

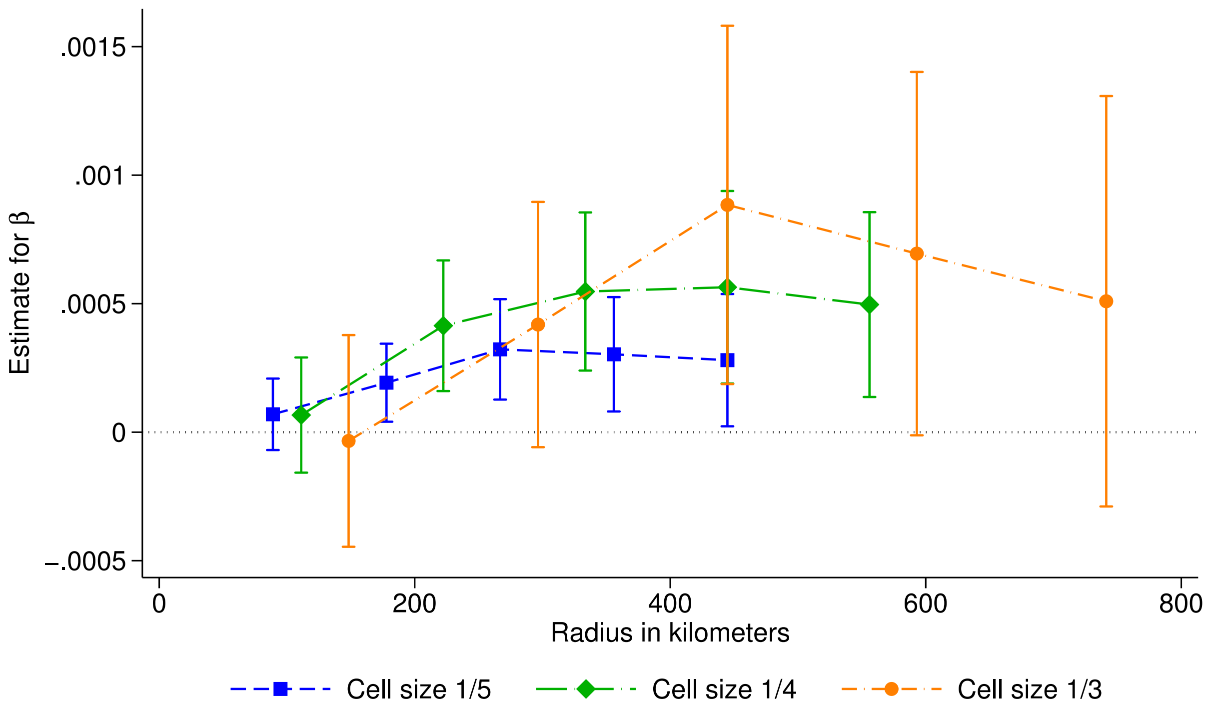

Figure IV: Coefficients for various cell sizes and radii.

Estimates for β from equation (3) and specification (8) from table III for various radii and cell sizes. Cell sizes are in degrees; radii are $r = 4, 8, 12, 16, 20$.

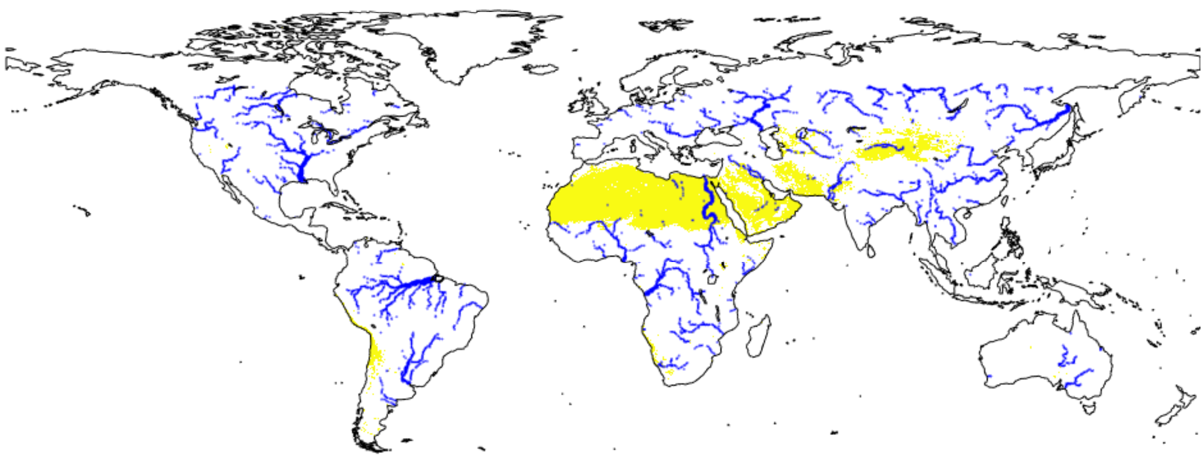

Figure V: Map of large rivers and deserts.

Areas with potential vegetation being deserts in yellow; the top one percent of river flow accumulation (rivers) in blue, with river flow accumulation scaled to magnitude.

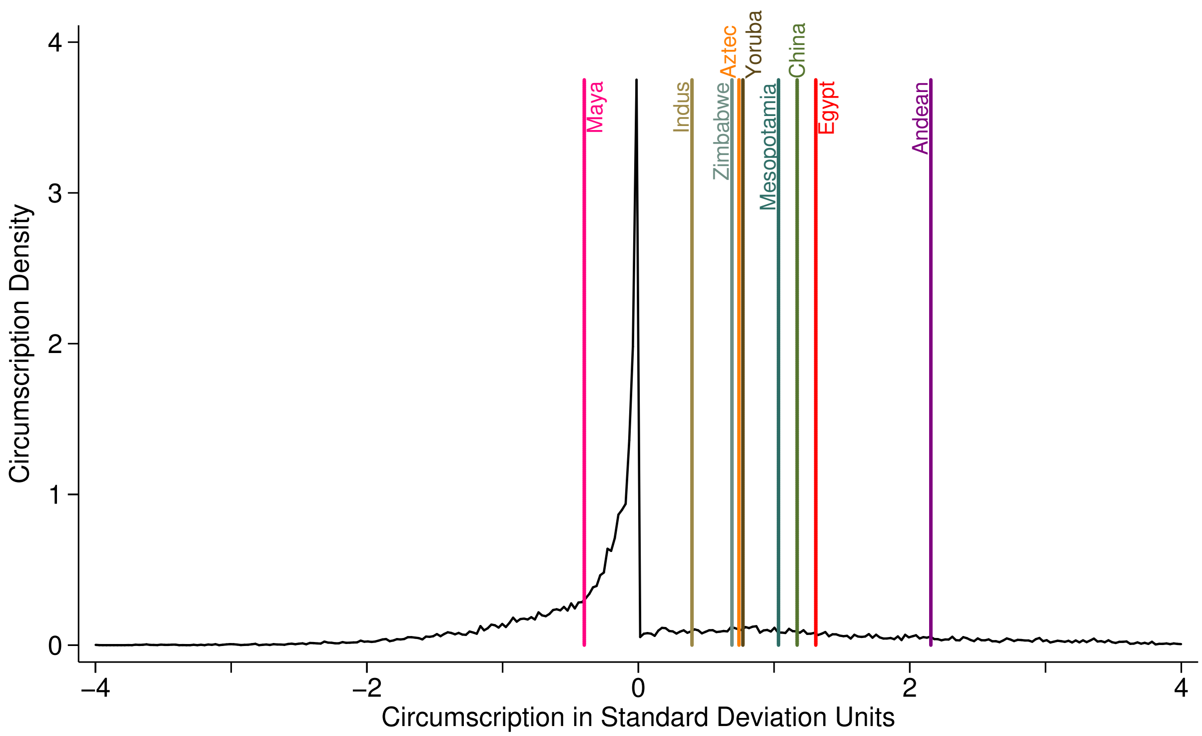

Figure VI: Density of circumscription and circumscription per civilization.

Estimated kernel density function of circumscription in black; average circumscription per civilization as vertical lines in various colors. The skew in the distribution of circumscription is due to the fact that most cells have at least one cell in their vicinity with positive land quality, even if their own land quality is zero.

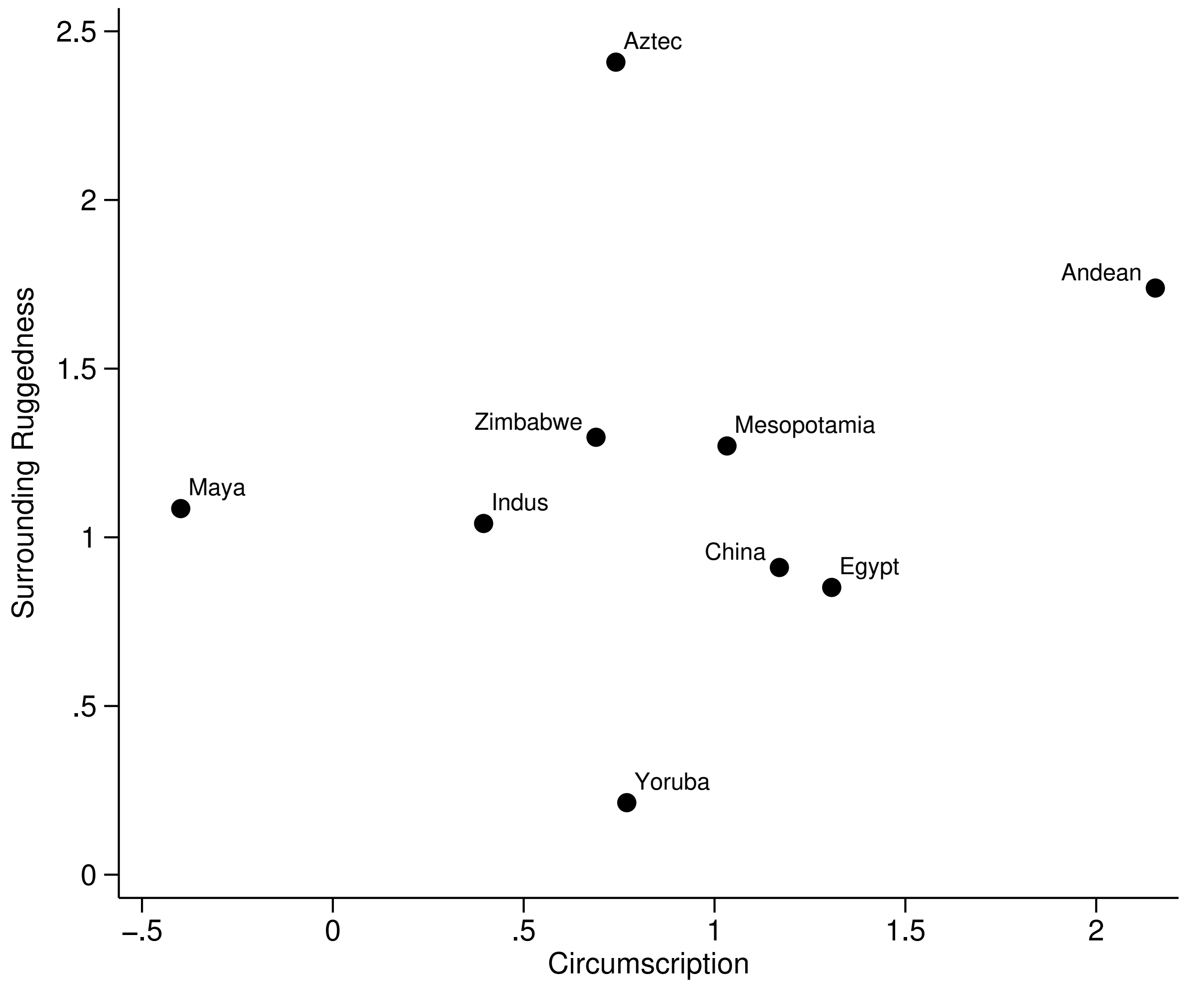

Figure VII: Scatter plot of ruggedness against circumscription by civilization.

Both variables are in standard deviation units.

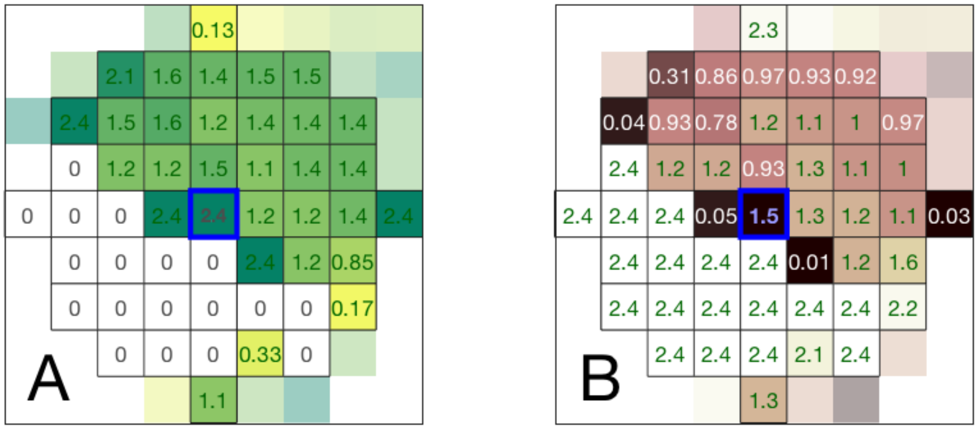

Figure VIII: Example map computing circumscription.

Panel A shows land quality in units of standard deviation using data from the Nile delta. How circumscribed is the cell with the blue frame? In panel B, we subtract the average value of all surrounding cells from the framed cell to arrive at 2.5. In other words, moving away from the framed cell into a random cell within 5 cell units results in land quality that is on average 2.5 standard deviation units lower.

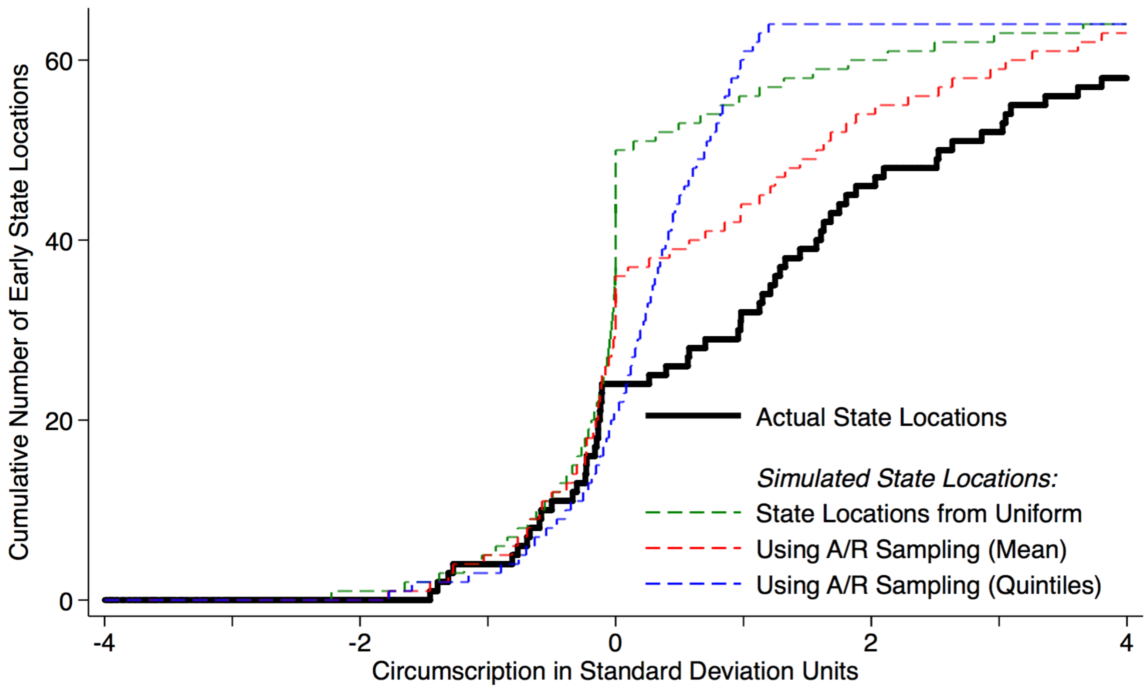

Figure IX: Empirical CDFs of actual and simulated state locations.

Empirical CDFs for actual state locations and simulated state locations using 1,000 iterations. Actual state locations in black (five sites are beyond the right edge of the figure). Simulated state locations are distributed according to three different assumptions. In green, state locations are uniformly drawn from all cells; in red, accept/reject sampling to match mean land quality of state cells (i.e. the sample is drawn from cells with the same mean land quality as state cells); in blue, accept/reject sampling to match the quintiles in land quality distribution of state cells.

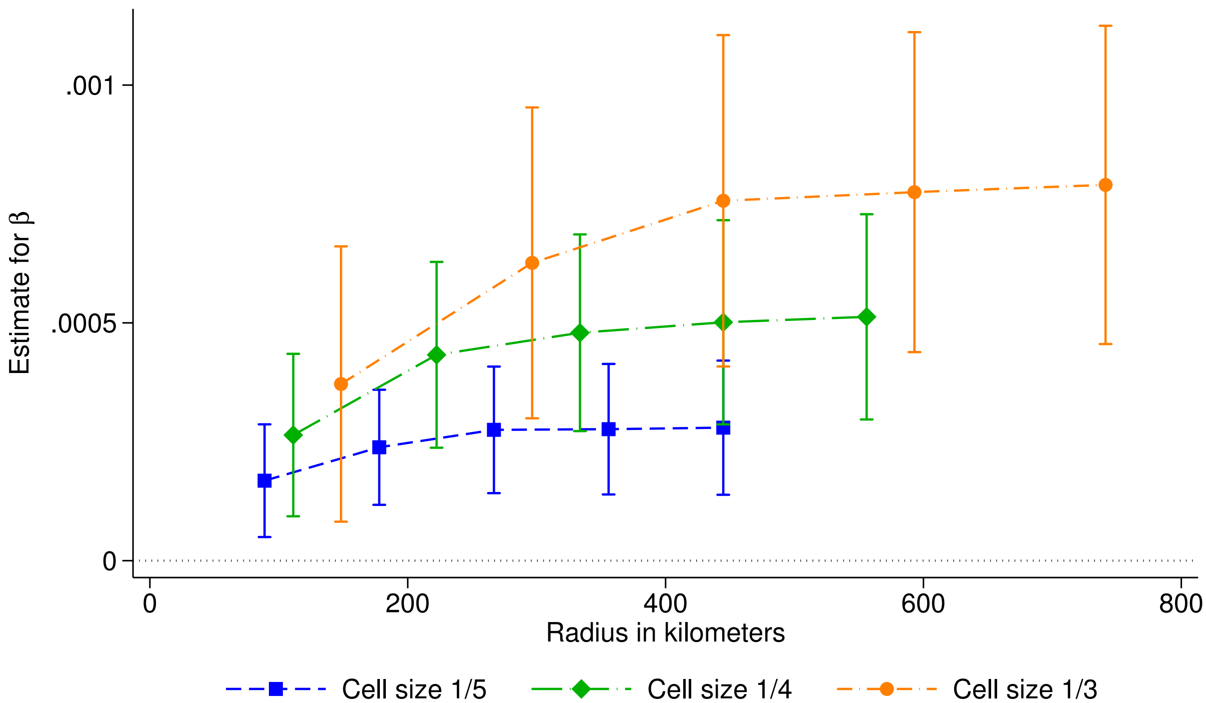

Figure X: Coefficients for various cell sizes and radii, alternative specification.

Estimates for β from equation (3) and specification (7) from table III for various radii and cell sizes. Cell sizes are in degrees; radii are $r = 4, 8, 12, 16, 20$.

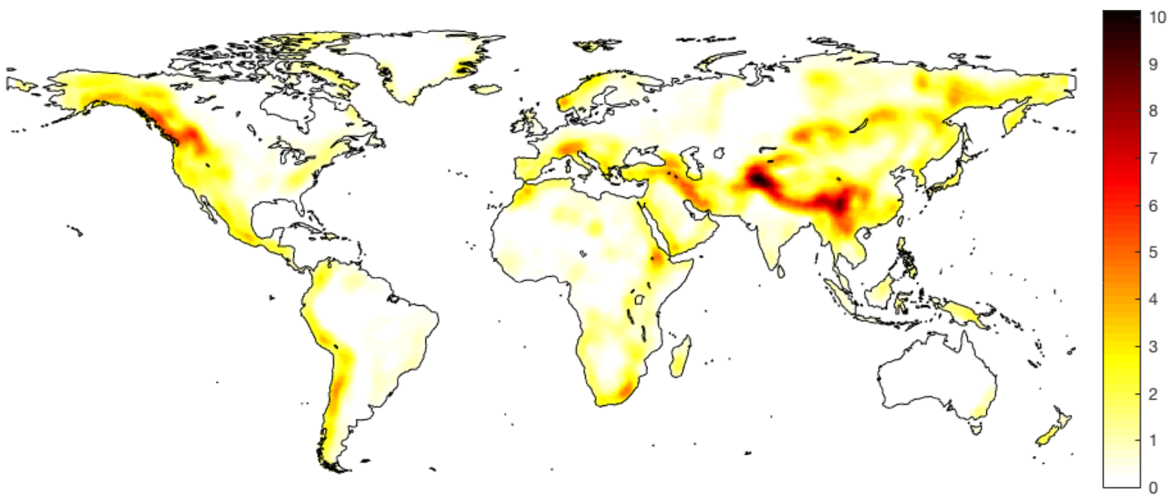

Figure XI: Map of surrounding ruggedness.

Surrounding ruggedness in standard deviation units. Standard radius of $r = 10$.

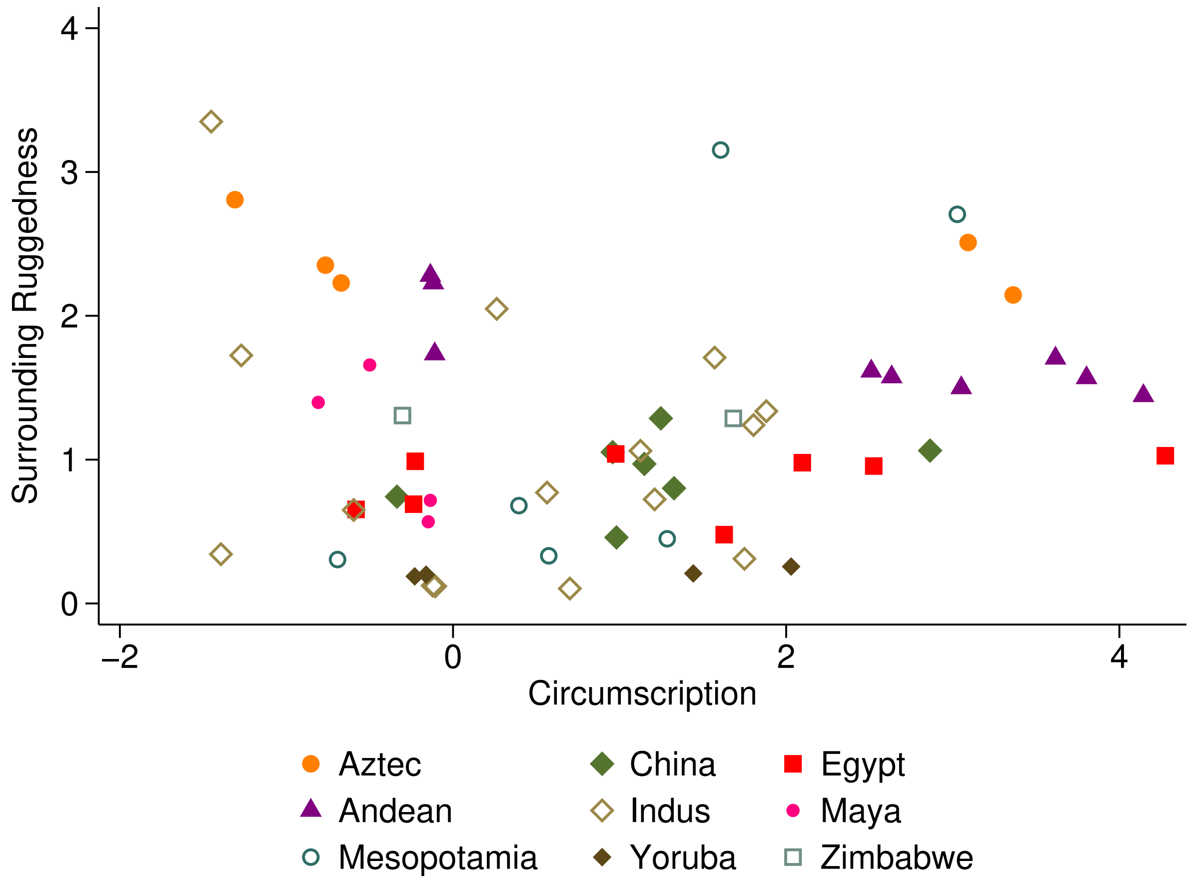

Figure XII: The extent of ruggedness and circumscription for all state sites.