The Emergence of Hierarchies and States: Productivity vs. Appropriability*

Joram Mayshar† Omer Moav‡ Zvika Neeman§ Luigi Pascali¶

January 9, 2018

Abstract

According to conventional theory, hierarchies and states were formed following the Neolithic Revolution because the increase in land productivity generated a food surplus. Regional differences in state hierarchies are therefore explained by differences in land productivity. We challenge this productivity-and-surplus theory, proposing that differences in appropriability, and not in productivity, account for differences in hierarchy between regions. To test our proposition, we follow the claim that cereals are easier to appropriate than most other crops. Utilizing several data sets, spanning several millennia, we find that the causal effect of the cultivation of cereals on hierarchy is significant, and leaves nothing that can be explained by land productivity per se.

Geography, Hierarchy, Institutions, State Capacity

JEL Classification Numbers: *D02, D82, H10, O43*

*We would like to thank Daron Acemoglu, Alberto Alesina, Yakov Amihud, Josh Angrist, Quamrul Ashraf, David Atkin, Simcha Barkai, Sascha Becker, Davide Cantoni, Ernesto Dal Bo, Carl-Johan Dalgaard, James Fenske, Raquel Fernandez, Oded Galor, Eric Ginder, Paola Giuliano, Stephen Haber, Naomi Hausman, Victor Lavy, Andrea Matranga, Stelios Michalopoulos, Suresh Naidu, Nathan Nunn, Omer Ozak, Joel Slemrod, Fabio Schiantarelli, Uwe Sunde, Jim Robinson, Yona Rubinstein, Felipe Valencia, Tom Vogl, Joachim Voth, John Wallis and David Weil for useful comments. We are also grateful for comments from participants in many seminars and workshops.

†Department of Economics, Hebrew University of Jerusalem. Email: msjoram@huji.ac.il

‡Department of Economics University of Warwick, School of Economics Interdisciplinary Center (IDC) Herzliya, CAGE and CEPR. Email: omer.moav100@gmail.com; Moav's research is supported by the Israel Science Foundation (Grant No. 73/11).

§Eitan Berglas School of Economics, Tel-Aviv University, Email: zvika@post.tau.ac.il

¶Pompeu Fabra University. Email: luigi.pascali@upf.edu

1 Introduction

Following the Neolithic Revolution – the transition of our ancestors from hunting-gathering to sedentary farming – complex hierarchies and states emerged in many regions. But in other regions these changes did not occur, despite the adoption of farming. This raises two related questions: What are the mechanisms that led to the development of complex hierarchies and states following the adoption of farming? And why did some regions develop complex hierarchies but others not?

Ever since Adam Smith and Karl Marx, scholars have attributed the emergence of hierarchy and states to one facet of agriculture: the surplus that was generated from increased productivity. According to this still dominant approach, agricultural surplus was a prerequisite for the rise of an elite class, bureaucrats, troops and other specialists who did not engage in food production for their own subsistence. It is thus differences in land productivity between regions that account for differences in the development of state hierarchy.

More recently, some scholars emphasize another facet of agriculture to account for the emergence of hierarchy: crop appropriability. Taylor (1973) and Testart (1982a) argue that the storage of crops was the key aspect of agriculture that facilitated the emergence of hierarchy. Allen (1997) argues that successful states in the ancient world depended on the ability of elites to extract a surplus from farmers, and that this ability was greatest when the surplus was storable. Scott (2009, 2017) distinguishes between valley and highland farming, between intensive agriculture and dispersed shifting horticulture, and between grains and root crops. He claims that only the former types of farming produce appropriable surplus, and thus facilitate the existence of states, while the latter forms enable societies to escape subjugation to states by avoiding the generation of appropriable surplus.

We provide theoretical arguments and empirical evidence that support the role of appropriability and challenge the conventional productivity and surplus theory. In particular, we contend that surplus is neither necessary nor sufficient for agriculture to give rise to hierarchy. This in itself, though, does not rule out the possibility that increased land productivity may explain hierarchy. However, as it turns out, our extensive empirical evidence reveals that, consistently with the appropriability approach, the cultivation of cereals had a significant causal effect on hierarchy, while land productivity had no such effect.

The most straightforward evidence in support of the proposition that a society's reliance on appropriable sources is necessary for the emergence of advanced hierarchy is provided by the observations that in farming societies that rely on roots and tubers, hierarchical complexity never exceeded the level that anthropologists define as "chiefdoms," and that all agriculture-based large states that we know of relied on cereals. As Scott puts it (2017, p. 21): "It is surely striking that virtually all classical states were based on grain, including millets. History records no cassava states, no sago, yam, taro, plantain, breadfruit, or sweet potato states."[1]

As Scott observes, there are multiple aspects of farming that contribute to appropriability. Here we focus on one of these by distinguishing within the main staple crops between more appropriable cereal grains and less appropriable roots and tubers. Indeed, cereal grains are seasonal, can be stored, and in fact, have to be stored in order to provide year-round nutrition. They are also easy to transport and as a result, are relatively easily appropriated. Roots and tubers, with their higher water content, are typically not seasonal, are bulky, have low energy density and are often perishable, and are thus not easily appropriable. Because cereals are appropriable by both bandits and by would-be elite, their cultivation generates a demand for protection, and, at the same time, also facilitates the supply of elite-provided protection. Specifically, the public-good nature of protection was overcome by the elite's ability to appropriate stored cereals, which would have served in part to cover the costs of providing such services.

Given this, we propose that once the opportunity to appropriate food sources arose, it led to the emergence of hereditary social hierarchy, and eventually to the state. We are not suggesting, however, that early elites were benign. As Olson (1993) observed, deterring bandits served both the farmers and the elite.

In addition to the empirical evidence examined below, we note two key observations that indicate shortcomings in the common productivity-and-surplus theory. First, productivity increased gradually also among hunter-gatherers, due to improved hunting techniques over time. That increase, however, was apparently translated in its entirety to higher population density, as Malthus would have predicted, without generating surplus or hierarchy. Since the Neolithic Revolution

[1] We focus on agriculture-based societies and exclude from our discussion states (like those of the Nabateans, the Venetians, and some African Kingdoms) that relied primarily on taxing trade. In Appendix C we defend the statement above by examining three purported counter-examples in Murdock's (1967) Ethnographic Atlas, where societies that depended on the cultivation of roots or tubers are coded as large states.

spanned several millennia (Purugganan and Fuller, 2010), one could expect that also this gradual increase in productivity would have been absorbed by increased population.[2] Why then was the adoption of agriculture different?

Second, we note the existence of regions that adopted agriculture but did not generate state institutions. In particular, New Guinea adopted agriculture at about the same time as Egypt (ca. 5000-4500 BCE, see Denham, 2011), cultivating bananas, taro and yam. But unlike Egypt, it didn't develop significant hierarchal complexity. Adhering to conventional wisdom, Diamond (1997) explains New Guinea's relative backwardness by its presumed low land productivity. However, as argued here, the distinctive feature of New Guinea's agricultural is not low land productivity, but rather its relatively high productivity in growing non cereal crops (see Appendix A). In fact, New Guinea witnessed a second wave of increased productivity in the 17th century, when sweet potatoes were introduced and rapidly displaced older crops to become the staple. Wiessner and Tumu (1998) record how the resulting substantial increase in crop productivity was transformed into substantial population growth and into prestige goods, such as the aggrandizing slaughter of pigs in communal festivals. But, that second wave of productivity increase left the highland population fractured, subject to endemic tribal warfare, and without any consolidation of power or a significant increase in social complexity. So why did New Guinea follow a very different path than Egypt's?

Both questions expose, we contend, serious shortcomings in the productivity and surplus theory, yet are easily explained by the appropriability theory. Hunter gatherers relied on hand to mouth food sources which are not easily appropriable, and, therefore, didn't develop hierarchies even after their productivity increased.[3] Similarly, we attribute New Guinea's low level of social complexity to its reliance on crops for which long-term storage is neither feasible (due to perishability) nor necessary (due to non-seasonality). Farmers' ability to cultivate highly productive crops that are by and large non-appropriable inhibits both the demand for socially-provided protection from bandits, and the emergence of protection-providing, yet appropriating, elite. To the extent that hierarchy is a prerequisite for advanced civilization, New Guinea's geography is thus a curse of plenty.

[2] Ashraf and Galor (2011) support the applicability of Malthus's theory by demonstrating that technological improvements before the Industrial Revolution had a positive effect on population size but no effect on long run per capita income.

[3] In the next section we consider the case of some hunting-gathering societies that anthropologists identify as "complex," and argue that this case is consistent with our proposed theory.

To illustrate our contention against the necessity or sufficiency of surplus, consider the following two contrasting scenarios. First, consider an early society that cultivates cassava with output above long-term subsistence (that is, with 'surplus'). Cassava is a perennial root that can be harvested year-round, yet rots shortly after harvest. This makes it difficult to confiscate, and practically impossible to transport for consumption by a distant elite.[4] It is thus unlikely that a complex hierarchy could emerge in this society, despite the availability of food surplus. Now consider another farming society that grows a cereal grain with no surplus: each family's (average) annual production equals its long-run subsistence needs. Since the crop has to be harvested within a short period and then stored, a tax collector, as well as bandits, could confiscate part of the crop. The durability of grains and their high caloric density also enables the transportation of grains for consumption by distant elite. Ongoing confiscation of food in this subsistence society can be expected to impact adversely the size of its farming population. Due to diminishing average product of labor, however, this would result in an equilibrium where total output exceeds the farming population's subsistence needs, and with the resulting surplus confiscated by the non-farming elite.[5]

The first of these scenarios and the New Guinea evidence imply that, had the Neolithic Revolution amounted to a transition to the cultivation of roots and tubers, the increase in productivity would not have resulted in the emergence of advanced hierarchy, and could have been absorbed by increased population. The second scenario demonstrates that, in contrast to conventional theory, the availability of surplus is not a precondition for taxation and hierarchy. Rather than surplus generating the elite, it is the elite that generate the food surplus on which it can flourish – once the opportunity to appropriate rises.

In our formal model, a fixed number of farmers can allocate their land between two crops, which we label cereals and tubers.[6] The productivity of the two crops is presumed to differ across

[4] In Appendix A we describe the characteristics of cassava and of other roots and tubers. The portability of these crops is hampered not only by their vulnerability to spoilage, but also by their bulkiness (due to ca. 70% moisture content). We also support there our various claims: (i) that reliance on roots and tubers is a major phenomenon in tropical regions; (ii) that roots and tubers are highly productive in the tropics; (iii) that their harvest is in general non-seasonal; and (iv) that after harvest they are significantly more perishable than cereals.

[5] The scenario of reduced population pertains to the case where no bandits exist. However, since stored grains are vulnerable to predation by bandits, if the government deters bandits, the population size with government is likely to be larger than without government – see the case of endogenous population in Appendix B.

[6] To focus on the empirical predictions, and for simplicity, we do not consider the decision whether to farm or to forage. For brevity, we often refer to tubers only, even when implying roots as well. What we are really after is a distinction among food types according to their degree of appropriability. We are aware that potatoes, for example, have been freeze dried in ancient Peru, and are altogether somewhat storable. We also ignore other important food

geographic locations. We assume that it is possible to tax cereals at some cost, but, for simplicity, that it is impossible to confiscate or tax tubers. As a result, cereals are cultivated only if their productivity advantage over tubers is sufficiently high to compensate for taxation by the state, or for the expected loss to bandits.

We distinguish between two regimes. In 'anarchy,' non-farmers can be foragers, where their average and marginal productivity is fixed, or bandits. Since we assume that bandits are unorganized, their number is determined endogenously, so that the average revenue from theft is equal to the alternative productivity in foraging. The probability that any farmer's cereals would be stolen is a function of the number of bandits. In the alternative 'hierarchy' regime, crops are taxed by a net revenue-maximizing elite that provides protection from bandits. We assume that to be viable and deter bandits, the state has to incur some fixed cost. Finally, we assume that the ruling elite can commit to its selected rate of taxation. To maximize the net tax revenue, the state thus employs tax collectors so that their marginal tax revenue is higher or equal to their wage – which equals the alternative income from foraging. Under hierarchy, the state will therefore employ (weakly) less tax collectors than the equilibrium number of bandits under anarchy.

The main prediction of the model is that a state cannot exist if tubers are sufficiently productive, since the potential tax revenue in that case is insufficient to cover the fixed cost of forming a government. This result illustrates the claim that it is relatively high productivity of the more appropriable cereals, rather than high agricultural productivity per se, which facilitates the development of hierarchy. The model also suggests that even though the elite is self-serving, whenever hierarchy exists it dominates anarchy in welfare terms. As a result, cereal based farming, which renders farmers vulnerable to taxation, leads not only to the development of a state, but also contributes to farmers' welfare. Anarchy is more distortionary than hierarchy for two reasons. First, the state's ability to commit to a lower tax rate encourages the cultivation of cereals when these are more productive. Second, since the state employs fewer tax collectors than the equilibrium number of bandits under anarchy, the forgone output in foraging is higher under anarchy.[7]

To establish empirically the advantage of the appropriability approach over the land-productivity

sources such as legumes, fruits, vegetables and livestock. To the extent that these are not easily appropriable, we would group them with tubers.

[7] We ignore the possibility that the non-benevolent state may contribute further to farmers' welfare if it contributes directly to agricultural productivity, for example through publicly provided irrigation infrastructure.

approach, we employ four different datasets that provide measures of hierarchy and statehood: a cross-section of pre-colonial societies, a panel of countries, a cross-section of archeological sites, and a panel of prehistorical archeological sites.

For our first cross section analysis, we use Murdock's (1967) Ethnographic Atlas on cultural, institutional and economic features of more than 1,200 pre-colonial societies from around the world. Our main outcome variable is the level of hierarchical complexity, which we relate to the major crop farmed in these societies. Since crop type might depend on hierarchy, we instrument for the cultivation of cereals by the potential productivity advantage of cereals over roots and tubers, calculated from the land suitability data provided by the Food and Agriculture Organization (FAO), under a rain-fed subsistence economy. As the productivity theory would predict, we find that land productivity by itself has significant positive effect on hierarchy. However, consistent with our critique of that theory, once we control for crop type this positive effect disappears. On the other hand, cultivating cereals has a considerable positive impact on hierarchical complexity, consistent with the appropriability theory. As it turns out, societies that cultivate roots or tubers have a similar level of hierarchy to non-farming (pastoral or foraging) societies. Data covering a subset of these societies provide information on the sources of political power and income of the politically dominant class (Tuden and Marshall, 1972). We find that farming surplus, evidenced by elite members deriving their income not from their own subsistence activities, is significantly higher in societies that grow cereals.

For our first panel data analysis we employ a dataset compiled by Borcan, Olsson and Putterman (2017). This dataset is based on present-day boundaries of 159 countries, with institutional information every half-century. Using the data describing the last millennium, we exploit the "Columbian Exchange" as a natural experiment. The new crops that became available as a result of the transfer of crops between the New and the Old World changed both land productivity and the productivity advantage of cereals over roots and tubers in most of the countries in our sample. Consistent with the appropriability theory and with our critique of the productivity theory, the panel regressions confirm that the productivity advantage of cereals over roots and tubers has a positive impact on hierarchical complexity, while land productivity does not.

Our third data source for indicators of social hierarchy consists of information on the location of ancient cities and other archeological sites. With this data, we employ two different cross-sectional approaches to test the appropriability and productivity theories. First, using the FAO-based data, we document that the cultivation of cereals, unlike land productivity, can explain the distribution of cities founded before AD 400 and the location of archaeological sites that testify to the presence of complex societies (e.g. pyramids, ancient temples, palaces, and mines). As an alternative to the FAO-based measures, we utilize data on the location of centers of origin of agriculture. We show that distance from these centers has a negative impact on the development of early civilization only when these were centers which domesticated cereals.

In our fourth empirical section we use data from the Archaeological Atlas of the World (Whitehouse and Whitehouse, 1975), which provides an alternative source of information on the location of ancient archaeological sites. This source reports also a radiocarbon dating of these sites, thus enabling us to measure the number of pre-Neolithic and post-Neolithic ancient cites in each area. Using difference-in-difference estimates for the impact of the introduction of crop types, we obtain once again support for the appropriability theory over the productivity theory, by employing either the two FAO-based measures, or the measures of distance to centers of domestication.

We acknowledge at the outset that our empirical tests cannot prove beyond doubt the validity of the appropriability theory, nor refute entirely the applicability of the conventional productivity-and-surplus theory. We discuss potential concerns regarding the data and the identification in the empirical section. Nevertheless, we contend that the robustness of our empirical results suffice to call into question the prevailing productivity-and surplus theory and to provide credence to the proposed alternative explanation.

The proposed appropriability mechanism, it should be emphasized, applies beyond antiquity. Since the modern transition away from agriculture is protracted and social institutions exhibit significant inertia, this theory may contribute to explaining current institutional differences. In particular, the identification of the cultivation of roots and tubers as a hindrance to the development of social hierarchy may contribute towards understanding the root causes for the underdevelopment of tropical regions, which are known to suffer from poorly functioning governments.[8]

Since the literature on the rise of hierarchy and states is extensive, we delay the literature review to section 5, after we present our model and evidence, so that we can better compare the alternative

[8] Bockstette, Chanda and Putterman (2002), Gennaioli and Rainer (2007), Spolaore and Wacziarg (2013), and Michalopoulos and Papaioannou (2013, 2014) demonstrate that deep rooted pre-colonial institutions affect current institutions and economic outcomes.

approaches. There we illustrate that the productivity-and-surplus theory, which we challenge, is indeed widely accepted. In section 2 we examine anthropological and archaeological evidence. We present the model in section 3 and our empirical results in section 4. Section 6 concludes.

2 Some Archeological and Anthropological Evidence

Anthropologists and archaeologists have long concluded that hunter-gatherer societies were fairly egalitarian and ostensibly leaderless. This led them to ponder the mechanisms that mediate the transition to agriculture and the emergence of hierarchy. In this section we review anthropological and archeological evidence and theories about the role of storage and appropriability.

The appropriability theory was anticipated by Taylor (1973) and Testart (1982a, 1982b). Taylor, in a neglected brief contribution, argues that the cultivation of grains forced the incipient agriculturalist into sedentism and to "social control that would assure the horde to its rightful owner." Holding that it thus contributed to "the very foundation of civilization," he suggests that the Neolithic Revolution ought to be called the "storage revolution." Summarizing extensive anthropological evidence on hunting-gathering societies, Testart (1982a, 1982b) identifies a positive association between social inequality and the prevalence of storage of seasonal food sources. When considering agricultural societies, he distinguishes between those based on cereals and those based on tubers, and attributes inegalitarian, complex social structures mostly to reliance on sedentism and storage (1982a, pp. 195-204).[9]

Tushingham and Bettinger (2013) study the transition of hunting-foraging aboriginal Californians from a long period of reliance on stored acorns to intensified reliance on drying and storing salmon. They theorize that in spite of its many advantages, reliance on salmon was initially avoided because it is a "front loaded" food source that requires much effort in procuring but little effort in preparing for consumption. Such a food source, they state, increases "the possibility that others will rob caches, which mobile foragers are not positioned to protect," and also increases the vulnerability of loss to "freeloaders" from the inside (pp. 533-534). They note that the transition to salmon coincided with increased complexity in the form of population concentration in permanent

[9] Testart (1982 c) offers a crude attempt to establish this association empirically on the basis of Murdock's Ethnographic Atlas.

villages.[10] While they do not mention hierarchy or agriculture, their analysis is perfectly consistent with the idea that increased vulnerability to appropriation contributes to social hierarchy.[11]

Indeed, diverse archaeological findings support the idea that the earliest phases of the transition to cereal farming in the Near East were correlated with communal storage and with the emergence of inequality and hierarchy. Semi-sedentary forms of living, dwellings, sickles, grinding stones, pestles and mortars, and storage facilities appear already in the pre-Neolithic, Natufian period, when cereals were collected but not yet sown. Active cereal cultivation emerged during the period known as Pre-Pottery Neolithic A (PPNA, ca. 9500-8500 BCE), when farmer-foragers collected wild grains on a large scale and also sowed (still prior to domestication).[12] Differentiated dwelling sizes and funerary assemblages reveal the appearance at that time of systematic inheritable inequality. Kuijt and Finlayson (2009) report the discovery of an elaborate, large communal storage pit in the Jordan Valley from about 9000 BCE. This finding reveals that storage was an integral part of the earliest phase of the transition to cereal farming. It also indicates organized social cooperation, attesting to the emergence of a form of hierarchy alongside the gradual intensification of cereal farming and its concomitant storage and sedentism.[13] The need to protect stockpiles is manifest also by the subsequent formation of early walled villages and urban centers, long before the formation of city states.[14]

[10] Cook's account of his voyages to the eastern shores of the Pacific Ocean (1784, volume II, book IV) provides a vivid eye-witness depiction of these villages.

[11] Chiwona-Karltun et al. (2002) provide an anecdotal illustration of the appropriability theory in a farming setting. They report that women in modern Malawi, and particularly single women, prefer to grow bitter and toxic cassava variants, even though these variants require significantly more post-harvest processing. This pattern is explained as due mostly to the advantages of this extra post-harvest drudgery, which protects these women against thievery, since thieves prefer the non-bitter variant. A Malawian woman is quoted: "We grow bitter, toxic cassava because it gives a certain level of food security. If we are to grow sweet cassava, look at our neighbors! Their whole field was harvested by thieves while they slept and now they have no food."

[12] The timing and location of the earliest transition to agriculture in the Fertile Crescent are commonly explained in that the end of the ice age and increased climatic seasonality generated evolutionary modifications in grasses that adapted to extended summer drought by developing larger seeds (Diamond, 1997). Also the onset of the Neolithic period is often attributed to climatic change (Bar-Yosef and Meadow, 1995; Matranga, 2017). An alternative explanation contends that it was food shortage due to population growth that led hunter-gatherers to engage in agriculture. Richerson, Boyd and Bettinger (2001, pp. 388-389) debunk this theory by employing a similar Malthusian argument to the one raised here against the idea that hierarchy emerged due to population pressure following the adoption of agriculture.

[13] Somewhat similar large round pits from PPNA were found elsewhere in the Jordan Valley and in several sites near the Euphrates (Mithen et al. 2011; Willcox and Strodeur 2012). These pits are identified as having served both for communal storage and as communal meeting places, possibly for ritual ceremonies. Some archaeologists identify storage as an indication of surplus, but a cereal-based farming society may be living at subsistence while engaging in intra-annual and even inter-annual storage.

[14] By focusing on the formation of states in southern Mesopotamia in the late fourth millennium BCE, Scott (2017),

3 A Model of Cereals and Hierarchy

The basic premise of the model is that regions of the world differ in their productivity of tubers relative to that of cereals. For simplicity we postulate that tubers, unlike cereals, cannot be expropriated by bandits or by tax collectors. We model farmers' choice of what crop to grow in two different regimes: anarchy and hierarchy, and derive conclusions on the circumstances that are conducive for the emergence of hierarchy.

The economy is populated by a measure one of farmers and a measure N of non-farmers. Our main exogenous variable, $\delta > 0$, measures the productivity advantage of cereals over tubers.[15] The productivity of cereals is normalized to unity. Thus, farmers can grow one unit of cereals, or $1 - \delta$ units of tubers, or any linear combination thereof. Hence, a farmer's output is $\beta + (1 - \beta)(1 - \delta) = (1 - \delta) + \beta\delta$, where $\beta \in [0, 1]$ is the fraction of land allocated to cereals. Output is measured in nutritional units independently of their source.

The income of non-farmers who engage in foraging is assumed to be constant and denoted: $s > 0$. In a state of anarchy, non-farmers can chose to be either foragers or bandits who expropriate crops from farmers. In a state of hierarchy, we assume that some non-farmers are hired by the state to serve as tax collectors, and are paid the wage s. We denote by λ the measure of bandits or tax collectors. N is assumed to be larger than the equilibrium level of λ. We also assume that in order to be viable and to deter bandits, the state has to possess monopoly over the use of force which entails a cost.

3.1 Anarchy

Under anarchy, farmers might be raided by bandits. A raided farmer loses his entire cereal crop, but none of his crop of tubers. Farmers are assumed to be risk neutral.[16] A farmer who faces a raid with probability τ, chooses the fraction of land allocated to cereal β to maximize his expected income I, weighing the productivity advantages of cereals over tubers, δ, against the disadvantage, as measured by the expropriation rate τ:

who emphasizes (like us) the role of cereal cultivation, misses the significance of the vulnerability to pillage already at much earlier, pre-states phases of prehistory.

[15] If $\delta \leq 0$ the analysis is trivial: tubers dominate cereals in providing both protection and higher productivity, so that farmers would only grow tubers in equilibrium, and the economy could only be in a state of anarchy.

[16] In Appendix D we show that our results are robust to the introduction of risk aversion.

$$I = (1-\tau)\,\beta + (1-\delta)\,(1-\beta) = (1-\delta) + \beta(\delta - \tau). \tag{1}$$

We assume that the rate of expropriation, τ, is a function of the measure of bandits λ: $\tau = \tau(\lambda)$. The function, $\tau(\lambda)$ is strictly increasing and strictly concave, and satisfies: $\tau(0) = 0$, $\lim_{\lambda \searrow 0} \tau'(\lambda) = \infty$, $\lim_{\lambda \nearrow \infty} \tau'(\lambda) = 0$ with $\lim_{\lambda \nearrow \infty} \tau(\lambda) = \overline{\tau} \leq 1$, with corresponding properties for the inverse function $\lambda(\tau)$. Bandits are uncoordinated (if they coordinate they become the ruling elite and the equilibrium becomes identical to that under hierarchy). Thus each bandit's expected income π is given by the total amount of confiscated cereals divided by the measure of bandits: $\pi = \tau(\lambda)\beta/\lambda$.

Definition. *Equilibrium consists of a pair (β, τ) such that:*

1. *β maximizes farmers' income I, given the confiscation rate τ;*
2. *given β, non-farmers are indifferent between being foragers or bandits, so that $\pi = s$.*[17]

Using the inverse relation $\lambda(\tau)$, the last condition can be restated as requiring: $\tau\beta/\lambda(\tau) = s$.

Define now a threshold rate δ_A by the implicit relationship:[18]

$$\frac{\delta_A}{\lambda(\delta_A)} = s.$$

Proposition 1. *The economy under anarchy has a unique equilibrium (β_A, τ_A) that is given by:*

$$(\beta_A, \tau_A) = \begin{cases} \left(\frac{\lambda(\delta)s}{\delta}, \delta\right) & if \quad \delta < \delta_A \\ \\ (1, \delta_A) & if \quad \delta \geq \delta_A \end{cases}.$$

That is, the equilibrium confiscation rate τ_A equal δ if tubers are grown, and tubers are not grown if $\delta > \delta_A$.

Proof. If $\delta > 0$, an equilibrium with no cereals ($\beta_A = 0$) can be ruled out since in that case $\pi = 0$, leading to $\lambda = 0$ and $\tau = 0$, which would lead to $\beta = 1$, a contradiction. This implies that the

[17] Micro-foundation for the shape of $\tau(\lambda)$ can be obtained by assuming that banditry is time consuming and that bandits are not coordinated, so that when their number increase, the probability of raiding the same farmers increases, and the marginal total loot declines.

[18] We use the subscript A to denote parameters and equilibrium values in a regime of anarchy, and similarly use the subscript H in a state of hierarchy.

equilibrium can either be one with cereals only ($\beta_A = 1$) or mixed ($0 < \beta_A < 1$), where both crops are cultivated.

If $\delta \geq \delta_A$, farmers cultivate only cereals ($\beta_A = 1$), even though this entails a maximal confiscation rate $\tau_A = \delta_A$ and a corresponding maximal number of bandits, $\lambda(\delta_A)$.

Our assumptions on $\tau(\cdot)$ imply that the confiscation rate, $\tau(\lambda)/\lambda$, is monotonically decreasing in τ, from infinity towards zero. Thus, when $\delta < \delta_A$, both cereals and tubers are cultivated and we have: $\delta/\lambda(\delta) > \delta_A/\lambda(\delta_A) = s$. Hence, there exists a unique $\beta_A \in (0, 1]$ such that $\pi_A \equiv \delta\beta_A/\lambda(\delta) = s$. The last condition, in conjunction with the condition $\tau_A = \delta$, defines the equilibrium combination (β_A, τ_A). ■

Income distribution. It follows from Proposition 1 that if cereals' productivity advantage is low ($\delta < \delta_A$), β_A, τ_A and $\lambda_A = \lambda(\tau_A)$ are strictly increasing in δ and all tend to zero when δ decreases towards zero. As a result, also the total expected amount of cereals confiscated by bandits, $\tau_A\beta_A$, strictly increase in δ. In that range, farmers' income equals $1 - \delta$, and thus decreases in δ. On the other hand, when the productivity advantage of cereals exceeds the threshold δ_A, all these variables become independent of δ, with farmers income equaling $1 - \delta_A$. In these two ranges combined, proposition 1 thus implies that $\tau_A\beta_A$, τ_A and λ_A all weakly increase in δ. In turn, even though bandits' welfare is equal to s independently of the value of δ, farmers' welfare weakly decreases with δ.

The effect of the reservation income s. The smaller is s, the larger the incentive for foragers to engage in banditry. This implies a higher threshold δ_A, meaning that farmers will raise tubers in a wider range of δ. Thus, for values of $\delta > \delta_A$, a lower s reduces farmers' income. However, for $\delta < \delta_A$, a smaller s has no effect on farmers income, or on τ and therefore on λ; it will reduce however the equilibrium value of β.

Two sources of inefficiency. Denote by Y_0 the maximal possible level of output in the economy, when all farmers cultivate only the more productive cereals (assuming $\delta > 0$) and all non-farmers engage in foraging. This maximal output level is: $Y_0 = 1 + Ns$. We can observe that the equilibrium (β_A, τ_A) introduces two deviations from this maximal output level: the first is due to farmers growing tubers (if $\delta < \delta_A$); and the other is due to the forgone output by banditry. Thus, equilibrium output is:

$$Y = Y_0 - (1 - \beta_A)\,\delta - s\lambda(\tau_A).$$

Inspection of the equilibrium values (β_A, τ_A) reveals that for large values of δ, the only distortion is the loss of output due to bandits being unproductive: $s\lambda_A = s\lambda(\tau_A)$, which equals the threshold level δ_A. For small values of δ, tubers are cultivated, $\tau_A = \delta$, and farmers are indifferent between the two crops. From the fact that expected revenue per-bandit is equal to $\tau_A\beta_A/\lambda(\tau_A) = s$ it follows that $s\lambda(\tau_A) = \tau_A\beta_A$. Thus we obtain:

Corollary 1. *The output loss $(Y_0 - Y)$ due to an anarchy regime is:*

$$(1 - \beta_A)\,\delta + \lambda_A s = \begin{cases} \delta & if \quad \delta < \delta_A \\ \\ \delta_A & if \quad \delta \geq \delta_A \end{cases}.$$

3.2 Hierarchy

We assume that in a state of hierarchy the elite (the state) chooses its tax policy to maximize the revenue net of the cost of tax collection. In order to facilitate comparison between the two regimes, we assume that the state has access to the same expropriation technology as bandits. Namely, the state cannot tax tubers, and if it employs a measure λ of tax collectors (hired from among the potential foragers) at cost s per tax collector, it can generate revenue of $\tau(\lambda)\beta$ from the farming sector. Adopting Weber's definition of the state, we also assume that a state has to have a monopoly power over the use of force, and thus be able to deter bandits. This deterrence power, we assume, entails a fixed cost $G_0 > 0$.

A key advantage that a state has is that it is farsighted and organized, and can thus commit not to expropriate farmers beyond a certain tax rate.[19] That is, the state selects a tax rate τ, and hires $\lambda(\tau)$ tax-collectors at cost $s\lambda(\tau)$, to maximize its net revenue, subject to the constraint that farmers respond to the tax rate:

[19] Another difference between bandits and the state is that bandits confiscate a farmer's entire cereal crop with probability τ, while an organized hierarchy taxes farmers at the rate τ with certainty. If farmers are risk neutral, as assumed here, this difference is unimportant.

$$\max_{\tau \geq 0} R(\tau) = \tau\beta - s\lambda(\tau),$$

subject to

$$\beta = \arg \max_{\beta' \in [0,1]} \left\{ (1-\delta) + \beta'(\delta - \tau) \right\}.$$

It is evident that $\beta = 0$ if $\tau > \delta$. This implies that in the optimum $\tau \leq \delta$. In addition $\beta = 1$ if $\tau < \delta$. Assuming that $\beta = 1$ when $\tau = \delta$, the state's problem becomes:

$$\max_{\tau} [\tau - s\lambda(\tau)], \text{ subject to } \tau \leq \delta.$$

The optimal tax rate under hierarchy is therefore: $\tau_H(\delta) = \min\{\delta, \delta_H\}$, where δ_H is the parameter that solves $s\lambda'(\delta_H) = 1$. Thus, at the high range of tubers' productivity, where $\delta < \delta_H$, $\tau_H = \delta$, the net revenue is $R(\tau_H(\delta)) = \delta - s\lambda(\delta)$, which is increasing in δ. Our assumption that the state has to sustain deterrence against bandits at a fixed cost $G_0 > 0$ sets a lower limit on the net revenue for the state to be viable. Clearly if G_0 exceeds the maximal revenue $R(\delta_H)$, the state is not viable. We thus assume that $R(\delta_H) > G_0$. Under this assumption we can define a viability threshold $\underline{\delta} < \delta_H$, such that: $R(\underline{\delta}) = G_0$, so that he state is viable when $\delta > \underline{\delta}$.

Proposition 2. (i) *If δ is small ($\delta < \underline{\delta}$), a state cannot exist. (ii) In the intermediate range where $\underline{\delta} \leq \delta < \delta_H$, the optimal tax rate is $\tau_H = \delta$. (iii) If δ is large ($\delta \geq \delta_H$), then the optimal tax rate is equal to δ_H.*

Income distribution. Under hierarchy, farmers grow only cereals. Thus, their income is $1 - \tau_H = 1 - \min\{\delta, \delta_H\}$, which is weakly decreasing in the cereal productivity advantage over tubers δ. Total tax receipts equals τ_H, and the net tax revenue to the elite is: $\tau_H - s\lambda(\tau_H) - G_0$. Both increase in δ up to the threshold δ_H, after which they remain constant.

Output Loss. Analogously to the case of anarchy, we obtain here:

$Y_0 - Y = (1 - \beta_H)\delta + s\lambda(\tau_H) + G_0$ and since $\beta_H = 1$, $Y_0 - Y = s\lambda(\tau_H) + G_0$. Thus:

Corollary 2. The output loss due to hierarchy is:

$$Y_0 - Y = \begin{cases} s\lambda(\delta) + G_0 & if \quad \delta < \delta_H \\ & \\ s\lambda(\delta_H) + G_0 & if \quad \delta \geq \delta_H \end{cases}.$$

.

3.3 Anarchy vs. Hierarchy

As explained in the previous sub-section, a state can only exist if tubers are sufficiently unattractive to farmers, that is, if the productivity advantage of cereals, δ, is above the threshold $\underline{\delta}$. The comparison between the regimes of anarchy and hierarchy depends on the relationship between the thresholds δ_A, δ_H and $\underline{\delta}$.

These threshold levels satisfy: $\delta_A > \delta_H > \underline{\delta}$. Since δ_H is defined by $\tau'(\lambda(\delta_H)) = s$, and δ_A is defined by $\delta_A/\lambda(\delta_A) = s$, $\delta_A > \delta_H$ follows from the strict concavity of $\tau(\lambda)$.

Proposition 3. *For $\delta > \underline{\delta}$ the state is viable and:*

(i) Hierarchy weakly Pareto dominates anarchy. (ii) The economy is more productive under hierarchy than under anarchy.

Proof. (i) Because the function $\tau(\cdot)$ is strictly concave, the marginal productivity of tax collectors (or bandits) is lower than the average productivity: $\tau'(\lambda) < \tau(\lambda)/\lambda$ and $\tau'(\lambda(\tau)) < \tau/\lambda(\tau)$. Recall that, $\lambda(\delta_H)$ is defined by $\tau'(\lambda(\delta_H)) = s$ and $\lambda(\delta_A)$ is defined by $\delta_A/\lambda(\delta_A) = s$. It therefore follows from the concavity of $\tau(\cdot)$ that $\delta_H < \delta_A$ and $\lambda(\delta_H) < \lambda(\delta_A)$.

Non-farmers earn the same income s irrespective of the regime. Suppose that $\delta > \underline{\delta}$. On the other hand, the implied tax rate on farmers under anarchy is larger than or equal than the tax rate under hierarchy. In the range where $\underline{\delta} \leq \delta \leq \delta_H$, the tax rate under both anarchy and hierarchy is δ; in the range $\delta_H \leq \delta < \delta_A$ the tax rate under anarchy δ is higher than the tax rate under hierarchy δ_H and in the range $\delta_A \leq \delta$ the tax rate under anarchy is δ_A, whereas under hierarchy it is lower δ_H. Hence, farmers are weakly better off in all cases under hierarchy than under anarchy. Finally, when $\delta > \underline{\delta}$, a hierarchy generates an additional surplus to the elite, since by construction: $\tau - s\lambda(\tau) - G_0 > 0$. ■

(ii) From corollaries 1 and 2 we obtain that the difference between total output under hierarchy to that under anarchy is equal to:

$$Y_H(\delta) - Y_A(\delta) = \begin{cases} \delta - s\lambda(\delta) - G_0 & if \quad \delta \in [\underline{\delta}, \delta_H] \\ \delta - s\lambda(\delta_H) - G_0 & if \quad \delta \in (\delta_H, \delta_A] \\ \delta_A - s\lambda(\delta_H) - G_0 & if \quad \delta > \delta_A \end{cases} .$$

When $\delta = \underline{\delta}$, by the definition of $\underline{\delta}$, $R(\underline{\delta}) = \underline{\delta} - s\lambda(\underline{\delta}) = G_0$, the output gap between the two regimes is zero . When $\underline{\delta} \leq \delta \leq \delta_A$, the output gap equals the rent enjoyed by the elite, which is increasing in δ. ■

The total output under hierarchy is weakly higher for two reasons. (1) Under hierarchy farmers cultivate only cereals. Thus they do not resort to self-protection through the cultivation of the less productive tubers, as they do (when $\delta < \delta_A$) under anarchy. (2) The state employs (weakly) fewer tax collectors than the scale of foragers who engage in banditry under anarchy, since with the state their marginal product is higher or equals their cost s, whereas under anarchy it is their average product that equal s.

The main predictions of the analysis

1. Farmers may choose to grow tubers even when tubers are less productive for the purpose of self-protection against appropriation by bandits or by tax collectors.

2. If the productivity advantage of cereals over tubers is sufficiently small, $\delta < \underline{\delta}$, a state cannot exist. This result illustrates our claim that it isn't low productivity that hinders the development of hierarchy and related institutions, but rather relatively high productivity of crops that are hard to expropriate. A hierarchy could emerge if the productivity advantage of cereals is sufficiently high: $\delta > \underline{\delta}$.

3. Whenever it exists, even a non-benevolent state hierarchy that monopolizes coercive force dominates anarchy from an efficiency point of view. This follows from our assumption that the state can commit to a tax rate that maximizes its net revenue, and that consequently farmers cultivate only the more efficient cereals.

3.4 Example

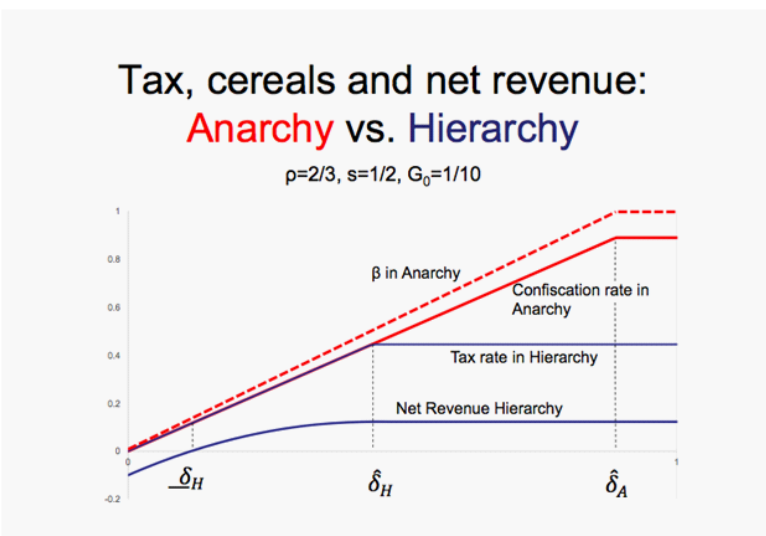
A simple example helps to illustrate our results diagrammatically. Consider the following specification for the expropriation function: $\tau(\lambda) = \rho\sqrt{\lambda}$, with $\rho \in (0,1)$.

In this case, $\delta_A = \rho^2/s$ and $\delta_H = \alpha^2/2s$.[20] The equilibrium under anarchy is given by

$$(\beta_A, \tau_A) = \begin{cases} \left(\frac{s\delta}{\rho^2}, \delta\right) & if \quad \delta < \delta_A \\ \\ \left(1, \frac{\rho^2}{s}\right) & if \quad \delta \geq \delta_A \end{cases}.$$

For $\underline{\delta} \leq \delta \leq \delta_H$ a state sets a tax rate equal to δ and generates net tax revenue: $R(\delta) = \delta - s\left(\frac{\delta}{\rho}\right)^2$, which increases in δ up to the point where $\delta = \delta_H$ after which $R(\delta) = R(\delta_H)$. Figure 1 presents the comparison between anarchy and hierarchy with respect to the tax rate and the production of cereals, as a function of δ. It also presents the net revenue of the elite in a regime of hierarchy.

Figure 1: Tax, cereals and net revenue: Anarchy vs. Hierarchy



[20] The lower limit for state existence, $\underline{\delta} > 0$, is implicitly defined by the quadratic equation: $\underline{\delta} - s\left(\frac{\delta}{\rho}\right)^2 = G_0$, where to have any solution we require that : $G_0 \leq \rho^2/4s$.

4 Evidence

In this section we provide supportive evidence for our main theoretical predictions. We employ four datasets that provide measures of hierarchy and statehood: a cross section of pre-colonial societies, a panel of countries, a cross section of ancient cities and archeological sites, and a panel of radiocarbon-dated prehistorical archaeological sites. Our main regressors are two measures of potential agricultural productivity: the productivity of the soil and the productivity advantage of cereals over roots and tubers – a measure corresponding to δ in our model. Consistently with the main prediction of our theory, our empirical investigation shows that it isn't low agricultural productivity that retards development of hierarchy, but rather high productivity of less appropriable crops. In combining agricultural productivity data with measures of hierarchy and statehood data, we follow a similar strategy that is employed by a growing recent literature, including: Alesina et al. (2013), Fenske (2013), Galor and Ozak (2016), and Nunn and Qian (2011).

4.1 Data

4.1.1 Ethnographic data

Murdock's (1967) Ethnographic Atlas provides a database of 1,267 societies from around the world. The database contains information on several cultural, institutional and economic features for these societies at an idealized moment of first contact with Europeans. From this sample, we remove 2 duplicate observations, 7 societies observed before AD 1500, and 10 societies for which the year of observation is missing, so that we are left with a total of 1,248 societies. These are matched to ethnic maps using either the geo-coordinates of each ethnicity provided by the Ethnoatlas or the maps on the spatial location of ethnicities constructed by Fenske (2013).[21]

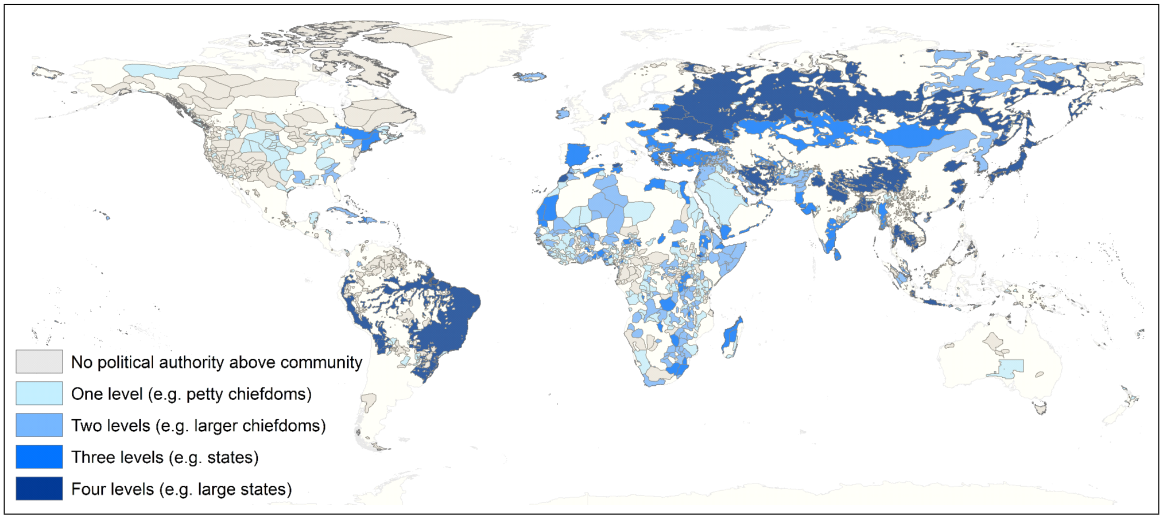
We measure pre-colonial hierarchical complexity using the variable "Jurisdictional Hierarchy beyond the Local Community."[22] This is an ordered variable with five possible levels: (i) no political authority beyond community, (ii) petty chiefdoms, (iii) larger chiefdoms, (iv) states, and (v) large states. We plot this measure of hierarchy in Figure 2 and present the summary statistics in the

[21] The ethnic maps in Fenske (2013) are constructed by combining Murdock's (1959) ethno-linguistic map for Africa with three other sources for the rest of the world (Heizer and Sturtevant, 1978; Global Mapping International, and Weidmann et al., 2010).

[22] Gennaioli and Reiner (2007) and Michaelopoulos and Papaioannou (2013) make a similar use of this variable.

first row of Table E.1 in the online appendix. The majority of our sample is composed of societies lacking any political integration above the local community, and groups where chiefs rule over very small districts. These societies prevail in North America, Australia and in Central Africa, but are rather rare in Northern Africa and in Asia, where large chiefdoms and states are more common.

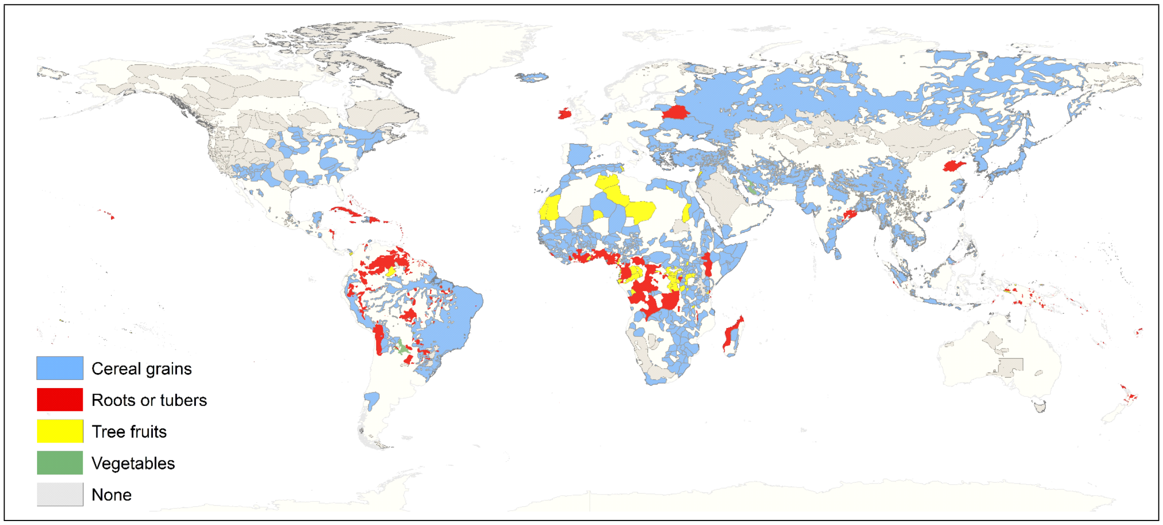
Figure 2: Jurisdictional hierarchy beyond the local community in pre-colonial societies



The Ethnoatlas also provides information on the reliance of these societies on agriculture for their diet, and on the major crop type of societies that practice agriculture. These two variables are plotted in Figure 3, with summary data in rows 2 and 3 of Table E.1 in the online appendix. As can be seen from Figure 5, approximately one fifth of the societies in the sample do not practice any form of agriculture. These societies are concentrated in North-West America, Central Asia, Australia and South-West Africa. The median society relies on agriculture for approximately 50 percent of its caloric needs. The great majority of the societies that practice some form of agriculture rely on either cereal grains (65.4 percent) or on roots and tubers (26.1 percent). The latter are concentrated in the tropics, while the former are scattered all over the world.[23] Using this information, we define a dummy that identifies societies whose primary crop is cereals and present summary statistics on the second row of Table E.1. Finally, the Ethnoatlas provides information on whether agriculture in farming societies was relying on intensive irrigation (row 4 in Table E.1).

[23] Some societies in the temperate zones grow potatoes - a tuber crop that is in fact similar in its relevant properties to cereals in that it is seasonal and storable.

Figure 3: Major crop in pre-colonial societies



The second source of ethnographic information is provided by the Standard Cross-Cultural Sample (SCCS), which is a derivative of the Ethnographic Atlas. The data are based on a representative sample, defined by Murdock and White (1969), of 186 societies taken from the Ethnoatlas. A large number of publications by diverse authors coded the SCCS societies for many different characteristics. Cumulative ethnographic codes and codebooks are published in the World Cultures electronic journal.

We use two variables from the SCCS (rows 5 and 6 in Table E.1). The first one, coded by Tuden and Marshall (1972), lists the sources of political power to the local elite. We create a dummy on "the existence of a farming surplus" that is equal to zero if the most prestigious members of the society derive their livelihood from their own subsistence activities and one otherwise. This dummy is plotted in figure E.1 in the appendix. The second variable is a measure of population density coded by Pryor (1985). Societies are categorized into 6 bins (the first bin contains societies with 0-1 persons per square mile, and the last contains societies with 500+ persons per square miles).

Table E.5 in the online appendix reports pairwise correlations among the variables collected for the pre-colonial societies in the Ethnographic Atlas. As expected, societies characterized by more complex hierarchies do generally display a higher reliance on agriculture (and in particular on cereals), a higher probability of producing a farming surplus and more dense populations.

4.1.2 Country-level data

At the country level, we construct a hierarchy index using data from Borcan, Olsson and Putterman (2014). The data cover 159 modern-day countries for every half century from AD 50 to 2000. The score is based on the following question: Is there a government above the tribal level? Borcan et al. (2014) assigned 1 point if the answer is yes, 0.75 if it is a chiefdom, and 0 if the answer is no. These data are merged with information on: the legal origin of the country (from La Porta et al., 1998); population density in 1500 (Acemoglu, Johnson and Robinson, 2002); mortality of early settlers (Acemoglu, Johnson and Robinson, 2001); the number of exported slaves (Nunn, 2008); climate and latitude (Nunn and Puga, 2012); and genetic diversity (Ashraf and Galor, 2013). Table E.2 in the online appendix provides summary statistics for these variables.

4.1.3 Location of ancient cities and archaeological sites

In an attempt to capture differences in social and hierarchical complexity further back in time, we also collected data on the location of 54,507 sites with archaeological ruins, including 2939 ancient cities. We use three different sources of information. Data on the location of cities and towns that were founded before AD 400 were provided by Daniel DeGroff.[24] A detailed map is presented in Figure E.5 in the online appendix. Data on archaeological ruins come from Ancientlocations.net, a collection of interesting locations of the ancient world. Locations are included if they existed prior to AD 476 (end of the West Roman Empire) in the Old World and prior to 1492 in the New World. The data are complemented with archaeological data from the Megalith Portal, a web community with input from thousands of photographers and archaeologists. Ruins are classified according to 57 categories, which allows us to disentangle archaeological evidence of complex societies (e.g. pyramids, mines, temples and palaces) against other types of evidence (e.g. standing stones). The original intent of the portal was to categorize archaeological ruins in Great Britain and it was only recently extended to cover the entire world: as a result it oversamples Europe. We therefore exclude types of ruins that are only found in Europe and its surroundings, and always show the robustness of our regressions when excluding Europe.

We aggregate data on the location of cities and archaeological ruins at the 1x1 decimal degree

[24] https://sites.google.com/site/ancientcitiesdb

raster square. The first 8 rows of Panel A in Table E.3, in the online appendix present descriptive statistics on the number of cities and relevant archaeological ruins in each terrestrial raster point.

4.1.4 Radiocarbon-dated prehistoric archaeological sites

David and Ruth Whitehouse's (1975) "Archaeological Atlas of the World" provides a database of the most relevant global prehistoric and proto-historic archaeological sites, which were known at that time These 4,215 sites are dated through radiocarbon, a method to determine the age of objects containing organic material by using the properties of this radioactive carbon isotope.[25]

We geo-reference these sites and, using the information in the map titles and accompanying text, classify them depending on whether they pre-date the Neolithic transition in the relevant location or not. The result is a list of 825 sites that belong to pre-transition years and 3,309 sites that belong to the post-transition years. (We exclude 8 sites for which either geo-reference was not possible or dates were uncertain and 73 sites for which we were uncertain on whether they belong to the pre- or post- transition years).

We compute the number of pre-Neolithic sites and post-Neolithic sites at the 1x1 decimal degree raster square of the world. The Atlas also classifies these sites according to eight different categories[26]. In the empirical analysis we will either use all sites, or alternatively, only sites of prehistoric settlements. Panel B of Table E.3, in the online appendix, presents the descriptive statistics for these variables.

4.1.5 Soil suitability data

The nature of our study requires detailed spatial data on the suitability of soil for different crops. The Global Agro-Ecological Zones (GAEZ) project from the Food and Agriculture Organization (FAO) provides global estimates of potential crop yields for different crops with cell size of 5'x5' (i.e. approximately 100 Km^2) based on two possible categories of water supply (rain-fed and irrigation) and three different levels of inputs (high, medium and low). In addition, it supplies two

[25] Although this database is approximately 40 years old, Maurer, Pischke and Rauch (2016) conclude that: "While there has been much additional excavation in the intervening period, there is little reason to believe that it is unrepresentative for the coverage of sites and locations."

[26] These categories are: 1. Undifferentiated sites and find-spots 2. Settlements 3. Funerary monuments 4. Religious monuments 5. Caves and rock shelters 5. Cave art and rock reliefs 6. Hoards and votive deposits 7. Mineral sources 8. Mineral workings 9. Sites which combine several of the above categories.

alternative projections of potential crop-yields: one is based on agro-ecological constraints, which could potentially reflect human intervention, and one based on agro-climatic conditions, which are arguably unaffected by human intervention. To prevent concerns of reverse causality, we consider potential yields based on agro-climatic conditions under rain-fed low-input agriculture.

Figure 4: Difference in potential yields (calories per hectare) of cereals versus roots and tubers.

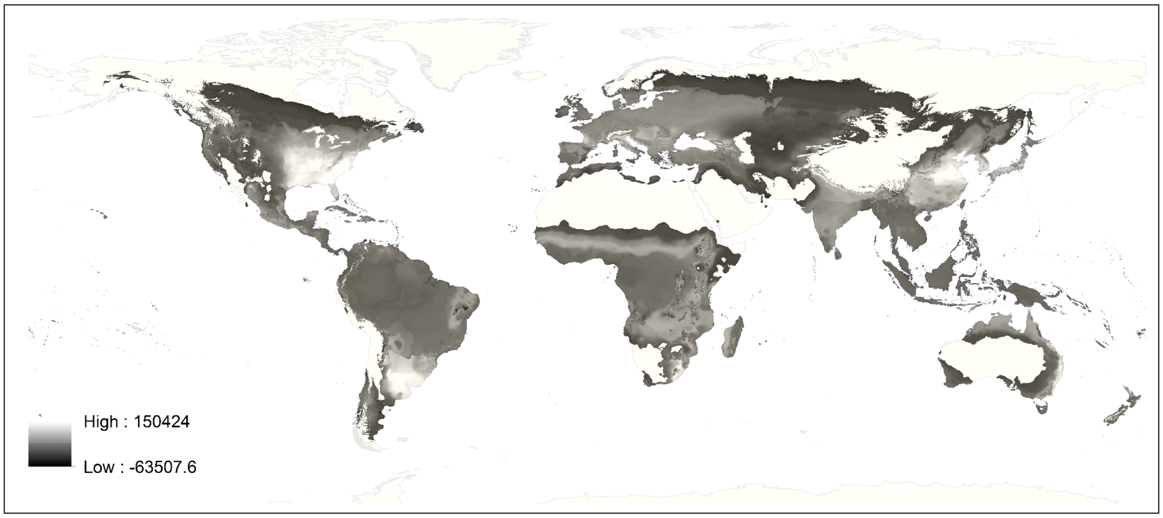


Figure 5:

GAEZ provides data on potential yields, in terms of tons per hectare per year, for 11 cereal grains and 4 roots and tubers. Following the same procedure as in Galor and Ozak (2016) for the crops relevant for our investigation, these yields are transformed from tons into calories using data on the caloric content of crops provided by the USDA National Nutrient Database for Standard Reference.[27] We then find the crop with the highest potential caloric yields for each raster point. The results are illustrated in figure E.4 in the Appendix. Cereal grains are the highest yielding crops in approximately 99 percent of the raster points in the sample, while roots and tubers are optimal in few very small areas in Siberia, Eastern Brazil and Central-East Africa. From these data we construct two measures: the productivity of land, measured as the maximum potential caloric yield per hectare; and the productivity advantage of cereals over roots and tubers, measured as the

[27] See Table E.4 in the appendix for the complete list of cereal grains, roots and tuber used in the empirical section and the correspondent caloric content.

difference between the maximum caloric yield of cereals and the maximum caloric yield of roots or tubers. The latter measure is described in Figure 5.

As robustness checks, we exploit two alternative measures of the productivity of the land, which have been widely used in the literature. The first one is an index developed by Ramankutty et al. (2002), which measures the fraction of land that is suitable for agriculture. The second one is a caloric suitability index developed by Galor and Ozak (2016), which captures the highest attainable potential caloric yields from 48 crops (which includes not only cereals, roots and tubers but also sugar crops, pulses, oil crops, vegetables, fruits, fiber crops and stimulant crops). Table E.5 in the appendix illustrates that our measure of the productivity advantage of cereals is positively correlated with our benchmark measure of land productivity (the correlation is slightly below 0.8), with the Ramankutty et al. index of suitable land (0.4) and with the Galor and Ozak caloric suitability index (0.8). We also construct a measure of the productivity advantage that comes from using the plow in agriculture. This equals the difference between the maximum caloric yield among crops that Alesina, Giuliano and Nunn (2013) identify as "plow-positive" (wheat, barley and rye) and those that they identify as "plow-negative" (sorghum, foxtail millet and pearl millet).

All these raster variables are attributed to the different societies in the Ethnoatlas by taking an average of their values within a 20-miles radius around the geo-coordinates reported in the Ethnoatlas;[28] and they are attributed to countries and the 1x1 decimal degrees raster squares by averaging them within these boundaries.

4.1.6 Other historical, demographic and geographic data

Larson et al. (2014) provide data on the 20 centers, in which the domestication of at least one plant or animal most likely took place (see Figure (E.5)) and the list of domesticates in each of these areas. We use these data to compute the distance of each raster point in the archeological data from the closest region of independent adoption of agriculture and from the closest region of independent domestication of cereal grains. Descriptive statistics on these two variables are reported in columns 11 and 12 in Table E.3.

Finally, GAEZ provides raster data on population density in 1995, precipitation and tempera-

[28] In the appendix we report the result of an alternative method, where we attribute these productivity measures to the different societies by using the maps on their spatial location constructed by Fenske (2013).

ture; the Global Digital Elevation Map (GDEM) provides raster data on elevation and ruggedness; the History Database of the Global Environment (HYDE) provides raster data on global estimates of population density between 1500 and 2000. These data are averaged within societies in the Ethnoatlas, countries and 1x1 decimal degree raster points.

4.2 Empirical results

4.2.1 The choice of crop

Our theory suggests that hierarchy would be correlated with the cultivation of appropriable crops, and in particular with the cultivation of cereals. Yet our theory does not imply a uni-causal relation between these two variables: the cultivation of cereals not only facilitates hierarchy, but also requires the protection that hierarchical institutions can provide. Thus, before advancing to examine hierarchy, in this subsection we examine the linkage between land suitability as measured by the GAEZ data, and actual crop choice in the Ethnographic Atlas database. To recall, our theory suggests that farmers choose the crop type on the basis of comparing the net caloric yield of cereals to that of the alternatives crops (where we focus on roots and tubers), taking into account the greater vulnerability of cereals to appropriation by bandits or the state.

In Table 1 we report regression results of the form:

$$Y_i = \alpha CerAdv_i + X_i'\beta + \varepsilon_i,$$

where the dependent variable Y_i is either a dummy variable that identifies societies choosing cereals as the main crop ($CerMain_i$) or a measure of the reliance of these societies on agriculture for their diet ($AgrRely_i$); $CerAdv_i$ is the caloric advantage of cereals in the land of society i (the difference between the maximum potential caloric yield of cereals and of roots or tubers), and X_i' is a vector of control variables.

Column 1 reports the bivariate relationship between cereal advantage and the choice to cultivate cereals as the main crop, without any controls. The association is positive and statistically significant. An increase in the productivity advantage of cereals over roots and tubers by one standard deviation is associated with an increase in the probability of growing cereals as the main crop by about 20 percent. Moreover, variation in this regressor alone is able to explain 13 percent of the variation in the dependent variable. Column 2 reports the results when adding as a control variable land productivity (the highest potential caloric yield across all 11 cereals and 4 roots/tubers), to address the concern that the productivity advantage of cereals might reflect land productivity. The impact of the productivity advantage of cereals is unchanged, land productivity doesn't have any significant impact on crop choice and the R^2 of the regression is practically unchanged. These results are repeated when introducing continental fixed effects (in column 3), and when resorting to logistic estimation to account for the binary nature of the dependent variable (columns 4 and 5).

In appendix E we report that these results survive a battery or robustness checks. These include controlling sequentially for: precipitation, temperature, elevation, and ruggedness, which are the main factors affecting crop productivity in the GAEZ dataset (Table E.6). We control for: geographical isolation (proxied by the distance to the nearest major river or coast), historical and current population density; evidence of intensive irrigation; and the productivity advantage from using the plow (Table E.7). In all cases, the qualitative results on the effect of cereal productivity advantage over roots and tubers are almost unaffected: the coefficients vary from 0.250 to 0.261 and are always statistically significant at the 1 percent confidence level.

In the last three columns of Table 1 the reliance of the society on agriculture is the dependent variable. As reasonably expected, land productivity increases the probability of reliance on farming. Interestingly, the productivity advantage of cereals has a negative effect on practicing agriculture, when controlling for land productivity. That is, for the same level of land productivity (measured in calories), cultivating roots and tubers is more rewarding than cultivating cereals. This is consistent with our theory as cereal is more vulnerable to expropriation.

4.2.2 Cereals and hierarchy: 2SLS estimates using the Ethnographic Atlas data

According to our theory, societies that grow cereals rather than roots or tubers are characterized by a more complex hierarchy and by generating a higher farming surplus. To test these predictions with the Ethnographic Atlas data, we estimate a regression of the form:

$$Y_i = \alpha_1 CerMain_i + \alpha_2 LandProd_i + X_i'\beta + u_i, \tag{2}$$

Table 1: Potential Crop Yields, Choice of Crops and Reliance on Agriculture

	Dependent variable is:							
	Major crop is cereal grains ($CerMain$ dummy)					Reliance on agriculture ($AgrRely$)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	Logit	Logit	OLS	OLS	OLS
$CerAdv$	0.205***	0.210***	0.253***	1.150***	1.617***	0.081***	-0.098***	-0.046**
	(0.029)	(0.063)	(0.059)	(0.339)	(0.380)	(0.022)	(0.029)	(0.022)
$LandProd$		-0.007	-0.137**	-0.119	-0.896**		0.230***	0.128***
		(0.083)	(0.069)	(0.384)	(0.407)		(0.046)	(0.035)
CONTINENT FE	NO	NO	YES	NO	YES	NO	NO	YES
Ave marg. effect of				0.282***	0.385***			
$CerAdv$				(0.081)	(0.092)			
r2	0.132	0.132	0.359			0.0733	0.235	0.387
pseudo r2				0.109	0.258			
N	982	982	982	982	982	1063	1063	1063

The table reports cross-sectional OLS and Logit estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

where Y_i is either a measure of hierarchy or an indicator for the presence of farming surplus in society i; $CerMain_i$, is, as mentioned above, a dummy variable that identifies societies that rely mainly on cereals for their subsistence; $LandProd_i$ is a measure of land productivity, and X_i' is a vector of control variables. This specification, however, encounters several problems.

First, the choice of the cultivated crop is influenced by the social institutions. According to our theory, cereals render farmers more vulnerable to theft, in particular in societies characterized by limited hierarchy and limited protection against bandits. To overcome this reverse causality concern, we exploit variations in potential, rather than actual, crop yields, which are derived from agro-climatic conditions that are presumably orthogonal to human intervention. Specifically, we run IV regressions, where we instrument for $CerMain_i$ by using the productivity advantage of cereals, $CerAdv_i$.

Second, there are several potential omitted variables that could be correlated with the main regressor and the measure of hierarchy. The disease environment, for instance, is correlated with both the cultivation of tubers (which is concentrated in the tropics) and is likely to be correlated with the quality of institutions (Acemoglu, Johnson and Robinson, 2001). A battery of robustness checks alleviates this concern. Moreover, in two of the following subsections, we conduct panel regressions that alleviate concerns regarding potential time-invariant omitted variables.

Table 2: Cereals, Surplus and Hierarchy - Reduced Form

	Dependent variable is:							
	Jurisdictional Hierarchy Beyond Local Community					Existence of Farming Surplus		
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Ord Logit	(6) OLS	(7) OLS	(8) Logit
CerAdv		0.244***	0.179	0.274**	0.495***	0.241***	0.202***	0.997***
		(0.069)	(0.120)	(0.107)	(0.149)	(0.0681)	(0.0742)	(0.384)
LandProd	0.239***		0.082	-0.188*	-0.224	-0.132	-0.0985	-0.479
	(0.075)		(0.141)	(0.108)	(0.178)	(0.0870)	(0.0985)	(0.463)
CONTINENT FE	NO	NO	NO	YES	YES	NO	YES	YES
Ave marg. effect of								0.249***
CerAdv								(0.096)
r2	0.0361	0.0416	0.0429	0.249		0.0911	0.157	
pseudo r2					0.121			0.124
N	952	952	952	952	952	140	140	140

The table reports cross-sectional OLS (columns 1-3 and 5-7), Ordered Logit (column 4) and Logit (column 8) estimates. The unit of observation is the society in Murdock's Ethnoatlas. Columns 1-4 report in parentheses Conley standard errors adjusted for spatial correlation, while columns 5-8 report robust standard errors. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Before presenting the 2SLS regressions that estimate the effect of cereals on hierarchy and surplus, we report in Table 2 the reduced form analysis. Column 1 shows a significant correlation between (potential) land productivity and the level of jurisdictional hierarchy in the societies in the Ethnoatlas, as predicted by the conventional productivity-and-surplus theory. Column 2 illustrates a significant correlation between the productivity advantage of cereals and hierarchy as predicted by the appropriability theory. Once both regressors are included (columns 3-8) the effect of the productivity advantage of cereals remains positive and significant (excluding in column 3) and the effect of land productivity disappears. An increase of one standard deviation in the productivity advantage of cereals increases the hierarchy index by 0.27 in the specification with continent fixed effects (column 4). In column 5, we use an ordered logit model to account for the ordinal nature of the dependent variable. A one standard deviation increase in the productivity advantage of cereals increases the log odds of being in a higher level of hierarchy by approximately 50 percent.

In the appendix (Table E.8), we relax the assumption of proportional odds, which is implicit in the standard ordered logit models, and estimate a generalized logit model.[29] As can be seen, the greatest impact of cereal advantage is to push societies from tribes and chiefdoms to states. More specifically, while an increase in one standard deviation in the productivity advantage of cereals increases the log odds of being in a level of hierarchy higher than a tribe by 32 percent, it increases the log odds of being in a level higher than a chiefdom by 65 percent and higher than a small state by 84 percent. In all cases, the impact of land productivity is either very small and not statistically significant, or negative.

Columns 6 to 8 in Table 2 provide further support for the appropriability hypothesis: the productivity advantage of cereals has a positive effect on the probability of having an economy that produces a farming surplus (elite consumption isn't based on direct subsistence). In all the regressions we include both the productivity advantage of cereals and land productivity. Only the former has a significant impact on surplus, independently on whether we control for continent fixed effects or not (columns 6 and 7), or use a logistic regression (column 8).

Table 3: Cereals and Hierarchy - OLS and 2SLS

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.707***	1.170***	0.863	1.040**	0.304**	0.892**	1.064**	0.993**
	(0.131)	(0.359)	(0.596)	(0.414)	(0.120)	(0.420)	(0.538)	(0.463)
LandProd			0.081				-0.037	
			(0.127)				(0.071)	
DEPENDENCE ON AGRICULTURE				0.334				-0.419
				(0.517)				(0.783)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	952	952	952	952	952	952	952	952
F excl instrum.		147.7	44.84	65.51		99.87	76.90	33.09

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

[29] The assumption of proportional odds means that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable.

Table 3 reports the OLS and 2SLS estimates of equation 2, when the dependent variable is hierarchy. The OLS estimates in column 1 show that cultivating cereals is associated with an increase of 0.70 in the hierarchy measure. Clearly, this positive association cannot be interpreted as causal. To overcome the reverse causality problem, we switch to the 2SLS estimates in the next three columns. Cultivating cereals as the main crop increases the hierarchy measure by more than one (column 2), which is equivalent, for instance, to a move from a tribe to a small chiefdom or from a large chiefdom to a state. In column 3, we add land productivity as a control variable. As can be seen, it does not have any significant effect on hierarchical complexity. It should be mentioned that in this specification the effect of cereals is rendered insignificant. It is significant in all other specifications. Column 4 includes the dependence of the society on agriculture instrumented with caloric advantage of cereals and land productivity. Results indicate that societies that practice agriculture are not characterized by more complex hierarchies unless they cultivate cereals. In columns 5-8, we repeat the analysis adding continent fixed effects in the regression. The 2SLS results are practically unchanged, with the exception of column 6 in which cultivating cereals has a significant effect on hierarchy despite controlling for land productivity.

The results of Table 3 survive a battery of robustness checks presented in the on line appendix. In Table E.9, we control sequentially for precipitation, temperature, elevation, and ruggedness – the main factors affecting crop productivity: none of them seems to explain our main result. Table E.10 addresses other potential channels, through which the cultivation of cereals might affect economic development and hierarchical complexity. In an attempt to rule out the possibility that cereals might affect development by increasing the tradability of food resources, we control for geographical isolation (proxied by distance from a major river and distance from the coast). There is some evidence that closeness to a river is associated with more complex hierarchies, but the impact of cereals on hierarchy is unaffected. We control also for historical and current population density,[30] which is associated in several theories with increased hierarchy (and with economic and technological development). Cereals still have a large impact on hierarchy when adding these

[30] We use two different proxies for historical population density. The first one, HYDE, is based on historical recostruction at the raster level and is available for the entire sample of societies in the Ethnoatlas. Because historical population reconstruction is unavoidably inexact, we also show the robustness of our results when using data from Pryor (1985), which are available for a much smaller sample (only 144 societies) and are based on completely different sources.

controls. The data, as one would expect, reveal a positive correlation between population density and hierarchical complexity. We control for evidence of agriculture based on intensive irrigation. This could be an important potential confounder as Bentzen et al. (2016) provide evidence of a causal impact of irrigation on autocracy. Our estimates confirm the validity of their results: societies that practice intensive irrigation are characterized by relatively more complex hierarchies. Our results on the impact of cereals are unchanged. Finally, results are robust to the inclusion of the productivity advantage of the plow, a variable that Alesina, Giuliano and Nunn (2013) have identified as an important determinant of gender roles.

The qualitative results are also practically unaffected (the coefficient varies between 0.750 and 1.471) when using ethnic boundaries as defined by Fenske (2013) to extract data on crop productivities (Table E.12), when the sample includes societies living in desertic soils (Table E.11), or when using either the Ramankutty et al. index of fertile land or the Galor and Ozak index of caloric suitability as alternative measures of land productivity (Table E.13). Table E.14, in the appendix, reports the OLS and 2SLS estimates of equation 2, when the dependent variable is the existence of a farming surplus in the society. The OLS estimates show that cultivating cereals is associated with an increase of 0.36 in the probability of producing a surplus. The coefficient more than doubles when turning to the 2SLS estimates and also in this case land productivity and reliance on agriculture do not affect the dependent variable. Results are robust to adding continent fixed effects. These results also survive a long list of robustness checks reported in (Tables E.15-E.19) in the appendix.

4.2.3 Panel data employing the Columbian Exchange in crop availability

The results thus far indicate that cultivating cereals has a large and significant effect on the development of complex hierarchical societies and that land productivity (instrumented with the same data as for cultivating cereals) does not. The analysis accounts for a large set of possible confounding factors, but we cannot rule out that unobservable characteristics that are systematically correlated with the productivity of different crops might be driving our results. To alleviate this potential concern, we exploit in this sub-section the exogenous change in the available crops in different locations of the world that was induced by the Columbian Exchange.[31]

Among the main four roots and tubers that we considered thus far, three were available before 1500 only in the New World: cassava, white potatoes and sweet potatoes. And among the eleven main cereals, only maize was available there. In the Old World, yam was the only available crop from among the four main roots and tubers, while all cereals, other than maize, were available. Accordingly, we re-compute for each location the productivity advantage of cereals over tubers and the absolute productivity of land both before the Columbian Exchange (prior to 1500) based on the relevant subset of crops, and after the exchange (after 1600) for the full set of crops.[32]

We use the country-level panel dataset constructed by Borcan et al. (2017). The unit of observation is the territory delimited by modern-day country borders for 159 countries every 50 years. We use the data for the last millennium. Since we lack observations on the major crop cultivated in these territories for the period of analysis, we only run the reduced form version of our empirical analysis: we regress the hierarchy index on the productivity advantage of cereals and on land productivity:

$$Hierarchy_{it} = \alpha_1 CerAdv_{it} + \alpha_2 LandProd_{it} + X'_{it}\beta + \eta_i + \eta_t + u_{it}. \tag{3}$$

Here, the dependent variable is an index of country i in year t and $CerAdv_{it} = CerAdv_{i,BeforeExchange}$ (the caloric advantage of cereals over roots and tubers before the Columbian Exchange) if $t \leq 1500$ and $CerAdv_{it} = CerAdv_{i,AfterExchange}$ (the caloric advantage after the Columbian Exchange) if $t \geq 1600$. Similarly, potential land productivity ($LandProd_{it}$), is calculated based on the pertinent crops available before and after the Columbian Exchange. X_{it} is a set of control variables. Country fixed effects control for all time invariant factors that differ between countries, and time period fixed effects control for any time pattern of hierarchical complexity that affects all countries simultaneously. The critical identification assumption is that there were no unobserved events in

[31] Up until now, our use of the GAEZ data was based on the presumption that all potential crops were known in every location. Since the Ethnographic Atlas data pertain to data collected by Anthropologists and travelers in the 18th-20th centuries, this presumption is likely not to pose a serious problem.

[32] The historical evidence points out that the New World's crops were adopted in Europe and Africa only in the seventeenth century. For instance, potato cultivation in the Old World began in the late seventeenth century by Irish peasants (Nunn and Qian, 2011), while the first accounts on the adoption of maize in Africa date back to the very end of the sixteenth century (Miracle, 1966). We thus exclude the years from 1500 to 1600. In appendix E we show that our results are robust when excluding the years between 1500 and 1750 (Table E.24).

the sixteenth century that are systematically correlated with the spatial variation in the change in the potential productivity advantage of cereals, and that had an independent effect on hierarchy.

We are aware that the change in crop availability induced by the Columbian Exchange was coincident with colonization. However, we contend that the colonization process does not seem to be driving our results: excluding colonies from the estimation sample doesn't have a quantitative effect on the estimates (see Table E.21). Moreover, the concern that changes in hierarchy were a result of colonization rather than of changes in the availability of crops, cannot explain the observation on the differential impact of the changes in cereal advantage and in changes in land productivity that we observe in Table 4.

Table 4: Cereals and Hierarchy - Panel Regressions

	Dep. Variable: Hierarchy Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CerAdv		0.189***	0.272***	0.282***	0.240***	0.255***	0.261***	0.197**
		(0.0683)	(0.0834)	(0.0760)	(0.0857)	(0.0889)	(0.0839)	(0.0795)
LandProd	0.141		-0.163	-0.193	-0.152	-0.115	-0.148	-0.165
	(0.0971)		(0.141)	(0.131)	(0.139)	(0.142)	(0.138)	(0.123)
Controls (x Year FE):								
Precipitation		NO	NO	YES	NO	NO	NO	NO
Temperature		NO	NO	NO	YES	NO	NO	NO
Elevation		NO	NO	NO	NO	YES	NO	NO
Ruggedness		NO	NO	NO	NO	NO	YES	NO
Abs Latitude		NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES	YES
TIME FE	YES	YES	YES	YES	YES	YES	YES	YES
r2	0.669	0.680	0.682	0.716	0.684	0.681	0.686	0.705
N	2869	2869	2869	2850	2812	2755	2869	2869

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. Robust standard errors, clustered at the country-level, in parentheses. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Column 1 in table 4 shows a positive but insignificant effect of land productivity on hierarchy, when cereal advantage is not controlled for. The sign of this coefficient turns negative but insignificant in all other specifications, when the cereal advantage is included in the regression. Column 2 confirms that the higher is cereal advantage, the higher is the country's hierarchy index. A one standard deviation increase in the productivity advantage of cereals increases the hierarchy index by 0.19. In the next six columns, we show that the results are robust when controlling for land productivity, and in addition also for precipitation, temperature, elevation, ruggedness and absolute latitude (interacted with the time-period fixed effects). In Table E.20, in the online appendix, we consider a host of additional factors (each interacted with time-period fixed effects) that might have impacted hierarchical complexity. Our choice of controls is driven by the determinants of long-term economic development that have been emphasized in the literature: legal origin of the country, population density in 1500, settlers mortality, the number of exported slaves, genetic diversity, distance to rivers and coast, endemicity of malaria, and the percentage of tropical land. The key results are essentially unaffected.

In Table E.21 in the appendix we exclude from the sample the cells in which the countries in our analysis were either colonies or protectorates. The estimated coefficient on the caloric advantage of cereals over roots and tubers become smaller by approximately a third. However, in all the specifications, the effect of cereal advantage remains positive and statistically significant, while the impact of land productivity on hierarchy remains insignificant. Table E.22 and E.23 in the appendix report further robustness checks. Specifically, in Table E.22, hierarchical complexity is proxied by a dummy that identifies societies with a government above tribal level. In Table E.23, land productivity is proxied by the caloric suitability index developed by Galor and Ozak (2016), which also varies depending on whether it is measured before or after the Columbian Exchange. Finally, in Table E.24, we exclude the years between 1500 and 1750, during which the Columbian Exchange of crops was not completed. In all the three cases, our main results are unaffected.

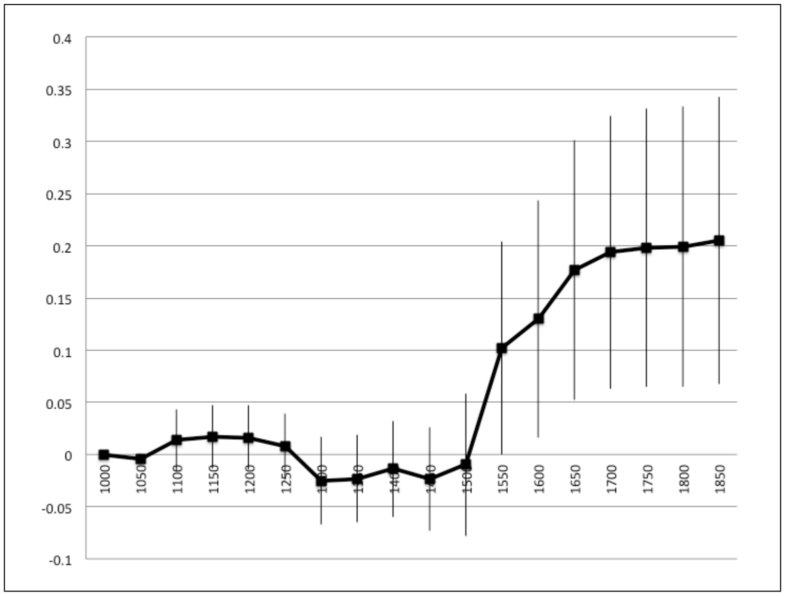
The estimation of equation (3) requires an assumption on the date in which the Columbian Exchange was completed. We also estimate a more flexible equation that takes the following form:

$$Hier_{it} = \sum_{j=1050}^{1850} \alpha_j (CerAdv_{i,AfterExchange} - CerAdv_{i,BeforeExchange}) + X_{it}'\beta + \eta_i + \eta_t + u_{it}. \quad (4)$$

This specification does not require any assumption about the timing of the Columbian Exchange and allows to test for potential pre-trends in the dependent variable. However, it does not capture the magnitude of the impact of the caloric advantage of cereals on hierarchy: because the main regressor is time invariant and equation (4) includes country and time-period fixed effects, the estimated αs must be measured relative to a baseline time-period, which we take to be the year 1000. The estimated coefficients and their 10 percent confidence intervals are reported in Figure

(6).[33]

Figure 6: Flexible estimates of the relationship between the change in the caloric advantage of cereals over roots and tubers due to the Columbian exchange and hierarchy.



The impact of the change in the productivity advantage of cereals over tubers enabled by the Columbian Exchange is constant over time between 1000 and 1500; it increases steadily during the sixteenth century and continues to increase, but at a lower rate until 1700, after which it stabilizes. This confirms that the Columbian Exchange produced a differential increase in hierarchy in those countries for which it also caused a larger increase in the productivity advantage of cereals over roots and tubers, and that the great majority of the full impact happened in the sixteenth century. It also rules out the possibility that pre-trends might be driving our results.

4.2.4 Cross-section investigation using the location of ancient cities and archaeological data

The cross-section results, based on the pre-colonial societies in the Ethnographic Atlas, and the panel estimates, based on the country-level data provided by Borcan et al. (2014), support our

[33] The 17 coefficients reported in Figure (6), can also be described as the estimated coefficients in 17 independent cross-country regressions, in which we regress the change in the hierarchy index between each of the 17 years in the sample (1050, 1100, .., 1850) and the year 1000 on the change in the caloric advantage of cereals over roots and tubers caused by the Columbian exchange.

appropriability theory. But both employ data from the last millennium. We now test our theory with earlier data, closer to the Neolithic adoption of farming.

Table (5) employs a dataset on the location of ancient cities (founded before AD 400 in the Old World and before AD 1492 in the New World). With the underlying presumption that the existence of a city is an indicator of hierarchy, we use a grid of the world land surface, in which the unit of observation is the 1x1 decimal degree raster, to test whether our theory can explain the geographical distribution of these ancient cities. We run regressions of the form:

$$Y_i = \alpha_1 CerAdv_{i,BeforeExchange} + \alpha_2 LandProd_{i,BeforeExchange} + X_{it}'\beta + u_{it}, \tag{5}$$

where the main regressors are measured on the basis of the pre-Columbian crops available in each location.

Table 5: Potential Crop Yields and the Location of Ancient Cities.

	Dependent variable is:							
	Presence of an ancient city (dummy)					Log(1+ number ancient cities)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	Logit	OLS	OLS	OLS
CerAdv		0.145***	0.129***	0.0340**	1.256**	0.186***	0.167***	0.0357**
		(0.0388)	(0.0380)	(0.0162)	(0.603)	(0.0538)	(0.0512)	(0.0180)
LandProd	0.024***	-0.0864***	-0.0744***	-0.0126	-0.208	-0.111***	-0.0966***	-0.0139
	(0.009)	(0.0267)	(0.0256)	(0.0128)	(0.542)	(0.0361)	(0.0333)	(0.0138)
CONTINENT FE	NO	NO	YES	NO	NO	NO	YES	NO
COUNTRY FE	NO	NO	NO	YES	YES	NO	NO	YES
Ave marg. effect of					0.0135**			
CerAdv					0.006			
r^2	0.0195	0.0841	0.0986	0.451		0.0773	0.0865	0.574
N	15927	15927	15927	15927	9032	15927	15927	15927

The table reports cross-sectional OLS and Logit estimates and the unit of observation is the 1x1 decimal degree square. Robust standard errors, clustered at the country-level, in parentheses. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Column 1 shows that higher land productivity is associated with a higher probability of finding an ancient city, as predicted by the conventional productivity-and-surplus theory. Column 2 illustrates that a higher productivity advantage of cereals over roots and tubers is also associated with an increase in the probability of finding an ancient city in that area. In addition, it shows once again that if we control for cereal advantage, the positive effect of land productivity vanishes. Results are robust to adding continent fixed effects presented in column 3. A one standard deviation increase in the cereal advantage is associated with an increase in the probability of finding a city by 15 percent without continent fixed effects and 13 percent in the specification with continent fixed effects.

A potential concern with these estimates is that data on the location of ancient cities might have different levels of accuracy in different (modern) countries. In column 4 we only exploit within country variation. From a qualitative point of view, adding country fixed effects does not change the results substantially, but the coefficient on cereal caloric advantage drops by a factor of 3. In column 5, we use a logit model to account for the binary nature of the dependent variable. The results are practically unchanged. Finally in columns 6-8 we look at the extensive margin: the productivity advantage of cereals is associated with a larger number of cities.

In Table (E.25) in the appendix we report results from a battery of robustness checks. Specifically, the estimates are practically unchanged when controlling sequentially for precipitation, temperature, elevation, ruggedness and absolute latitude. Moreover, results are robust to excluding Europe or deserts from the sample. (Figure (E.5) in the appendix shows that ancient cities are concentrated in Europe, while they are absent in desert areas).

The test reported in Table (5) gets us closer in time to the Neolithic period. We now attempt to connect our analysis with the domestication of plants during the Neolithic. The underlying idea is that the probability that any given crop would reach a certain area would be positively associated with the geographic distance between this area and the nearest area in which that crop was domesticated. Therefore, our theory predicts that raster points that are geographically close to centers where cereals were first domesticated, relative to the distance to areas of tuber domestication, would be more likely to adopt cereal farming, and thus develop hierarchies and build cities. Global data on the diffusion of crops during the Neolithic transition are not available, but archaeologists and botanists have identified some 20 centers in which independent domestication took place and from which domesticated crops spread to the rest of the world (see map E.5 and Larson et al., 2014). We use these data to compute for each raster point two new measures: the distances to the nearest center of independent domestication of agriculture ($DistanceAgr$), and the distance to the nearest center of independent domestication of cereal grains ($DistanceCer$).

Using the same dependent variables on early urbanization as in Table (5), we now estimate regressions of the form:

$$Y_i = \alpha_1 DistanceCer_i + \alpha_2 DistanceAgr_i + X'_{it}\beta + u_{it}. \tag{6}$$

Table 6: The Origin of the Neolithic Transition and the Location of Ancient Cities.

	Dependent variable is:						
	Presence of an ancient city (dummy)				Log(1+ number ancient cities)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	Logit	OLS	OLS	OLS
DistanceCer		-0.00214***	-0.00143**	-0.187***		-0.00253***	-0.00147**
		(0.000597)	(0.000604)	(0.0333)		(0.000767)	(0.000647)
DistanceAgr	-0.00120***	0.000909	0.000253	0.112***	-0.00117***	0.00133	0.0000658
	(0.000343)	(0.000676)	(0.000566)	(0.0379)	(0.000409)	(0.000900)	(0.000680)
CONTINENT FE	NO	NO	YES	YES	NO	NO	YES
Ave marg. effect of				-0.002***			
DistanceCer				(0.001)			
r2	0.00949	0.0307	0.0495		0.00512	0.0220	0.0376
N	15927	15927	15927	15116	15927	15927	15927

The table reports cross-sectional OLS and Logit estimates and the unit of observation is the 1x1 decimal degree square. Robust standard errors, clustered at the country-level, in parentheses. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

The first column of Table (6) shows that distance from the nearest area of independent adoption of agriculture is negatively correlated with urbanization. The second column shows that distance from the nearest area of cereal domestication is negatively correlated with urbanization, while distance from the nearest area of independent adoption of agriculture is not, once the distance from cereal domestication is included in the regression. This result, too, is robust when considering within-continent variation (column 3), using logit rather than OLS (column 4), or when looking at the extensive margin (columns 5-7). In the appendix table E.27, we show that results are robust to controlling sequentially for precipitation, temperature, elevation, and ruggedness. In Table E.28, we extend the analysis of Table 7 by controlling for: absolute latitude, irrigation potential, productivity advantage from use of the plow, and population density in 1995. In the last two columns of this table, we limit the analysis to Asia and Africa and exclude deserts. The results are robust with one exception. When controlling for absolute latitude the coefficient on distance to the areas of cereal domestication drops by half and is no longer statistically significant (column 1). The great majority of the centers of cereal domestication are concentrated in a very limited latitude band (between 10 and 40 degrees N), in which there are almost no centers of domestication of roots and tubers. Thus, the data limit our ability to disentangle the impact of distance to the closest cereal domestication center and absolute latitude. There are, however, a couple of exceptions to this rule: in Lingnan (South China, lat: approx. 25 N) yams and taro were domesticated but not cereals, and in the Sudanese savannah (lat: approx. 5 N) sorghum was domesticated. When limiting the analysis to Africa and Asia - the continents of these two areas - we find that, even after controlling for absolute latitude, distance to the closest centers of early domestication of cereals matters, while distance to the closest centers of early domestication of roots and tubers does not (column 2).

In Table (E.29) in the online appendix, we replace the left hand side variable - previously ancient cities - with alternative archaeological sites (which predate the 476 CE fall of the Roman Empire in the Old World, or the 1492 discovery of the Americas in the New World). Once again, we view the existence of some of these site categories as a proxy for hierarchy/civilization. Column 1 documents a negative and statistically significant correlation between the presence of ancient archaeological sites and the distance from the nearest area of domestication of cereals. As before, distance to the nearest center of domestication of other crops does not seem to matter when controlling for distance to centers of cereal adoption. This result is confirmed when archaeological sites are pyramids, ancient temples, ancient mines, ancient palaces and ancient sculptured stones, but not when the archaeological sites are ancient standing stones. The results reported in this table are also valid when examining the extensive margin and focusing on the number of archaeological sites in the cell (Table (E.30) in the appendix) and when excluding Europe (Table (E.31) in the appendix).

In Table (E.32), in the online appendix, we present regression results with the same left hand side archaeological sites variables, regressed on land productivity and the productivity advantage of cereals over roots and tubers. The latter has a positive and significant effect on the presence of ancient archaeological ruins, ancient mines, ancient palaces, ancient sculptured stones, and standing stones, but not when archaeological sites are pyramids or ancient temples. In all regressions, land productivity doesn't have a positive effect on any of these variables.

4.2.5 Difference-in-difference using radiocarbon-dated prehistoric archaeological sites

The results in the previous section are still not definite for two main reasons: 1. They are based on a cross-section and, even though we control for a large set of confounders, we cannot exclude that the cereal productivity advantage and the proximity to centres of cereal domestication could be correlated with unobservable geographic characteristics affecting the location of ancient cities and archaeological sites. 2. The data used are obtained from online sites that are not always connected with academic sources.

To handle these two limitations, in this last subsection, we move to the radiocarbon-dated prehistoric and proto-historic archaeological sites listed in David and Ruth Whitehouse's (1975) "Archaeological Atlas of the World". We assign each of these sites to a 1x1 decimal degree raster point of the world land surface and count the number of pre-Neolithic sites and post-Neolithic sites in each of these points. We then run the following difference-in-difference regressions:

$$Y_{i,t} = \alpha_1 CerAdv_{i,BeforeExchange} * Post_t + \alpha_2 LandProd_{i,BeforeExchange} * Post_t + X'_{it}\beta + \eta_i + \eta_t + u_{it}, \tag{7}$$

and:

$$Y_{i,t} = \alpha_1 DistanceCer_i * Post_t + \alpha_2 DistanceAgr_i * Post_t + X'_{it}\beta + \eta_i + \eta_t + u_{it}. \tag{8}$$

where the subscript i indicates the raster point of the world; the subscript t indicates whether the site pre-dates or not the Neolithic transition; η_i and η_t are cell and period fixed effects, and $Post_t$ is a dummy variable that identifies archaeological sites dating after the Neolithic transition.

Column 1 of Table (7) shows that a higher productivity advantage of cereals over roots and tubers is associated with a relative increase in the probability of finding a post-Neolithic site rather than a pre-Neolithic site, confirming that the Neolithic transition led to more visible traces of human societies but only in areas where agriculture started with cereals. In addition, it illustrates that if we control for cereal advantage, land productivity does not produce any significant positive effect. Column 2 shows that distance from the nearest area of cereal domestication is associated with a decrease in the probability of finding a post-Neolithic site rather than a pre-Neolithic site. Again, once distance from cereal domestication is included in the regression, distance from the nearest area of independent domestication does not produce any significant effects. The results reported in the rest of Table (7) confirm the results in columns 1 and 2, with the difference that the dependent variable is either the number of archaeological sites (columns 3 and 4), the presence of a prehistoric settlement (columns 5 and 6), or the number of prehistorical settlements in the area (columns 7 and 8).

Table 7: Potential Crop Yields and the Location of Ancient Cities.

	Dependent variable is:							
	archaeol. site (dummy)		Log(1+ # archaeol. site)		ancient settlem. (dummy)		Log(1+ # ancient settlem.)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post X CerAdv	0.0683**		0.117***		0.0668**		0.101***	
	(0.0293)		(0.0417)		(0.0283)		(0.0365)	
Post X LandProd	-0.0266		-0.0581**		-0.0226		-0.0456*	
	(0.0208)		(0.0281)		(0.0203)		(0.0247)	
Post X DistanceCer		-0.00152**		-0.00193***		-0.00141**		-0.00165***
		(0.000601)		(0.000719)		(0.000559)		(0.000632)
Post X DistanceAgr		0.000381		0.000909		0.0000127		0.000430
		(0.000643)		(0.000850)		(0.000568)		(0.000705)
CELL FE	YES	YES	YES	YES	YES	YES	YES	YES
TIME FE	YES	YES	YES	YES	YES	YES	YES	YES
r2	0.0710	0.0253	0.0800	0.0265	0.0738	0.0316	0.0765	0.0294
N	31854	31854	31854	31854	31854	31854	31854	31854

The table reports difference-in-difference OLS regressions. The unit of observation is the 1x1 decimal degree square either before or after the Neolithic transition. Robust standard errors, clustered at the country-level, in parentheses. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

4.3 Caveats

We acknowledge that our extensive empirical tests do not provide full proof for our claims regarding the validity of the appropriability theory and against the productivity and surplus theory. Clearly, we cannot conduct a randomized control experiment to confirm this theory. We, therefore, rely on available data. Yet, none of the datasets that we employ is perfect. The FAO measures of potential productivity pertain to modern data. Even if the caloric content per unit weight, as well as soil and climatic conditions, may not have changed much over the last ten millennia, the potential land yield has evidently changed significantly. In particular, more extensive learning by doing and communication can be presumed to have increased the potential yield advantage of cereals over roots and tubers. The Ethnographic Atlas cross-sectional data on pre-colonial societies also pertains to a period much later than the early transition to farming. It is based on data collected by ethnographers, and as such is not a random sample of societies, but rather biased towards relatively isolated societies with relatively low levels of hierarchy. The panel data on state institutions before and after the Columbian Exchange suffers from the obvious problems that by the time of the Columbian exchange, most countries were already organized as states, and that the exchange itself coincided with the onset of colonialism. The two datasets on excavated archaeological sites are also far from perfect. These datasets are biased towards the Euro-Asian core. In addition, the first archaeological dataset is from a non-academic source, while the second is outdated.

We also acknowledge that our tests are based on discretely distinguishing cereals from roots and tubers, which provides only a crude proxy for a more continuous, multi-faceted measure of appropriability. Similarly, for lack of a suitable measure of surplus, we engage only rudimentarily with the issue of whether surplus may have existed prior to the emergence of hierarchy. We have also focused entirely on caloric productivity, whereas nutrition requires much more than calories.

And still, we contend that the robustness of our empirical results, using alternate datasets and multiple alternative specifications, in addition to the non-quantitative evidence discussed in the first two sections, and our theoretical arguments, suggest that one should reject the prevailing productivity-and surplus explanation for the emergence of hierarchy, and, pending further studies, prefer the proposed appropriability explanation.

5 Alternative Theories for the Emergence of Hierarchy and States

In this section we review the extensive literature that links the transition to agriculture and the shift away from egalitarianism towards hierarchy. We note though, that almost all the literature in this field refrains from differentiation between crop types. As a result, those who seek to generalize from the experience of cereal-based farming, typically fail to explain the social developments in tuber-based countries like New Guinea. On the other hand, those who base their models on ethnographic evidence, often generalize from the experience of chiefdoms that typically rely on non-cereal modes of farming.

Both the appropriability and productivity-and-surplus explanations for the emergence of hierarchy can be traced to Adam Smith. According to Smith, government and property protection first emerged with the transition to pastoralism and the need to protect livestock from theft (Smith 1978, p. 16).[34] But, when he turned to agriculture, Smith reverted to emphasize the role of productivity

[34] Smith, though erred in believing that the domestication of livestock preceded farming. In addition, while livestock is vulnerable to theft it is less amenable to feed troops and a bureaucracy, and was thus rarely used as a major tax

as generating surplus, division of labor and trade and, as a result, a demand for an extended role of government (1978, p. 409).[35] For Smith and his intellectual heirs, surplus had to be available before the landlord, the capitalist or the ruler could seize it.

Engels similarly stated that in pre-agricultural societies "Human labor power... yielded no noticeable surplus as yet over the cost of its maintenance" (1978, p. 65), and that it was only the surplus generated by the adoption of agriculture that triggered the transition to a class society. According to Childe (1936), the transition to agriculture resulted in food surplus which enabled individuals to specialize in non-farming activities. This division of labor then led to political integration and eventually to urban centers and to the formation of city-states under a state bureaucracy. Also Lenski (1966) contends that farming technologies generated a surplus that hunter-gatherers could not produce, and that this surplus, which intensified with the transition from horticulture to intensive agriculture, then led to the emergence of social power.[36]

Given our focus on the distinction between tropical and temperate crops, we note that many scholars have sought to explain what lies behind the relative underdevelopment of tropical countries.[37] Adopting the productivity-and-surplus approach, Diamond (1997), concludes (p. 92): "In short, plant and animal domestication meant much more food ... The resulting food surpluses ... were a prerequisite for the development of settled, politically centralized, socially stratified, economically complex, technologically innovative societies." In a recent survey, Price and Bar-Yosef (2010; p. 160) reach a similar conclusion: "Cultivation also supported a stable economy with surplus that resulted in the formation of elite groups as predicted by Lenski (1966)." Nowadays the productivity and surplus explanation seems the default explanation in text books and scientific

base.

[35] The idea that agriculture generated surplus and that surplus led to government was expressed already by earlier seventeenth century social thinkers. See Meek (1976) and Aspromourgos (1996).

[36] From our perspective, Lenski's preoccupation with the distinction between horticulture and agriculture and his underlying presumption that the former is but a prior, less developed, form of the latter, reflects a misperception. Once one realizes that the forms of horticulture witnessed by ethnographers are highly biased towards farming societies that are based on roots and tubers, it becomes clearer that horticulture (especially in the observed societies) and chiefdoms may represent a long-term geographic-conditioned state of affairs, rather than a stage in the transition to agriculture (based mostly on the cultivation of cereals) and to state institutions.

[37] Sachs, Mellinger and Gallup, (2001), Olsson and Hibbs (2005) and Spolaore and Wacziarg (2013) provide empirical attempts to link income per capita across countries with geographic variables. Nowadays, two main features of the tropics are typically argued to have impeded its development: low agricultural productivity and a high burden of disease. Weil (2007, 2010) finds that the effect of health on growth is rather small and cannot explain the extent of the gap between tropical and non-tropical countries. Here we question the productivity explanation, and provide an alternative, geographical/institutional explanation.

writing: Diamond (1997) and Price and Bar-Yosef (2010) are two salient examples.

Another influential variant theory posits that the increased productivity of agriculture accounts for the emergence of hierarchy via the growth of population, rather than through the generation of surplus.[38] Johnson and Earle (2000) claim that rising population agglomeration led to increased conflict and necessitated the reorganization of society into increasingly complex social forms in order to contain violence.[39] This functionalist, Hobbesian theory was adopted by North, Wallis and Weingast (2009), who purport to explain the evolution of human history from the Neolithic age to modern times in terms of the institutions that are formed to contain humans' natural proclivity to violence.[40] Motivated by the contrasting political structures in the valleys of Peru and in Amazonia, Carneiro's (1970) 'circumscription theory' offers a variant of this conflict argument. Postulating that an elite can extract ongoing surplus only when the subjects of taxation are geographically entrapped, Carneiro contends (p. 735) that states could not emerge in the Amazon Basin because "the vanquished could flee to a new locale, subsisting there about as well as they had subsisted before, and retaining their independence." Whereas "in Peru ... this alternative was no longer open to the inhabitants of defeated villages. The mountains, the desert, and the sea ... blocked escape in every direction."

Carneiro's puzzlement over limited social complexity in Amazonia is reminiscent of Diamond's concern with regard to the underdevelopment of New Guinea. We note, however, that the environmental theory of one is inconsistent with the geographical evidence of the other. Diamond's

[38] Other scholars who find fault with the surplus theory point out that an increase in productivity may be dissipated in various ways without creating surplus. Pearson (1957) contended that cultural needs would evolve to eliminate any surplus. Sahlins (1972) argued that hunter-gatherers could have easily procured food beyond their immediate needs, but deliberately refrained from doing so by preferring leisure. He inferred that the first farmers could have similarly responded to increased productivity by working less hard. Sahlins concluded (p. 140): "Leadership continually generates domestic surplus," claiming (like us) that it was hierarchy that generated surplus and not vice versa. Sahlins, however, did not resolve the key questions: what accounts then for the rise of hierarchy and why did its emergence correlate with agriculture?

[39] We note however, that, as observed above, agglomeration in villages accompanied the transition to cereal farming even at the pre-domestication stage, and probably preceded the increase in population density that by and large followed the domestication of plants and animals. From our perspective, the early tendency for agglomeration was a result from a need to protect crop stockpiles. Moreover, we also note that in the test-case of New Guinea, increased land productivity and population growth did not give rise to the type of institutions envisioned by these models, nor did they decrease the prevalence of violence.

[40] Our approach can be considered to be neo-Hobbesian, in the sense of combining the functional role of government in protecting individuals from theft and banditry, with the recognition that this need for protection arose simultaneously with the non-functional increased ability of the would-be rulers to appropriate. Thus, it avoids teleology. In the spirit of Olson (1993), hierarchy arises because it becomes feasible to tax, and because it serves the elite's interest to protect farmers from expropriation by bandits.

theory about the advantage of an east-west orientation of land mass can hardly resolve Carneiro's comparison between Peru and Amazonia, and Carneiro's theory fails to resolve Diamond's concern about limited social complexity in the Pacific tropical islands. The appropriability theory offers a consistent explanation: whereas agriculture in the tropical Amazon and the Pacific islands was based on tuber crops, farming in the western valleys of the Andes relied mostly on maize. The formation of the Mayan state societies in the non-circumscribed tropical lowlands of Mexico, where maize was first domesticated and became the staple crop, provides additional support for the theory on the preponderant role of cultivating cereals in the emergence of hierarchy.

Dow and Reed (2013) provide another variant of conflict theory, suggesting that hierarchy emerged after those who gained control over land excluded outsiders, and employed them as workers. Boix (2015) suggests that the introduction of agriculture caused inequality between inside farmers, who were able to benefit from the new productive technology, and outsiders, leading the latter to raid the farmers. As he has it, this conflict ended either in dictatorships by the outside bandits who turned stationary (see Olson, 1993), or in republics managed by the insiders themselves. Finally, Dal Bó, Hernández and Mazzuca (2015) theorize that farmers' increased insecurity due to pillage by outsiders discouraged investment, and this inefficiency was only resolved through the development of state defense capacity.

Other theories invoke alternative functional explanations for the coincidence between the emergence of farming and hierarchy. Wittfogel's (1957) influential theory contends that strong despotic hierarchies were required in riverine environments in order to realize agricultural potential through the public construction and management of large irrigation projects. Indeed, Bentzen, Kaarsen and Wingendr (2017) show empirically that environments with the potential for irrigation systems have had greater inequality in the past, as well as more authoritarian states in the present. But Wittfogel's many critics point out that irrigation projects in early civilizations were constructed by local communities, prior to the emergence of a strong central state, and were also typically managed locally rather than centrally.[41]

[41] Mayshar, Moav, and Neeman (2017) contend that, in contrast to Wittfogel's original causal theory, it is not that a need for irrigation led to a despotic state, but rather that irrigation systems enabled control and expropriation by the central state – analogous to the interpretation here that food storage facilitated confiscation. The findings by Bentzen et al. are consistent with our appropriability approach, since they do not address the direction of causality on whether hierarchy preceded or followed irrigation.

Another functional theory focuses on the demand for law and order to facilitate trade. On the basis of evidence from Africa, Bates (1983) argues that ecologically diverse environments increase the returns from trade and thus generate a demand for hierarchy.[42] Fenske (2014) and Litina (2014) provide evidence for this theory. We interpret their findings as consistent with the appropriability approach, since trade also facilitates taxation. Our focus here, however, is on societies in which agriculture was the predominant potential tax base.

Finally, a number of scholars reverse the standard causal direction and maintain that hierarchy preceded agriculture, and actually led to agriculture. Cauvin (2000) argues that the willingness of hunter gatherers to abandon their traditional ways of life and to engage in farming was conditioned by a prior change in collective psychology and by the rise of centralized religion. Acemoglu and Robinson (2012, pp. 139-142) suggest that an institutional innovation among the semi-sedentary Natufians in the ancient Near East enabled a political elite to gain power, and then, in effect, to cause "the transition first to sedentary life and then to farming" (p. 140). In suggesting that hierarchy was the cause of surplus, rather than its consequence, this theory resembles ours. However, it is diametrically different in that we seek to explain the emergence of hierarchical institutions, while taking the transition to farming as given.[43]

We conclude this brief survey of the literature with theories that differentiate between alternative farming forms and relate these differences to social institutions. Nunn and Qian (2011) show how the adoption of the potato in Europe in the mid-sixteenth century led to population growth and to substantial social changes. They argue that these changes were due to the higher caloric yield of the potato in regions that are highly suitable for its cultivation. Our current perspective leads us to suggest a complementary mechanism whereby European farmers adopted the potato because it provided them with greater immunity against taxation/theft, which translated into growth of the farming population. Consistent with this mechanism, Mcneill (1999, pp 71-72) reports that European farmers resisted adopting the potato after it reached Europe, yet during the Dutch Wars

[42] Algaze (2008) offers a similar theory with regards to ancient southern Mesopotamia.

[43] In support of their theory of the precedence of political innovations, Acemoglu and Robinson (2012) review also the case of the Bushong people in the Congo, who became the subjects of a kingdom in the seventeenth century by an entrepreneurial leader who imported and developed maize farming technology. Without contesting the pertinence of this example, we note, in the spirit of Scott (2009, 2017), that their neighboring Lele people, who share the same environment, avoided the cultivation of cereals and were able to resist subjugation to hierarchy – either to the Bushong, or to a copycat leader of their own.

in 1557-1609, "villagers along the route [of the Spanish army] swiftly discovered that by leaving the tubers in the ground and digging them only as needed for their own consumption, they could safely survive even the most ruthless military requisitioning. Foraging parties were unwilling to dig for their food when stores of grain were available in barns." Mayshar, Moav and Neeman (2017) study another aspect of appropriability, arguing that geographical attributes that contribute to the transparency of cereal farming may alleviate principal-agent problems and facilitate more onerous taxation, and increased state capacity.[44] Finally we note that the empirical approach adopted here is similar to that of Alesina, Giuliano and Nunn (2013), who demonstrate how early farming techniques, and in particular the use of the plow, impact current perceptions of gender roles.

6 Concluding Remarks

The prevailing scholarly view attributes the emergence of hierarchy to the increased productivity of agriculture. It is commonly presumed that this productivity increase generated food abundance, which in turn led to population increase, specialization in crafts, trade, and the rise of elite. We do not challenge the perception that the transition away from egalitarianism towards hierarchy was correlated with the shift to agriculture and to higher productivity. But we contend that the causal mechanism that relates agriculture to hierarchy may have had little to do with the increase in productivity. Noting that states failed to develop in regions that cultivate roots and tubers, we examine here an alternative theory that the key feature of the Neolithics that brought about the rise of a non-food producing elite was farmers' increased vulnerability to appropriation, as a result of their reliance on seasonal stored cereals.

This claim is consistent with the observation that rudimentary hierarchies also emerged among some hunter-gathering societies that relied on storing seasonal food sources. This theory accounts also for cross regional differences in the extent of hierarchical institutions between temperate regions,

[44] In that paper, Mayshar, Moav and Neeman purport to explain key institutional differences between the cereal-dependent states in Egypt, Southern Mesopotamia, and Northern Mesopotamia. De la Sierra (2013) presents alternative evidence on the significance of transparency for appropriability. Studying the mining regions in the modern Democratic Republic of Congo, he shows that a rise in the price of the substance coltan, which is produced from bulky and transparent ores, led to the monopolization of power and to the cessation of conflict between rival armed groups in the coltan-rich regions, whereas an increase in the price of gold, which is easier to conceal and hence less transparent, did not. Similarly, Buonanno et al. (2015) show the effect of a rise in the price of sulphur on the emergence of the Sicilian Mafia.

whose geographical conditions are particularly suited to growing appropriable cereals, and other regions - particularly in the tropics - where the advantage of cereals over roots and tubers is minimal.

The main testable prediction of our theoretical model is that the key variable that accounts for the emergence and persistence of hierarchy in farming societies is not absolute land productivity, but rather sufficient productivity advantage of cereals over roots and tubers. The nature of the question at hand and the available data do not provide an empirically fool-proof test for the correctness of the appropriability theory and our critique of the conventional productivity theory. Nevertheless, in addition to the multiple supporting considerations presented in the first two sections, the findings of our many empirical investigations clearly support the appropriability theory and not the productivity theory.

To conclude, we note that Acemoglu, Johnson and Robinson (2002, 2012) question the role of geography in accounting for current income disparities. They argue that the underdevelopment of countries close to the equator is an outcome of the extractive institutions that happened to be established there by colonialist powers. Besley and Persson (2009, 2014), on the other hand, contend that underdevelopment is closely related to low fiscal capacity, which might be overcome by investment in fiscal administration.[45] Our conclusions here have significant implications for this debate over what accounts for the underdevelopment of tropical regions. Whereas conventional theories suggest that it is low agricultural productivity, our theory suggests that the root cause for this underdevelopment in the tropics is the relatively high productivity of crops and alternative food sources that provide the population with substantial immunity against appropriation, and that have thus contributed to low fiscal capacity and delayed the formation of hierarchical states.

[45] Gennaioli and Voth (2015) emphasize how investment in state capacity since the Middle Ages responded to conflict. Dincecco and Prado (2012) and Dincecco and Katz (2014) show that state capacity is persistent, and has a positive effect on economic performance.

References

[1] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001) “The Colonial Origins of Comparative Development: An Empirical Investigation,” American Economic Review, 91, 1369-1401.

[2] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2002) “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution,” Quarterly Journal of Economics 117: 1231-1294.

[3] Acemoglu, Daron and James A. Robinson (2012), Why Nations Fail: The Origins of Power, Prosperity, and Poverty, Random House.

[4] Alesina, Alberto, Giuliano, Paola and Nunn Nathan (2013) “On the Origins of Gender Roles: Women and the plow,” Quarterly Journal of Economics, 128, 155–194.

[5] Algaze, Guillermo (2008), Ancient Mesopotamia at the Dawn of Civilization: The Evolution of an Urban Landscape, University of Chicago Press.

[6] Allen, Robert C. (1997). “Agriculture and the Origins of the State in Ancient Egypt.” Explorations in Economic History 34, 135-154.

[7] Ashraf, Quamrul H. and Oded Galor (2011) “Dynamics and Stagnation in the Malthusian Epoch,” American Economic Review, 101, 2003-2041.

[8] Ashraf, Quamrul H. and Oded Galor (2013) “The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development,” American Economic Review 103(1), 1-46.

[9] Aspromourgos, Tony (1996), On the Origins of Classical Economics: Distribution and Value from William Petty to Adam Smith, Routledge.

[10] Bar-Yosef, Ofer and Richard H. Meadow, (1995), “The Origins of Agriculture in the Near East,” in Last Hunter–First Farmers: New Perspectives on the Prehistoric Transition to Agriculture, edited by T. Douglas Price and Anne Brigitte Gebauer, School of American Research Press, 39-94.

[11] Bates, Robert H. (1983) Essays on the Political Economy of Rural Africa, University of California Press

[12] Bentzen, Jeanet S, Kaarsen, Nicolai, and Asger M. Wingender (2017) “Irrigation and Autocracy,” Journal of the European Economic Association, 15(1), 1-53.

[13] Besley, Timothy and Torsten Persson (2009) "The Origins of State Capacity: Property Rights, Taxation and Politics," American Economic Review, 99, 1218-1244.

[14] Besley, Timothy and Torsten Persson (2014) "Why Do Developing Countries Tax So Little?" Journal of Economic Perspectives, 28, 99-120.

[15] Bockstette, Valerie, Areendam Chanda and Louis Putterman, (2002) "States and Markets: The advantage of an Early Start", Journal of Economic Growth, 7, 347-369.

[16] Boix, Carles (2015) Political Order and Inequality: Their Foundations and Their Consequences for Human Welfare, Cambridge University Press.

[17] Borcan, Oana, Ola Olsson and Louis Putterman, (2014) "State History and Economic Development: Evidence from Six Millennia." Brown University - Department of Economics -Working Paper Series.

[18] Buonanno P., Durante, R., Prarolo G. and Vanin, P. (2015) "Poor Institutions, Rich Mines: Resource Curse and the Origins of the Sicilian Mafia, Economic Journal, 125, 175-202.

[19] Carneiro, Robert (1970) "A Theory of the Origin of the State," Science, 169, 733–738.

[20] Cauvin, Jacques, (2000) The Birth of the Gods and the Origins of Agriculture, translated by Trevor Watkins, Cambridge University Press.

[21] Childe, V. Gordon (1936) Man Makes Himself, Watts & Co.

[22] Chiwona-Karltun, Linley, Chrissie Katundu, James Ngoma, Felistus Chipungu, Jonathan Mkumbira, Sidney Simukoko, and Janice Jiggins. (2002) "Bitter cassava and women, an intriguing response to food security." LEISA Magazine, 18, 1–15.

[23] Conley, Tim (1999), "Estimation with Cross Sectional Dependence," Journal of Econometrics 92(1), 1-45.

[24] Cook, James (1784), A Voyage to the Pacific Ocean Undertaken by the Command of His Majesty for making Discoveries in the Northern Hemisphere, volume II, London.

[25] Dal Bó, Ernesto, Pablo Hernández and Sebastián Mazzuca (2015) "The Paradox of Civilization: Pre-Institutional Sources of Security and Prosperity," NBER working paper 21829.

[26] De la Sierra, Raul S. (2013), On the Origin of States: Stationary Bandits and Taxation in Eastern Congo. Working paper.

[27] Denham, Tim (2011) "Early Agriculture and Plant Domestication in New Guinea and Island Southeast Asia," Current Anthropology, 52, S379-S395.

[28] Diamond, Jared (1997) Guns, Germs, and Steel: The Fates of Human Societies, Norton, New York.

[29] Dincecco, Mark and Mauricio Prado (2012) "Warfare, Fiscal Capacity, and Performance," Journal of Economic Growth, 17, 171-203.

[30] Dincecco, Mark and Gabriel Katz (2016) "State Capacity and Long-Run Economic Performance," The Economic Journal, 126 (590), 189-218.

[31] Dow Gregory K. and Clyde G. Reed (2013) "The Origins of Inequality: Insiders, Outsiders, Elites, and Commoners," Journal of Political Economy, 121, 609-641.

[32] Fenske, James (2013) "Does land abundance explain African institutions?" Economic Journal 123, 1363-90.

[33] Fenske James (2014) "Ecology, trade and states in pre-colonial Africa," Journal of the European Economic Association, 12, 612–640.

[34] Galor Oded and Omer Ozak, (2016),"The Agricultural Origins of Time Preference" American Economic Review, 106 (10). 3064–3103

[35] Gennaioli, Nicola and Ilia Rainer (2007) "The Modern Impact of Pre-Colonial Centralization in Africa." Journal of Economic Growth, 12, 185-234.

[36] Gennaioli, Nicola and Hans-Joachim Voth (2015) "State Capacity and Military Conflict," Review of Economic Studies, 82(4), 1409-1448.

[37] Heizer, R.F. and Sturtevant, W.C. (1978) Handbook of North American Indians, Smithsonian Institution, Washington, D.C.

[38] Johnson, Allen W and Earle, Timothy (2000) The Evolution of Human Societies: from Foraging Group to Agrarian State. Stanford University Press; Second edition.

[39] Kuijt, Ian and Bill Finlayson (2009) "Evidence for Food Storage and Predomestication Granaries 11,000 Years Ago in the Jordan valley," Proceedings of the National Academy of Science, 106, 10966-10970.

[40] La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny. (1998). "Law and Finance," Journal of Political Economy 106, 1113-1155.

[41] Larson, G. et al. (2014). "Current perspectives and the future of domestication studies." Proceedings of the National Academy of Sciences 111(17), 6139–6146.

[42] Lenski, Gerhard, E. (1966) Power and Privilege: A Theory of Stratification, McGraw-Hill.

[43] Litina, Anastasia (2014) “The Geographical Origins of Early State Formation,” Working Paper.

[44] Matranga, Andrea (2017), “The Ant and the Grasshopper: Seasonality and the Invention of Agriculture.” Working paper.

[45] Maurer, Stephan, Jorn-Steffen Pischke and Ferdinand Rauch (2016), “Of Mice and Merchants: Trade and Growth in the Iron Age,” mimeo.

[46] Mayshar, Joram, Omer Moav, and Zvika Neeman, (2017) “Geography, Transparency and Institutions.” American Political Science Review, 111(3), 622-636.

[47] McNeill, William H. (1999) “How the Potato Changed the World’s History”, Social Research, 66, 67-83.

[48] Meek, Ronald L. (1976) Social Science and the Ignoble Savage, Cambridge University Press.

[49] Michalopoulos, Stelios and Elias Papaioannou (2013) “Pre-Colonial Ethnic Institutions and Contemporary African Development.” Econometrica, 81, 113-52.

[50] Michalopoulos, Stelios and Elias Papaioannou (2014) “National Institutions and Subnational Development in Africa,” Quarterly Journal of Economics, 129, 151-213.

[51] Miracle, Marvin P. (1966) Maize in Tropical Africa. University of Wisconsin Press.

[52] Mithen, Steven J. et al. (2011) “An 11,600 Year-old Communal Structure from the Neolithic of Southern Jordan,” Antiquity, 85, 350–364.

[53] Murdock, George P. (1959) Africa: its Peoples and their Culture History. New York, McGraw-Hill.

[54] Murdock, George P. (1967) Ethnographic atlas, Pittsburgh, University of Pittsburgh Press.

[55] Murdock, George P., Douglas R. White (1969) “Standard Cross-Cultural Sample.” Ethnology 9: 329–369.

[56] North, Douglass C., John Joseph Wallis and Barry R. Weingast (2009) Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History, Cambridge University Press.

[57] Nunn, Nathan, (2008) “The Long-Term Effects of Africa’s Slave Trades.” Quarterly Journal of Economics 123(1), 139-176.

[58] Nunn, Nathan and Nancy Quian, (2011) “The Potato’s Contribution to Population and Urbanization: Evidence from an Historical Experiment.” Quarterly Journal of Economics 126, 593-650.

[59] Nunn, Nathan and Diego Puga, (2012) "Ruggedness: the Blessing of Bad Geography in Africa." Review of Economics and Statistics 94, 20-36.

[60] Olson, Mancur (1993) "Dictatorship, Democracy, and Development," American Political Science Review, 87, 567-576.

[61] Olsson, Ola and Douglas A. Hibbs Jr. (2005) "Biogeography and long-run Economic Development," European Economic Review, 49, 909-938.

[62] Pearson, Harry W. (1957) "The economy has No Surplus: Critique of a Theory of Development," in Trade and Market in the Early Empires: Economies in History and Theory, edited by Karl Polanyi, Conrad Arensberg and Harry W. Pearson, Free Press, 320-341.

[63] Price, T. Douglas and Ofer Bar-Yosef (2010) "Traces of Inequality at the Origins of Agriculture in the Ancient Near East," in Pathways to Power: New Perspectives on the Emergence of Social Inequality, edited by T. Douglas Price and Gary Feinman, Springer, 147-168.

[64] Pryor, Frederic. L. (1985). "The Invention of the Plow." Comparative Studies in Society and History 27: 740-744.

[65] Purugganan, Michael D. and Dorian Q. Fuller (2011) "Archaeological data reveal slow rates of evolution during plant domestication," Evolution, 65, 171–183.

[66] Ramankutty, Navin, Jonathan A. Foley, John Norman and Kevin McSweeney (2002). The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. Global Ecology and Biogeography 11(5), 377-392

[67] Richerson, Peter J., Robert Boyd and Robert L. Bettinger (2001) "Was Agriculture Impossible During the Pleistocene but Mandatory During the Holocene? A Climate Change Hypothesis," American Antiquity, 66, 387-411.

[68] Sachs, Jeffrey D., Andrew D. Mellinger, and John L. Gallup (2001) "The Geography of Poverty and Wealth" Scientific American, 284, 71-74.

[69] Sahlins, Marshall (1972) Stone Age Economics, de Gruyter.

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

[71] Scott, James C. (2017) Against the Grain: A Deep History of the Earliest States. Yale University Press

[72] Smith, Adam (1978) Lectures on Jurisprudence, edited by R.L. Meek, D.D. Raphael and P.G. Stein, Oxford University Press.

[73] Spolaore, Enrico and Romain Wacziarg (2013) "How Deep Are the Roots of Economic Development?" Journal of Economic Literature, 51, 1–45.

[74] Taylor, Walter W. (1973) "Storage and the Neolithic Revolution," in Estudios Dedicados a Prof. Dr. Luis Pericot, Barcelona: Publicaciones Eventuales, 23, 193–197.

[75] Testart, Alain (1982a) Les Chasseurs-Cueilleurs ou L'origine des Inégalités, Societe d'Ethnographie.

[76] Testart, Alain (1982b) "The Significance of Food Storage among Hunter-Gatherers: Residence patterns, Population densities, and Social Inequalities," Current Anthropology, 23, 523-537.

[77] Testart, Alain (1982c) "Les Tubercules Sont-ils aux Céréales comme la Sauvagerie est à la Civilisation?" Journal D'agriculture Traditionnelle et de Botanique appliquée, 29, 349-354.

[78] Tuden, A., Marshall, C. (1972) Political Organization: Cross-Cultural Codes 4.Ethnology 11: 436-64.

[79] Tushingham, Shannon and Robert L. Bettinger (2013) "Why Foragers Choose Acorns Before Salmon: Storage, Mobility, and Risk in Aboriginal California," Journal of Anthropological Archaeology, 32, 527-537.

[80] Weidmann, Nils.B., Jan K. Rod, and Lars-Erik Cederman, (2010). "Representing Ethnic Groups in Space: a New Dataset," Journal of Peace Research 47, 491-9.

[81] Weil David (2007) "Accounting for the Effect of Health on Economic Growth," Quarterly Journal of Economics, 122, 1265-1306.

[82] Weil, David (2010) "Endemic Diseases and African Economic Growth," Journal of African Economies, 19, supplement, 1-29.

[83] Whitehouse, David and Ruth Whitehouse (1975) "Archaeological Atlas of the World," Thames and Hudson.

[84] Wiessner, Polly and Akii Tumu (1998) Historical Vines: Enga networks of exchange, ritual, and warfare in Papua New Guinea, Smithsonian Institution Press.

[85] Willcox, George, and Danielle Stordeur (2012), "Large-scale Cereal Processing before Domestication During the Tenth Millennium cal BC in Northern Syria," Antiquity, 86, 99-114.

[86] Wittfogel, Karl A. (1957) Oriental Despotism: A Comparative Study of Total Power, Yale University Press.

For Online Publication

Appendix A: Cereals vs. Roots and Tubers

In this appendix we seek to provide evidence in support of our various factual claims on the distinction between cereals and roots/tubers: (i) that reliance on roots and tubers is a major phenomenon in tropical regions; (ii) that roots and tubers are highly productive in the tropics; (iii) that their harvesting is in general non-seasonal; (iv) that after harvest they are significantly more perishable than cereals; and (v) that there exist significant climatic and soil variations in the productivity of cereals and of roots and tubers. (vi) that there moisture content is very high, making them bulky to transport.

.

Table A.1: Major staple crops produced in regions of the world in 1961

	World 1961		Sub-Sahara 1961		Papua New-Guinea 1961	
	Average Cal Yield (mil Kcal/ha)*	Total Energy Produced (10^{12} kcal)**	Average Cal Yield (mil Kcal/ha)*	Total Energy Produced (10^{12} kcal)**	Average Cal Yield (mil Kcal/ha)*	Total Energy Produced (10^{9} kcal)**
Rice	6.92	798	4.57	12	13.16	6
Wheat	3.78	772	2.39	6		
Maize	7.09	748	3.66	53	3.34	0
Barley	4.68	255	2.79	3		
Oats	3.19	122	2.86	0		
Sorghum	3.02	139	2.53	29	5.62	0
Rye	3.92	119	0.60	0		
Millet	2.24	97	2.17	24		
Soybean	1.66	40	0.55	0		
Potato	9.41	208	5.14	1	5.21	0
Cassava	11.85	114	9.10	50	15.94	84
Sweet Potato	6.32	84	4.44	3	3.44	254
Yam	8.54	10	8.65	9	19.56	142
Taro	6.63	5	0.54	0	7.51	136
Banana	9.47	19	5.21	3	11.17	276
Plantain	6.37	16	5.65	11		
Total of above		3,545		203		898
Population (mil)		3,083		223		2

The average caloric content of each crop is based on the figures in table E.4. For Soybeans, Taro, Bananas and Plantains the respective figures are 1.47, 1.12, 0.89, and 1.22, based on http://ndb.nal.usda.gov/ndb/, accessed Feb 2017. * Calculated on the basis of the caloric content of each crop multiplied by the average land yield, as reported by the FAO (http://faostat3.fao.org/download/Q/*/E accessed Feb 2017). ** Calculated on the basis of the average caloric yield multiplied by the crop area, as reported by the FAO (http://faostat3.fao.org/download/Q/*/E accessed Feb 2015).

Table A1 presents summary data on the main staple crops in sub-Saharan Africa, in Papua New Guinea (PNG) and the world in 1961 – the earliest year for which the Food and Agriculture Organization (FAO) provides the information.[46] In relying on relatively recent data, our presumption is that the soil and climatic conditions have not changed significantly since the Neolithic period. Of course, we recognize that the plants that provide most of the calories that humans consume have undergone major modifications since antiquity, and that their availability was greatly impacted by the post-Columbian migration of species between the continents.

(i) The data in Table A1 reveals that roots and tubers provided 37.9 percent of the total calories produced by the main staple crops in sub-Saharan Africa in 1961. We note that cassava alone provided about 45 percent of the total calories produced by these crops in Nigeria in 2013.

(ii) The table reveals further that the average caloric yield of cassava and yam in sub-Saharan Africa (9.10 and 8.65 mil Kcal/Ha) exceeded the comparable world average yield of the three main cereals, rice, maize and wheat (equal to 6.82, 7.09 and 3.56 mil Kcal/Ha, respectively).

(iii) The table reveals that the yield of non-cereals in PNG is generally high.

(iv) The seasonality of cereals is well known. They have to be sown and reaped in a relatively fixed period in the year, and usually once a year. On the other hand, roots and tubers are generally perennial and may be harvested at any time during the year. In fact, cassava can be left intact in the ground for two years. This gives farmers great flexibility around the timing of the harvest, and prevents the need for significant storage. Rees et al. (2012, p. 394) report: "Harvest time [of cassava] ranges from six to 24 months, and roots can be left in the ground until needed, making cassava a very useful food security crop."[47]

(v) Harvested grains are storable with relatively little loss from one harvest to the next, and even over several years. On the other hand, roots and tubers are in general perishable once out of the ground, though to different degrees. In particular, cassava starts to rot at ambient African temperature within 2-3 days of being harvested. The rapid deterioration is often hastened by abrasions caused by uprooting and transportation. Rees et al. (2012, p. 394) summarize the evidence: "Despite their agronomic advantages over grains, which are the other main staple food crops, root crops are far more perishable. Out of the ground, and at ambient temperatures these root crops have shelf lives that range from a couple of days for cassava. . . , two to four weeks for sweet potato, to between four and 18 weeks for the natural dormancy of yams. . . " Cassava's fast rotting upon harvest can be overcome only by freezing or by laborious processing that turns the

[46] Given a rough estimate of 1 million calories required per person per year (2740 kcal per day), the columns on total energy produced provide a crude estimate of the population (in millions) whose energy needs could be supported by each crop (ignoring the feeding of animals, seed requirements and wastage). It is evident that (other than in PNG) the total energy produced by the listed major crops could roughly feed the entire population – but in PNG it could feed less than half the population.

[47] See also Lebot (2009) and Bradshaw (2010).

moist root into dry flour.

(vi) Lebot (2009) lists the optimal annual rainfall for cassava, yams and sweet potato as ranging from 750 to 1500 mm of rain, and the optimal temperature as 20-30 degrees centigrade. This reveals that while these crops are cultivable in the tropics, they cannot be cultivated in temperate climates.

(vii) According to Lebot (pp. vi, 78) the moisture content of cassava is 63% of the weight, and that of sweet potato and yam is 71% and 74% respectively.

By these considerations, even though the potato is biologically a tuber, when it comes to the degree of appropriability, it could be considered a quasi-cereal, since it is cultivable in temperate climates, is seasonal, and is relatively non-perishable upon harvesting.

References

[1] Bradshaw, J. E. (editor), (2010) *Root and Tuber Crops*, Springer.

[2] Lebot, Vincent (2009) *Tropical Root and Tuber Crops: Cassava, Sweet Potato, Yams and Aroids,* CABI.

[3] Rees, Debbie et al. (2012) "Tropical Root Crops," in *Crop Post-Harvest: Science and Technology – Perishables,* edited by Debbie Rees, Graham Farrell and John Orchard, Blackwell-Wiley, 392-413.

For Online Publication

Appendix B: Surplus and appropriation – the role of population

We develop here a simple model to illustrate our Malthusian critique of the surplus theory for explaining the rise of hierarchy following the Neolithic Revolution. In this model, when population size is exogenous, both an increase in the degree of appropriability and a rise in productivity (generating surplus) lead to larger net tax revenue as a share of output. However, when the population is endogenous, according to the Malthusian framework, an increase in appropriability raises the share of net taxes, while a rise in productivity does not.

Denote the total size of the farming population by N. The production function is assumed to be Cobb-Douglas:

$$Y = (AX)^{\alpha} N^{1-\alpha} = A^{\alpha} N^{1-\alpha},$$

where A denotes the level of technology, X is the constant size of land which we normalize to one, and $0 < \alpha < 1$.

We assume that the cost of taxing a share τ of total income Y is given by:

$$\frac{Y \cdot C(\tau, m)}{z},$$

where m represents per-capita surplus income. The parameter $z > 0$ represents the degree of appropriability, so that a higher z implies a lower cost of taxation. The function $C(\tau, m)$ is continuous and differentiable, and increasing and convex in the tax rate τ. ($C_1 \geq 0, C_{11} > 0$). In adapting the standard surplus approach, we assume that resistance to tax payment is lower the higher is surplus income. As a result, the cost of taxation is assumed to decrease in surplus income, or $C_2 < 0$. Surplus income is:

$$m = (1 - \tau) \left(\frac{A}{N} \right)^{\alpha} - s,$$

where s is subsistence income. The share of total net taxes out of total income, denoted by π, is:

$$\pi(\tau, m, z) = \tau - \frac{C(\tau, m)}{z}.$$

The government chooses the tax rate τ to maximize its net revenue $\Pi = \pi Y$. We assume the existence of an interior solution for the tax rate, τ^*, where the first and second order conditions are satisfied. Our aim is to examine how π is affected by changes in productivity A and in the degree of appropriability z.

B1. The case of a fixed population

Given our assumptions, when the population is constant, Y is independent of τ. The optimal tax rate τ^* thus maximizes π and satisfies the first order condition:

$$\frac{1}{z}\frac{dC(\tau, y)}{d\tau}\bigg|_{\tau=\tau^*} = \frac{C_1(\tau^*, m) - C_2(\tau^*, m)\left(\frac{A}{N}\right)^{\alpha}}{z} = 1.$$

Consider the effect of an increase in the appropriability parameter z. By the envelope theorem:

$$\frac{d\pi(\tau^*, m, z)}{dz} = \frac{\partial \pi(\tau^*, m, z)}{\partial z} = \frac{C(\tau^*, m)}{z^2} > 0.$$

Consider next the effect of an increase in productivity A. By a similar argument:

$$\frac{d\pi(\tau^*, m, z)}{dA} = \frac{\partial \pi(\tau^*, m, z)}{\partial m} \cdot \frac{dm}{dA} = -\frac{C_2(\tau^*, m)}{z} \cdot \frac{\alpha(m+s)}{A} > 0.$$

Thus, we have:

Proposition B1. *With a fixed population, both an increase in appropriability z and an increase in productivity A raise the share of taxes out of income π.*

B2. The case of Malthusian population

In a Malthusian setting the population size adjusts to keep agents' per capita surplus income m at zero. Thus:

$$N = \frac{(1-\tau)Y}{s}.$$

This implies:

$$Y = A\left(\frac{1-\tau}{s}\right)^{\frac{1-\alpha}{\alpha}} \equiv Y(\tau, A);\ m \equiv 0.$$

Denote:

$$\pi^*(\tau, z) \equiv \pi(\tau, 0, z) = \tau - \frac{C(\tau, 0)}{z}.$$

In this case, the tax rate has a negative effect on output through its effect on the size of the farming population N.

The optimal tax rate $\tau^* = \tau^*(z, A)$ maximizes $\Pi = \pi^*(\tau, z)Y(\tau, A)$. Our assumptions imply that it is implicitly defined by the first order condition:

$$F(\tau, z, A) \equiv Y(\tau, A)\frac{\partial \pi^*(\tau, z)}{\partial \tau} + \pi^*(\tau, z)\frac{\partial Y(\tau, A)}{\partial \tau} = Y\left(1 - \frac{C_1(\tau, 0)}{z}\right) - \pi^*(\tau, z)\, Y \frac{1-\alpha}{\alpha(1-\tau)} = 0.$$

Thus, at the optimum τ^*:

$$\frac{\partial \pi^* (\tau, z)}{\partial \tau} = -\frac{\pi^* (\tau, z)}{Y (\tau, A)} \cdot \frac{\partial Y (\tau, A)}{\partial \tau} = \pi^* (\tau, z) \cdot \frac{1-\alpha}{\alpha (1-\tau)} > 0.$$

In addition,

$$\frac{d\pi^* (\tau^* (z, A), z)}{dz} = \frac{\partial \pi^* (\tau^* (z, A), z)}{\partial \tau} \frac{d\tau^* (z, A)}{dz} + \frac{\partial \pi^* (\tau^*, z)}{\partial z} = \frac{\partial \pi^* (\tau^*, z)}{\partial \tau} \frac{d\tau^*}{dz} + \frac{C(\tau^*, 0)}{z^2}.$$

To prove that this expression is positive, it is sufficient to prove that $\partial \tau^* / \partial z$ is positive. By the Implicit-Function Theorem, for $F (\tau, z, A)$ defined above:

$$\frac{\partial \tau^*}{\partial z} = - \left. \frac{\partial F / \partial z}{\partial F / \partial \tau} \right|,$$

and by the second-order conditions: $\partial F / \partial \tau < 0$. Thus,

$$sign \left[\frac{\partial \tau^*}{\partial z} \right] = sign \left[\left. \frac{\partial F}{\partial z} \right| \right].$$

Now,

$$\frac{\partial F}{\partial z} = Y \cdot \frac{C_1 (\tau, 0)}{z^2} + \frac{C (\tau, 0)}{z^2} \cdot Y \cdot \frac{1-\alpha}{\alpha (1-\tau)} > 0.$$

Similarly,

$$\frac{d\pi^* (\tau^* (z, A), z)}{dA} = \frac{\partial \pi^* (\tau^* (z, A), z)}{\partial \tau} \frac{d\tau^* (z, A)}{dA}.$$

Once again by the Implicit Function Theorem: $sign \left[\frac{\partial \tau^*}{\partial A} \right] = sign \left[\left. \frac{\partial F}{\partial A} \right| \right]$. But

$$\frac{\partial F (\tau, z, A)}{\partial A} = \frac{\partial \pi^* (\tau, z)}{\partial \tau} \cdot \frac{\partial Y (\tau, A)}{\partial A} + \pi^* (\tau, z) \cdot \frac{\partial^2 Y (\tau, A)}{\partial \tau \partial A}.$$

Since $\frac{\partial Y(\tau, A)}{\partial A} = \frac{Y(\tau, A)}{A}$ and $\frac{\partial^2 Y(\tau, A)}{\partial \tau \partial A} = \frac{\frac{\partial Y(\tau, A)}{\partial \tau}}{A}$, we have:

$$\frac{\partial F (\tau, z, A)}{\partial A} = \frac{F (\tau, z, A)}{A}.$$

Since the first order conditions require $F (\tau, z, A) = 0$, it follows that $\frac{\partial \tau^*}{\partial A} = 0$ so that

$$\frac{d\pi^* (\tau^* (z, A), z)}{dA} = 0.$$

Thus, we have:

Proposition B2. *With Malthusian population, an increase in appropriability z raises the share of taxes in the economy π, but an increase in productivity A leaves that share intact.*

For Online Publication

Appendix C: Outliers in the Ethnographic Atlas database

We checked the Ethnographic Atlas database for possible outliers. Out of the total Atlas sample of 1412 societies, only 1259 had data on hierarchy above the local level (variable 33). Out of these 1259 societies, only 959 are coded as reliant on specified food sources (variable 29): 634 relied mostly on cereal grains, 259 on roots and tubers, 82 on fruit trees and 4 on vegetables. The Atlas defines a "large state" as possessing at least four levels of hierarchy above the local community, and defines those with three levels as "states" (Gray, 1998). The following table provides the distribution of these societies by the major crop type.[48]

	Total	% Cereal	% Roots & tubers	% Fruit trees	% Other
All societies	979	64.8	26.5	8.4	0.4
Above hierarchy = 4	32	90.6	9.4	0	0
Above hierarchy = 3	82	79.3	13.4	7.3	0

It is evident that societies that rely on grain crops form a significantly larger share of those with higher levels of hierarchy than in the total sample. Still, we were intrigued by the three societies (9.4%) that are coded as "large states" that are reliant on roots and tubers. In this appendix we examine the three apparent exceptions to our statement that all pre-modern large states relied on cereals. Our conclusion is that our statement stands.

1. The first case is identified by Murdock (1981, p. 39) as "Shantung Chinese." The anthropological data were based on Yang's (1945) detailed report about the small village of Taitou, apparently pertaining to the 1930s. The province where this village (which no longer exists) was located is now known as Shandong. Shandong, on the eastern edge of China (in between Beijing and Shanghai) is now one of the more prosperous provinces. There is little doubt that the village of Taitou was at the time part of a major state: China. We take issue, however, with the presumption that this state, or even that province, should be classified as based on roots and tubers, as the Ethnographic Atlas has it.

To begin with, Yang reports (1945 p. 16) that the main crops cultivated in the village of Taitou were "wheat, millet, barley, soybean, corn, sweet potatoes, and peanuts." Thus it is not clear why the village was coded as reliant on the intensive cultivation of roots, even though this may not be entirely wrong. According to Yang, the land of the village was multi-cropped, and sweet potatoes

[48] All of the societies with maximal hierarchies had a specified major crop type, and two societies with above hierarchy coded as 3 had unspecified crop type (thus making the total 84, instead of 82).

were the principal crop during the summer months. More precise information can be obtained from the data collected by Buck (1937). Buck provides detailed statistics from another village, Tsimio, located about forty kilometers from Taitou. According to these data (pp. 73, 75), sweet potatoes supplied 43.3 percent of the total calories consumed in Tsimio, and seeds supplied 54.4 percent of the calories. But Tisimio (and presumably also Taitou) was unrepresentative of Shandong as a whole. In the other eleven villages that entered Buck's sample from the province, one listed the percentage of calories supplied by seeds as 76.0 (with sweet potatoes supplying 20.8 percent), and all of the other ten villages listed seeds as supplying at least 91 percent of the calories consumed. We conclude that the coding of the "shantung Chinese" society as reliant on the intensive cultivation of roots is wrong.

But the problem at hand goes beyond coding. The Ethnographic Atlas focuses mostly on cultural issues such as kin relations, with the underlying presumption that the data from one location pertains for the society as a whole. Yet the number of levels of hierarchy presiding over any location depends mostly on the region that encompasses that specific location. Even if Shandong's farming were based solely on roots, it could easily belong to a complex polity whose hierarchy relies on taxing cereals. This is evidently the case for China, whose state relied from the very start on the cultivation of cereals: mostly on wheat in the north (where Shandong is located), and mostly on rice in the south. Furthermore, sweet potatoes reached China only after the Columbian Exchange, while Shandong was already an integral part of the Chinese hierarchical state almost two millennia earlier.

2. The data source for "Byelorussians (White Russian)" society (latitude: 55N, longitude 28E) is identified in Ethnology, (vol. 4, April 1965) as "Unpublished ethnographic notes" by Melvin Ember (1954), pertaining to 1910. We were unable to obtain this source, but employed alternative sources about the Belarus.[49]

Vakar (1956, p. 30) describes the territory as a "low marshy plane, sloping slightly to the south and east. Only in the north and northeast are there elevation of the soil," with the highest point at 345 meters above sea level. "In general ... the soil ... is not very favorable for cultivation. Forest and shallow lakes cover more than one half the whole land. ... Crops are liable to be damaged by humidity rather than droughts. ... fish and potatoes being the staff of life for most Belarusian peasants." This characterization is confirmed by the soil-suitability data according to which the soil suitability for agriculture in Belarus is in the 25th percentile below the average, and

[49] The following summary is based mostly on Vakar (1956), and The Great Soviet Encyclopedia (GSE), vol. 3, 1973.

the productivity of cereals is below average. The fact that potatoes became the main cultivated crop from the 16th century is thus not surprising. In 1913 the land yielded 4.0 million tons of potatoes and only 2.6 million tons of all cereal grains; at the same time, "livestock raising is the leading branch of agriculture in Byelorussia" (GSE, p. 627). Thus, we do not question the Atlas's coding of Belarus as based on roots and tubers (at least since Columbus).

The multiple rivers that cross Belarus provided the main waterway for early trade between the Black Sea in the south and the Baltic Sea in the north. The area was settled between the 6th and the 8th centuries by East Slav tribes, whose economy was apparently based on trade in forest products like furs, game, honey, beeswax and amber and on primitive (cereal) agriculture. As it turns out, almost throughout its history, the territory of Belarus was ruled from outside by its neighboring states, rather than from within. In the 10-12 centuries it was controlled from Kiev (now in the Ukraine) in the south, under the Grand Duke of the Kieven Rus' – but several local feudal vassal principalities existed (all centered in the relatively elevated north).[50] From the 13th until the 18th centuries, the territory was ruled from the north and the west: first by the Grand Duchy of Lithuania, then by the Polish-Lithuanian Commonwealth, and effectively by Poland. Finally, from 1795 until 1990 Byelorussia was ruled from the east, as part of the Russian Empire and later as a Republic within the USSR. Thus we don't question the coding of Belarus society as subjected to full state hierarchy, however, we note that as a dependent territory almost throughout its history, its form of agriculture can hardly be considered to have shaped the degree of complexity of the states that controlled it. Indeed, Borcan et al. (2014) assign Belarus a statehood score of 0.5 from 850 until 1950 – where, in their scoring system "Band/tribe is marked by a rule score of 0, paramount chiefdom is assigned 0.75 and fully-fledged state receives the value 1."

3. The third case concerns the Bubi society on the tropical island of Fernando Po (now called Bioko, and part of the state of Equatorial Guinea). This volcanic island, whose area is about 2,000 square kilometers, is located in the Atlantic Ocean, about 30 kilometers off the coast of present Cameroon. The Bubis' staples were yams and cocoyams (=taro), in addition to fish and game. Thus, also in this case we do not contest the Atlas identification of this society as reliant on roots and tubers. However, we find problematic the characterization of this society (presumably in 1920) as a large state society.[51] Indeed, according to the Atlas, the form of agriculture is not extensive,

[50] Some historians identify the second half of the 11th century as the "first Independent Belorussian State," under the prince ruling in Polock (=Polacak= Polatsk) – see Vakar (1956) and Plokhy (2006 pp. 12, 46-54). If so, Belarus regained independence only with the breakup of the USSR in 1991.

[51] One indication for the problematic identification of the Bubi society as a "large state" is its lack of urban centers. Of the 32 Atlas societies coded as "large states," 27 had cities with population exceeding 50,000 people, 3 more had towns with 5,000 or more people, and one (the Siamese) had towns in the range of 1,000-5,000 people. The Bubi

but rather "extensive or shifting agriculture." Written sources about the island by missionaries and travelers since the 19th century are surprisingly abundant. The Atlas was coded on the basis of several such sources (Ethnology vol. 5, Jan 1966 pp. 132-133). For our purposes of examining mostly the Bubis' political institutions, we find it expedient to rely mostly on Sundiata (1994, 1996, 2011).

The island of Fernando Po was a Spanish colony, but the British set a station in its northern coast from 1827 until 1943. The colonists (including freed slaves that they brought to the island) settled in the coast, while the native Bubis concentrated in the interior hilly slopes of the island where they managed their own affairs. According to Sundiata (1996 p. 7), in 1912 the island's Bubi population numbered only 6,800.[52] The coastal settlers traded with the Bubis by purchasing from them palm oil for export, in exchange for various imported wares (like knifes).

The Bubis were divided in the 19th century into 28 districts, with several villages in each district. The districts often fought with one another over women and over imported objects (Sundiata, 1996 pp. 75-79). According to Sundiata (1996 p. 75), "in spite of their fairly simple material culture, the nineteenth-century Bubi had an elaborate system of socioeconomic stratification." Within chiefdoms, individuals were distinguished as nobles or as commoners, as well as by age groups. The chiefs' role was mostly in settling disputes, imposing and collecting fines, and leading fights.

The long-standing political fragmentation of the Bubis underwent substantial change in the 1840s. Sundiata (1996, p. 80-82) attributes this development to recurring incidents after the British left the island, whereby foreign interlopers from the coasts raided the Bubis inland. As a result, Moka, the chief of one of the southern districts, was able to form a loose confederacy of the various chiefs and was recognized as "Great Chief," (or king) with the task of mediating intra-ethnic strife and containing the incursions.[53] The island under Moka might be considered a hierarchical kingdom.[54] We doubt that this should be considered a case of a stable state society based on the

society is coded as "missing data," but in fact had not had villages exceeding 1,000 people.

[52] An 1837 report by British adventurers states "The native population of Fernando Po may be about five thousand, divided into tribes, that were formerly constantly at war with each other" (Laird and Oldfield, p. 301). That report also notes the Bubis' "free and independent bearing" and adds that "they still go perfectly naked" (p. 302). The missionary Clarke (1848, p. v) estimates the island's total population (including the non-Bubis) at "20,000 souls."

[53] According to Sundiata (1996 p.80-81), Moka "had seventy wives and concubines" and reinforced his position by exercising magical powers. He refused to be seen by any Europeans and "limited his use of European wares to guns and machetes," scorning "the use of European cloth, rum, salt and tobacco."

[54] Baumann (1888), a German traveler, provides a firsthand report of Moka's "kingdom" in 1886. He describes Moka's capital village as consisting of scattered hut complexes, surrounded by fields of yam (p. 103). He also narrates how, on the way there, a Bubi chief asked if he would become their king and resist Moka, telling him how much they resented being subjected to Moka (p. 33).

cultivation of tubers, but rather than an ad hoc confederacy of chiefdoms, motivated primarily by external issues related to warfare and trade (as was the case with several earlier African kingdoms). The fragility of the Bubis' presumed state is evident in that little of it survived into the twentieth century, after cocoa plantations (manned by imported laborers) started to dominate the Island's economy and the Spanish colonial power asserted its rule on the Island.

References

[1] Baumann, Oscar (1888) Eine Afrikanische Tropen-Insel: Fernando Póo und die Bube, Eduard Hölzel.Borcan, Oana, Ola Olsson and Louis Putterman, (2014) "State History and Economic Development: Evidence from Six Millennia," Brown University, Department of Economics Working Paper Series.

[2] Buck, John Lossing (1937) Land Utilization in China: Statistics, Shanghai, North-China Daily News.

[3] Clarke, John (1848), Introduction to the Fernandian Tongue.

[4] Gray, J. Patrick (1998), "Ethnographic Atlas Codebook," World Cultures, 10 86-136.

[5] The Great Soviet Encyclopedia (1973) "Byelorussian Soviet Socialist Republic," volume 3, translation of third Edition [1970], New York, Macmillan.

[6] Laird, Macgregor and R.A.K. Oldfield (1837) Narrative of Expedition into the Interior of Africa, London, Richard Bentley.

[7] Plokhy, Serhii (2006) The Origins of the Slavic Nations: Premodern Identities in Russia, Ukraine, and Belarus, Cambridge University Press.

[8] Sundiata, Ibrahim (1994) "State Formation and Trade: The Rise and Fall of the Bubi Polity, c.1840-1910," The International Journal of African Historical Studies, 27, 505-523.

[9] Sundiata, Ibrahim K. (1996) From Slaving to Neoslavery: the Bight of Biafra and Fernando Po in the Era of Abolition, 1827-1930, University of Wisconsin Press.

[10] Sundiata, Ibrahim (2011) "Moka of Bioko (late 1820s-1899): The Chief Who United a Central African Island" in The Human Factor in Modern Africa, edited by Dennis Cordell, Rowman and Littlefield, pp. 47-66.

[11] Yang, Martin C (1945) A Chinese Village, New York, Columbia University Press.

For Online Publication

Appendix D: Risk-Averse Farmers

In this appendix we illustrate the robustness of the model's qualitative predictions when farmers are risk averse. The results are in a sense even stronger, given that risk-averse farmers under anarchy seek more protection by choosing a smaller share of cereals. Farmers' risk aversion does not affect the analysis of the model under a regime of hierarchy since in this case the tax rate that the state imposes is certain. We chose to illustrate the case of anarchy with risk-averse farmers by examining a case where a simple analytic solution can be obtained. For that purpose, we employ the specification of the expropriation function, $\tau(\lambda) = \rho\sqrt{\lambda}$, as in the model's example, and consider the case where farmers have a log-utility function: $u(I) = \log(I)$. Farmers under anarchy thus chose $\beta \geq 0$ to maximize the expected utility:

$$U(I) = (1-\tau)\log\left(\beta + (1-\delta)(1-\beta)\right) + \tau \log(1-\delta)(1-\beta).$$

The solution is

$$\beta_A = \max\left\{\frac{\delta - \tau}{\delta}, 0\right\}.$$

Non-farmers' freedom to enter banditry implies: $s = \tau\beta/\lambda(\tau)$. And thus:

$$\tau_A = \frac{\rho^2 \beta_A}{s}.$$

Solving for the equilibrium values of (β_A, τ_A) yields (when $\beta_A > 0$):

$$\beta_A = \frac{s\delta}{\rho^2 + s\delta};\ \tau_A = \frac{\rho^2\delta}{\rho^2 + s\delta}.$$

Inspection of the equilibrium values of (β_A, τ_A) reveals that as δ tends to zero, both β_A and τ_A tend to zero. As δ increases towards one, τ_A approaches $\rho^2/(\rho^2 + s)$ and β_A approaches $s/(\rho^2 + s)$. This implies that even in the limit, when the productivity of tubers approaches zero, they are still grown by farmers.

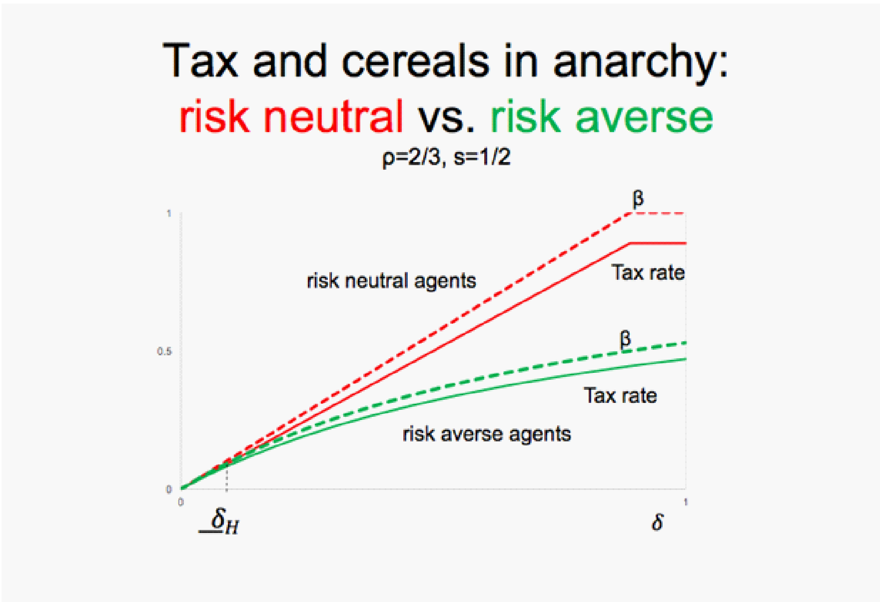
Compared to the model with risk neutrality (in the preceding sub-section), the introduction of risk aversion implies that farmers reduce the cultivation of cereals β_A, and increase the share of land devoted to tubers as a device for self-insurance. Consequently the confiscation rate τ_A is lower, and the measure of banditry λ_A is smaller as well.

While the former effect tends to increase overall inefficiency, the total efficiency effect of introducing risk aversion in a regime of anarchy is positive. To recall from corollary 1, under risk neutrality the overall inefficiency $(1-\beta_A)\,\delta+s\lambda_A$ is equal to δ. This is smaller than the inefficiency under risk aversion, which under our specification is equal to $(1-\beta_A)\,\delta+\lambda_A s=\delta-\beta_A\,(\delta-\tau_A)<\delta$. Correspondingly, the expected income of each farmer under anarchy is also higher under risk aversion, because

$(1-\tau_A)\,(\beta_A+(1-\delta)\,(1-\beta_A))+\tau_A\,(1-\delta)\,(1-\beta_A)=1-\delta+(\delta-\tau_A)\,\beta_A$ is equal to $1-\delta$ under risk neutrality, but is strictly larger under risk aversion because under risk aversion $\tau_A<\delta$. The reason for this is that under risk neutrality farmers in a mixed equilibrium are indifferent between growing cereals and tubers and so derive an identical income of $1-\delta$. In contrast, under risk aversion, farmers derive a strictly larger expected income from cereals to compensate for the risk associated with cereals, which pushes their expected income higher.[55] The figure illustrates the difference between the two types of equilibrium: the case of risk neutral farmers and risk averse farmers.

[55] This implies that risk neutral farmers would benefit if they could commit to grow less cereals in equilibrium, which we assume they cannot. The problem is that when a farmer decides how much cereal to grow, he ignores the negative externality this imposes on other farmers through contributing to the measure of bandits.

Figure E.1: Output: Anarchy vs. Hierarchy



For Online Publication

Appendix E: Data Tables and Additional Evidence

Table E.1: Descriptive Statistics: societies in Ethnoatlas

	SOURCE	**Mean**	**p50**	**SDev**	**Min**	**Max**	**N**
Hierarchy	Ethnoatlas	1.89	2.00	1.04	1.00	5.00	1,059
CerMain	Ethnoatlas	0.54	1.00	0.50	0.00	1.00	1,092
Dependence on agriculture	Ethnoatlas	0.45	0.50	0.27	0.03	0.93	1,178
Dependence on agriculture	Ethnoatlas	0.14	0.00	0.35	0.00	1.00	863
Farming surplus	Tuden and Marshall (1972)	0.49	0.00	0.50	0.00	1.00	162
Population density (categorical)	Pryor (1985)	3.83	4.00	1.57	2.00	7.00	168
LandProd (std)	authors	0.00	0.23	1.00	-1.92	2.66	1,179
CerAdv (std)	authors	0.00	-0.13	1.00	-1.73	4.16	1,179
Precipitation (std)	FAO-GAEZ	0.00	-0.13	1.00	-1.39	10.65	1,179
Temperature (std)	FAO-GAEZ	0.00	0.37	1.00	-2.57	1.32	1,179
Elevation (std)	GDEM	0.00	0.17	1.00	-9.24	3.58	1,179
Ruggedness (std)	GDEMs	0.00	-0.35	1.00	-0.90	6.41	1,179
Absolute Latitude (std)	Ethnoatlas	0.00	-0.43	1.00	-1.21	3.36	1,179
Distance to major river (std)	Fenske (2013)	0.00	-0.63	1.00	-0.63	1.58	1,179
Distance to coast (std)	Fenske (2013)	0.00	-0.30	1.00	-1.11	3.14	1,179
Pct Malaria	MAP	0.17	0.06	0.21	0.00	0.69	1,179
Population density 1995 (std)	FAO-GAEZ	0.00	-0.38	1.00	-0.62	7.23	1,161
Historical Population Density (std)	HYDE	0.00	-0.23	1.00	-0.30	25.85	1,179
Plow Advantage (std)	FAO-GAEZ	-0.00	0.31	1.00	-2.83	2.61	1,179
% Fertile land	Ramankutty et al (2002)	-0.00	-0.03	1.00	-1.43	2.53	1,134
Caloric Suitability Index (std)	Galor and Ozak (2015)	0.00	0.28	1.00	-1.95	2.63	1,179

FAO GAEZ v3 database downloaded on 15/01/2016. std - a standardized variable that has been rescaled to have a mean of zero and a standard deviation of one.

Table E.2: Descriptive Statistics: Countries X 50 years

	SOURCE	**Mean**	**p50**	**SDev**	**Min**	**Max**	**N**
Hierarchy	Borcan et al. (2014)	0.72	1.00	0.45	0.00	1.00	2,869
LandProd (std)	authors	0.00	0.35	1.00	-1.64	2.69	2,959
CerAdv (std)	authors	0.00	-0.00	1.00	-1.49	3.12	2,959
Precipitation (std)	FAO-GAEZ	0.00	-0.29	1.00	-1.38	2.89	2,940
Temperature (std)	FAO-GAEZ	0.00	0.20	1.00	-2.68	1.52	2,884
Elevation (std)	GDEM	0.00	-0.33	1.00	-1.10	4.65	2,845
Ruggedness (std)	GDEM	0.00	-0.31	1.00	-1.12	4.25	2,959
Absolute Latitude (std)	Nunn and Puga (2012)	0.00	-0.17	1.00	-1.51	2.18	2,959
Legal Origin: English common law	La Porta et al. (1999)	0.27	0.00	0.44	0.00	1.00	2,959
Legal Origin: French civil law	La Porta et al. (1999)	0.45	0.00	0.50	0.00	1.00	2,959
Legal Origin: Socialist law	La Porta et al. (1999)	0.22	0.00	0.41	0.00	1.00	2,959
Legal Origin: German civil law	La Porta et al. (1999)	0.03	0.00	0.18	0.00	1.00	2,959
Legal Origin: Scandinavian law	La Porta et al. (1999)	0.03	0.00	0.18	0.00	1.00	2,959
Population density 1500 (std)	Acemoglu et al. (2002)	0.00	-0.05	1.00	-2.96	2.78	2,959
Mortality of early settlers (std)	Acemoglu et al. (2002)	0.00	-0.11	1.00	-2.91	2.56	1,519
Slaves exported (std)	Nunn (2008)	0.00	-0.26	1.00	-0.26	9.01	2,959
Genetic Diversity (std)	Ashraf and Galor (2013)	0.00	0.24	1.00	-3.66	1.74	2,675
Distance to major river (std)	www.pdx.edu/econ/	0.00	-0.29	1.00	-0.89	7.63	2,845
Distance to coast (std)	www.pdx.edu/econ/	0.00	-0.41	1.00	-0.75	4.48	2,845
Pct Malaria	MAP	0.65	0.94	0.41	0.00	1.00	2,883
% country with tropical climate (std)	Nunn and Puga (2012)	0.35	0.00	0.43	0.00	1.00	2,959
Caloric Suitability Index (std)	Galor and Ozak (2015)	0.00	0.29	1.00	-1.82	2.93	2,959

FAO GAEZ v3 database downloaded on 15/01/2016. std - a standardized variable that has been rescaled to have a mean of zero and a standard deviation of one.

Table E.3: Descriptive Statistics: 1x1 decimal degree pixel

	SOURCE	Mean	p50	SDev	Min	Max	N
Panel A: *cross-sectional data-15,927 raster points*							
Cities founded before 400 AD	DeGroff (2009)	0.16	0.00	1.36	0.00	76.00	15,927
Archaeological sites	ANCIENTLOCATIONS.NET	0.24	0.00	2.58	0.00	138.00	15,927
Pyramids or Mastaba	MEGALITHIC.CO.UK	0.01	0.00	0.75	0.00	87.00	15,927
Temples	MEGALITHIC.CO.UK	0.04	0.00	0.64	0.00	46.00	15,927
Mines	MEGALITHIC.CO.UK	0.01	0.00	0.21	0.00	22.00	15,927
Palaces	MEGALITHIC.CO.UK	0.00	0.00	0.06	0.00	5.00	15,927
Sculptured Stones	MEGALITHIC.CO.UK	0.02	0.00	0.83	0.00	101.00	15,927
Standing Stones	MEGALITHIC.CO.UK	0.04	0.00	0.71	0.00	45.00	15,927
LandProd (std)	authors	-0.70	-1.28	1.26	-1.78	3.76	15,927
CerAdv (std)	authors	-0.93	-1.40	1.03	-2.67	2.96	15,927
DistanceCer	authors	28.71	24.72	22.78	0.00	270.26	15,927
DistanceAgr	authors	19.02	15.51	17.57	0.00	234.51	15,927
Precipitation (std)	FAO-GAEZ	0.00	-0.32	1.00	-1.06	9.24	15,862
Temperature (std)	FAO-GAEZ	0.00	-0.32	1.00	-1.33	1.81	15,833
Elevation (std)	GDEM	-0.00	-0.34	1.00	-0.87	6.01	15,927
Ruggedness (std)	GDEM	-0.00	0.35	1.00	-3.38	1.10	15,927
Absolute Latitude	authors	40.52	41.50	22.20	0.50	83.50	15,927
Irrigation Potential (std)	authors	0.00	0.67	1.00	0.67	1.74	8,214
Plow Advantage (std)	authors	0.00	0.05	1.00	-3.00	3.44	15,927
Population density 1995 (std)	FAO-GAEZ	0.00	-0.59	1.00	-0.76	3.58	15,861
Panel B: *panel data-15,927 raster points with observations either before or after Neolithic transition*							
Prehistoric archaelogical sites	D. and R. Whitehouse (1975)	0.11	0.00	0.71	0.00	23.00	31,854
Prehistoric settlements	D. and R. Whitehouse (1975)	0.08	0.00	0.59	0.00	21.00	31,854

FAO GAEZ v3 database downloaded on 15/01/2016. std - a standardized variable that has been rescaled to have a mean of zero and a standard deviation of one.

Table E.4: Caloric content of cereals, roots and tubers

Crop	Energy	Crop	Energy
Barley	3.52	Sorghum	3.39
Buckwheat	3.43	Sweet Potato	0.86
Cassava	1.6	Wetland Rice	3.7
Foxtail Millet	3.78	Wheat	3.47
Indigo Rice	3.7	White Potato	0.77
Maize	3.65	Yams	1.18
Oat	2.46	Sorghum	3.39
Rye	3.38		

Values are in kilo calories per 100g. Source: Galor and Ozak (2015) and USDA Nutrient Database for Standard Reference (R25). The data source in table A1 is different, and therefore the caloric content reported there is slightly different as well.

Table E.5: Pairwise correlations of the main variables used in the empirical analysis on the societies in the Ethnoatlas

Variables	*Hier.*	*CerMain*	Dep. agric.	Farm. surp.	H. Pop dens.	*LandProd*	*CerAdv*	% Fertile land	Caloric suit. ind.
Hierarchy	1.0								
CerMain	0.3	1.0							
Dependence agriculture	0.4	0.5	1.0						
Farming surplus	0.6	0.4	0.3	1.0					
Hist Pop density (Pryor)	0.6	0.5	0.7	0.4	1.0				
LandProd	0.2	0.3	0.4	0.2	0.3	1.0			
CerAdv	0.2	0.4	0.3	0.3	0.2	0.8	1.0		
% Fertile land	0.2	0.2	0.3	0.2	0.3	0.4	0.5	1.0	
Caloric suitability index	0.2	0.3	0.5	0.2	0.3	1.0	0.8	0.5	1.0

Table E.6: Potential Crop Yields and Choice of Crops. Robustness checks: Controlling for Geography.

	Dep. Variable: Major crop is cereal grains (dummy)			
	(1)	(2)	(3)	(4)
CerAdv	0.214***	0.274***	0.248***	0.250***
	(0.057)	(0.054)	(0.059)	(0.059)
LandProd	-0.088	-0.174***	-0.132*	-0.133*
	(0.067)	(0.064)	(0.069)	(0.070)
Precipitation	-0.058*			
	(0.031)			
Temperature		0.066**		
		(0.033)		
Elevation			0.030*	
			(0.017)	
Ruggedness				0.012
				(0.026)
CONTINENT FE	YES	YES	YES	YES
r2	0.367	0.364	0.362	0.359
N	982	982	982	982

The table reports cross-sectional OLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.7: Potential Crop Yields and Choice of Crops. Robustness checks: Controlling for Isolation, Population Density and the Plow.

	Dep. Variable: Major crop is cereal grains (dummy)						
	(1)	(2)	(5)	(6)		(7)	(8))
CerAdv	0.252***	0.250***	0.253***	0.261***	0.254***	0.253***	0.257***
	(0.058)	(0.060)	(0.059)	(0.071)	(0.059)	(0.058)	(0.049)
LandProd	-0.137**	-0.135*	-0.136**	-0.222***	-0.139**	-0.204***	-0.205***
	(0.067)	(0.069)	(0.069)	(0.086)	(0.069)	(0.060)	(0.061)
Major River	-0.028*						
	(0.017)						
Distance Coast		0.016					
		(0.035)					
Hist Pop Dens (HYDE)			-0.015				
			(0.016)				
Hist Pop Dens (PRYOR)				0.206***			
				(0.038)			
Pop Dens 1995					-0.004		
					(0.025)		
Irrigation						0.191***	
						(0.049)	
Plow Advantage							-0.148***
							(0.033)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.362	0.360	0.360	0.383	0.348	0.396	0.398
N	982	982	982	144	966	800	982

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.8: Cereals and Hierarchy - Reduced Form using generalized ordered logit

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community			
	Hierarchy<=1 vs Hierarchy>1	H<=2 vs H>2	H<=3 vs H>3	H<=4 vs H>4
CerAdv	0.327*	0.542***	0.674***	0.841**
	(0.173)	(0.172)	(0.230)	(0.407)
LandProd	0.0596	-0.392**	-0.485*	-0.597
	(0.198)	(0.199)	(0.281)	(0.515)

The table reports the estimates from a generalized ordered logit. The unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.9: Cereals and Hierarchy - 2SLS. Robustness checks: Controlling for Geography.

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community			
	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
CerMain	0.911	0.750*	1.102**	1.071**
	(0.624)	(0.407)	(0.553)	(0.545)
LandProd	-0.008	0.051	-0.045	-0.039
	(0.081)	(0.062)	(0.074)	(0.075)
Precipitation	-0.001			
	(0.001)			
Temperature		-0.248***		
		(0.072)		
Elevation			-0.069*	
			(0.039)	
Ruggedness				-0.008
				(0.050)
CONTINENT FE	YES	YES	YES	YES
N	952	952	952	952
F excl instrum.	49.13	83.83	74.16	74.51

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.10: Cereals and Hierarchy - 2SLS. Robustness checks: Controlling for Isolation, Population Density and the Plow.

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
CerMain	1.073**	1.078**	1.021**	1.471*	0.992**	0.933*	1.029**
	(0.519)	(0.545)	(0.504)	(0.811)	(0.472)	(0.555)	(0.453)
LandProd	-0.040	-0.038	-0.056	0.006	-0.085	0.012	0.080
	(0.070)	(0.071)	(0.071)	(0.119)	(0.072)	(0.078)	(0.066)
Major River	0.122***						
	(0.038)						
Distance to Coast		-0.024					
		(0.058)					
Hist Pop Dens (HYDE)			0.211**				
			(0.102)				
Hist Pop Dens (PRYOR)				0.276			
				(0.192)			
Pop Dens 1995					0.290***		
					(0.048)		
Irrigation						0.445*	
						(0.242)	
Plow Advantage							0.259***
							(0.093)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
N	952	952	952	142	936	770	952
F excl instrum.	76.84	74.70	77.41	14.22	76.15	84.42	85

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.11: Cereals and Hierarchy - 2SLS. Robustness checks: Sample Including Societies Living in Desertic Soils.

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.712***	1.200***	0.831**	0.999***	0.313***	0.839***	1.180***	1.092***
	(0.0596)	(0.206)	(0.360)	(0.262)	(0.0703)	(0.273)	(0.322)	(0.284)
LandProd			0.0667				-0.0489	
			(0.0520)				(0.0418)	
DEPENDENCE ON				0.327				-0.513
AGRICULTURE				(0.257)				(0.434)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	1059	1059	1059	1059	1059	1059	1059	1059
F excl instrum.		130.2	44.59	56.16		81.93	64.09	51.98
A-R Test (p-val)		0.000	0.0183	0.000		0.00163	0.000	0.000

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. All societies included in the Ethnoatlas, for which the relevant data are available, are included in the sample. "A-R Test" is the Anderson-Rubin test: the null hypothesis that the endogenous regressor is equal to zero. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.12: Cereals and Hierarchy - 2SLS. Robustness checks: Potential Calorie Yields Refer to Ethnic Boundaries in Fenske (2013)

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.707***	1.104***	0.752	0.872**	0.304**	0.839**	0.897**	0.898**
	(0.131)	(0.364)	(0.483)	(0.414)	(0.120)	(0.395)	(0.436)	(0.440)
LandProd			0.104				-0.015	
			(0.099)				(0.060)	
DEPENDENCE ON				0.569				-0.225
AGRICULTURE				(0.520)				(0.892)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	952	942	942	952	952	942	942	942
F excl instrum.		156.3	55.98	52.60		120.1	88.82	20.90

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.13: Cereals and Hierarchy - 2SLS. Robustness checks: Controlling for Alternative Measures of Land Suitability for Agriculture

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community			
	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
CerMain	1.009***	0.723	0.867	1.121*
	(0.372)	(0.478)	(0.636)	(0.585)
% fertile land	0.073	0.057		
(Ramankutty et al. 2002)	(0.061)	(0.054)		
Caloric Suitability Index			0.081	-0.049
(Galor and Ozak, 2016)			(0.138)	(0.078)
CONTINENT FE	NO	YES	NO	YES
N	952	952	952	952
F excl instrum.	106.3	70.49	38.25	65.04

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. All societies included in the Ethnoatlas, for which the relevant data are available, are included in the sample. Standard errors (in parentheses) are adjusted for spatial correlation using Conley's (1999) method. *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.14: Cereals and Surplus - OLS and 2SLS

	Dependent variable: Existence of a farming surplus							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.359***	0.940***	0.846***	0.846***	0.299***	1.005***	0.797**	0.799**
	(0.0791)	(0.260)	(0.273)	(0.275)	(0.0901)	(0.316)	(0.314)	(0.317)
LandProd			0.0186				0.0361	
			(0.0626)				(0.0611)	
DEPENDENCE ON				0.191				0.438
AGRICULTURE				(0.663)				(0.775)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	139	139	139	139	139	139	139	139
F excl instrum.		16.08	17.37	5.486		15.35	12.44	4.338

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.15: Cereals and Surplus - 2SLS. Robustness checks: Controlling for Geography.

	Dependent variable: Existence of a farming surplus			
	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
CerMain	0.686*	0.718**	0.855***	0.834**
	(0.385)	(0.284)	(0.329)	(0.327)
LandProd	0.0567	0.0525	0.0211	0.00806
	(0.0722)	(0.0663)	(0.0639)	(0.0717)
Precipitation	-0.0546			
	(0.0727)			
Temperature		-0.0326		
		(0.0607)		
Elevation			-0.0934***	
			(0.0340)	
Ruggedness				-0.100
				(0.0637)
CONTINENT FE	YES	YES	YES	YES
N	139	139	139	139
F excl instrum.	9.260	17.77	12.12	12.20

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.16: Cereals and Surplus - 2SLS. Robustness checks: Controlling for Isolation, Population Density and the Plow.

	Dependent variable: Existence of a farming surplus						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
CerMain	0.736**	0.802**	0.732**	0.798**	0.825**	0.637*	0.786***
	(0.318)	(0.319)	(0.324)	(0.312)	(0.327)	(0.349)	(0.283)
LandProd	0.0449	0.0358	0.0254	0.0333	0.0214	0.111**	0.0395
	(0.0594)	(0.0615)	(0.0584)	(0.0518)	(0.0603)	(0.0509)	(0.0570)
Major River	0.0560						
	(0.0418)						
Distance to Coast		-0.00556					
		(0.0472)					
Hist Pop Dens (HYDE)			0.0689*				
			(0.0375)				
Hist Pop Dens (PRYOR)				0.0115			
				(0.0861)			
Pop Density 1995					0.0287		
					(0.0360)		
Irrigation						0.0638	
						(0.134)	
Plow Advantage							0.00735
							(0.0526)
N	139	139	139	139	137	111	139
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
F excl instrum.	11.05	11.87	10.51	14.08	10.75	12.15	16.75

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.17: Cereals and Surplus - OLS and 2SLS. Robustness checks: Potential Calorie Yields Refer to Ethnic Boundaries in Fenske (2013).

	Dependent variable: Existence of a farming surplus							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.359***	0.909***	0.894***	0.846***	0.299***	0.953***	0.845**	0.864***
	(0.0791)	(0.274)	(0.297)	(0.275)	(0.0901)	(0.318)	(0.336)	(0.303)
LandProd			0.00286				0.0196	
			(0.0657)				(0.0657)	
DEPENDENCE ON AGRICULTURE				0.191				0.210
				(0.663)				(0.723)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	139	138	138	138	139	138	138	138
F excl instrum.		15.52	17.23	5.486		16.90	13.56	4.786

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.18: Cereals and Surplus - OLS and 2SLS. Robustness checks: Sample Including Societies Living in Desertic Soils.

	Dependent variable: Existence of a farming surplus							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
CerMain	0.368***	0.630***	0.871***	0.871***	0.294***	0.657**	0.814***	0.821***
	(0.0733)	(0.220)	(0.279)	(0.283)	(0.0849)	(0.260)	(0.300)	(0.316)
LandProd			-0.0368				-0.0215	
			(0.0501)				(0.0473)	
DEPENDENCE ON AGRICULTURE				-0.362				-0.244
				(0.488)				(0.540)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	161	161	161	161	161	161	161	161
F excl instrum.		18.58	17.37	14.46		19.68	14.27	7.531
A-R Test (p-val)		0.00711	0.000	0.000		0.0109	0.00391	0.00191

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. All societies included in the Ethnoatlas, for which the relevant data are available, are included in the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.19: Cereals, Surplus and Hierarchy - 2SLS. Robustness Checks: Different Measures of Soil Suitability for Agriculture.

	Dependent variable: Existence of a farming surplus			
	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
CerMain	1.168***	1.270***	0.878***	0.843**
	(0.368)	(0.419)	(0.303)	(0.354)
% fertile land	-0.0819	-0.0844		
(Ramankutty et al. 2002)	(0.0574)	(0.0590)		
Caloric Suitability Index			0.0124	0.0285
(Galor and Ozak, 2016)			(0.0671)	(0.0652)
CONTINENT FE	NO	YES	NO	YES
N	139	139	139	139
F excl instrum.	8.528	10.39	14.30	10.30

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. All societies included in the Ethnoatlas, for which the relevant data are available, are included in the sample. Robust standard errors in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.20: Cereals and Hierarchy - Panel Regressions - Robustness Checks

	Dep. Variable: Hierarchy Index								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CerAdv	0.160*	0.127	0.206*	0.274***	0.162*	0.245***	0.258***	0.273***	0.254***
	(0.0892)	(0.0843)	(0.116)	(0.0833)	(0.0831)	(0.0928)	(0.0957)	(0.0840)	(0.0675)
LandProd	-0.0507	0.0471	-0.261	-0.176	-0.00456	-0.121	-0.133	-0.199	-0.211**
	(0.133)	(0.132)	(0.192)	(0.143)	(0.148)	(0.151)	(0.151)	(0.145)	(0.102)
Controls (xYear):									
Legal Origin	YES	NO	NO	NO	NO	NO	NO	NO	NO
Pop Density 1500	NO	YES	NO	NO	NO	NO	NO	NO	NO
Settlers Mortality	NO	NO	YES	NO	NO	NO	NO	NO	NO
Slave Exports	NO	NO	NO	YES	NO	NO	NO	NO	NO
Genetic Diversity	NO	NO	NO	NO	YES	NO	NO	NO	NO
Distance River	NO	NO	NO	NO	NO	YES	NO	NO	NO
Distance Coast	NO	NO	NO	NO	NO	NO	YES	NO	NO
Pct Malaria	NO	NO	NO	NO	NO	NO	NO	YES	NO
Tropical Land	NO	NO	NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
TIME FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
r2	0.699	0.714	0.707	0.683	0.682	0.678	0.679	0.681	0.744
N	2869	2869	1501	2869	2603	2755	2755	2793	2869

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.21: Cereals and Hierarchy - Panel Regressions. Robustness checks: Excluding Colonies

	Dep. Variable: Hierarchy Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CerAdv	0.128*	0.186**	0.230***	0.162**	0.182**	0.178**	0.135*
	(0.0660)	(0.0786)	(0.0735)	(0.0816)	(0.0857)	(0.0788)	(0.0772)
LandProd		-0.111	-0.179	-0.0997	-0.0884	-0.0879	-0.115
		(0.136)	(0.131)	(0.136)	(0.138)	(0.134)	(0.119)
Controls (x Year FE):							
Precipitation	NO	NO	YES	NO	NO	NO	NO
Temperature	NO	NO	NO	YES	NO	NO	NO
Elevation	NO	NO	NO	NO	YES	NO	NO
Ruggedness	NO	NO	NO	NO	NO	YES	NO
Abs Latitude	NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES
YEAR FE	YES	YES	YES	YES	YES	YES	YES
r2	0.773	0.774	0.789	0.774	0.770	0.777	0.786
N	2414	2414	2398	2365	2329	2414	2414

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. The sample excludes those cells 50yearsXcountry in which countries were either colonies or protectorates. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.22: Cereals and Hierarchy - Panel Regressions. Robustness checks: a Different Measure of Hierarchy

	Dep. Variable: Government above tribal level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CerAdv	0.188***	0.270***	0.280***	0.235***	0.252***	0.259***	0.192**
	(0.0683)	(0.0835)	(0.0758)	(0.0855)	(0.0890)	(0.0840)	(0.0791)
LandProd		-0.159	-0.189	-0.150	-0.110	-0.145	-0.161
		(0.140)	(0.131)	(0.138)	(0.142)	(0.138)	(0.122)
Controls (x Year FE):							
Precipitation	NO	NO	YES	NO	NO	NO	NO
Temperature	NO	NO	NO	YES	NO	NO	NO
Elevation	NO	NO	NO	NO	YES	NO	NO
Ruggedness	NO	NO	NO	NO	NO	YES	NO
Abs Latitude	NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES
YEAR FE	YES	YES	YES	YES	YES	YES	YES
r2	0.672	0.674	0.707	0.677	0.673	0.677	0.699
N	2869	2869	2850	2812	2755	2869	2869

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. The dependent variable is a dummy that identifies those territories characterized by a supra-tribal government. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.23: Cereals and Hierarchy - Panel Regressions. Robustness checks: a Different Measure of Soil Suitability for Agriculture

	Dep. Variable: Hierarchy Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CerAdv	0.189***	0.272***	0.282***	0.240***	0.255***	0.261***	0.197**
	(0.0683)	(0.0834)	(0.0760)	(0.0857)	(0.0889)	(0.0839)	(0.0795)
Caloric Suitability Index		-0.163	-0.193	-0.152	-0.115	-0.148	-0.165
		(0.141)	(0.131)	(0.139)	(0.142)	(0.138)	(0.123)
Controls (x Year FE):							
Precipitation	NO	NO	YES	NO	NO	NO	NO
Temperature	NO	NO	NO	YES	NO	NO	NO
Elevation	NO	NO	NO	NO	YES	NO	NO
Ruggedness	NO	NO	NO	NO	NO	YES	NO
Abs Latitude	NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES
TIME FE	YES	YES	YES	YES	YES	YES	YES
r2	0.680	0.682	0.716	0.684	0.681	0.686	0.705
N	2869	2869	2850	2812	2755	2869	2869

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.24: Cereals and Hierarchy - Panel Regressions. Robustness Checks: Excluding Years 1500-1750

	Dep. Variable: Hierarchy Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CerAdv	0.198***	0.272***	0.282***	0.235***	0.249***	0.260***	0.190**
	(0.0720)	(0.0889)	(0.0811)	(0.0912)	(0.0946)	(0.0892)	(0.0846)
LandProd		-0.145	-0.176	-0.140	-0.0889	-0.130	-0.148
		(0.149)	(0.140)	(0.146)	(0.150)	(0.146)	(0.129)
Controls (x Year FE):							
Precipitation	NO	NO	YES	NO	NO	NO	NO
Temperature	NO	NO	NO	YES	NO	NO	NO
Elevation	NO	NO	NO	NO	YES	NO	NO
Ruggedness	NO	NO	NO	NO	NO	YES	NO
Abs Latitude	NO	NO	NO	NO	NO	NO	YES
COUNTRY FE	YES	YES	YES	YES	YES	YES	YES
YEAR FE	YES	YES	YES	YES	YES	YES	YES
r2	0.711	0.712	0.743	0.715	0.711	0.716	0.735
N	2416	2416	2400	2368	2320	2416	2416

The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. The years 1500-1750 are excluded from the regression. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.25: Potential Crop Yields and the Location of Ancient Cities. Robustness checks: Controlling for Geography and Population Density.

	Dependent variable: Presence of an ancient city (dummy)			
	(1)	(2)	(3)	(4)
CerAdv	0.124***	0.167***	0.129***	0.129***
	(0.0367)	(0.0458)	(0.0380)	(0.0386)
LandProd	-0.0693***	-0.120***	-0.0751***	-0.0743***
	(0.0253)	(0.0333)	(0.0257)	(0.0261)
Precipitation	-0.00393			
	(0.00735)			
Temperature		0.0588***		
		(0.0163)		
Elevation			-0.00178	
			(0.00624)	
Ruggedness				-0.00883
				(0.00747)
CONTINENT FE	YES	YES	YES	YES
r2	0.0987	0.116	0.0986	0.100
N	15862	15833	15927	15927

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.26: Potential Crop Yields and the Location of Ancient Cities. Robustness checks: Controlling for Geography and Population Density.

	Dependent variable: Presence of an ancient city (dummy)					
	(1)	(2)	(3)	(4)	(5)	(6)
CerAdv	0.168***	0.109***	0.104***	0.0946***	0.0634***	0.112***
	(0.0474)	(0.0400)	(0.0278)	(0.0301)	(0.00546)	(0.0410)
LandProd	-0.121***	-0.0824***	-0.0522***	-0.0910***	-0.0405***	-0.0694**
	(0.0358)	(0.0313)	(0.0179)	(0.0232)	(0.00440)	(0.0292)
Abs Latitude	-0.00257***					
	(0.000908)					
Irrigation Potential		0.000316				
		(0.0122)				
Plow Advantage			0.0315***			
			(0.00962)			
Pop Dens 1995				0.0891***		
				(0.0128)		
CONTINENT FE	YES	YES	YES	YES	YES	YES
r2	0.112	0.113	0.113	0.172	0.0705	0.105
N	15927	8214	15927	15861	12052	8942

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. In column 5 the sample excludes Europe, in column 6 the sample excludes deserts. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.27: The Origin of the Neolithic Transition and the Location of Ancient Cities. Robustness checks: Controlling for Geography and Population Density.

	Dependent variable is the presence of an ancient city (dummy)			
	(1)	(2)	(3)	(4)
DistanceCer	-0.00141**	-0.000710	-0.00148**	-0.00132**
	(0.000601)	(0.000922)	(0.000679)	(0.000630)
DistanceAgr	0.000250	0.000340	0.000323	0.000101
	(0.000576)	(0.000653)	(0.000609)	(0.000617)
Precipitation	0.00153			
	(0.00715)			
Temperature		0.0412**		
		(0.0186)		
Elevation			-0.00719	
			(0.00863)	
Ruggedness				-0.00600
				(0.00666)
Pop Dens 1995				
Abs Latitude				
CONTINENT FE	YES	YES	YES	YES
r2	0.0497	0.0593	0.0505	0.0501
N	15862	15833	15927	15927

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.28: The Origin of the Neolithic Transition and the Location of Ancient Cities. Robustness checks: Controlling for Geography and Population Density.

	Dependent variable is the presence of an ancient city (dummy)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DistanceCer	-0.000710	-0.00354**	-0.00294***	-0.00136*	-0.00168**	-0.00144**	-0.00264***
	(0.00112)	(0.00141)	(0.00103)	(0.000710)	(0.000651)	(0.000591)	(0.000922)
DistanceAgr	0.000339	-0.000386	0.000528	0.000271	0.00194***	0.000564	0.000369
	(0.000662)	(0.00170)	(0.00125)	(0.000664)	(0.000636)	(0.000468)	(0.00103)
Absolute Latitude	-0.00159	0.00148					
	(0.00110)	(0.00139)					
Irrigation Potential			0.00536				
			(0.0105)				
Plow Advantage				0.0352***			
				(0.00984)			
Pop Density 1995					0.0819***		
					(0.0130)		
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.0548	0.0460	0.105	0.0689	0.153	0.0642	0.0836
N	15927	5763	8214	15927	15861	12052	8942

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. In column 6 the sample includes only Asia and Africa, in column 7 the sample excludes deserts. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.29: The Origin of the Neolithic Transition and Archeological Ruins.

	Dependent variable is a dummy that identifies evidence of:						
	ancient archaeolog. sites	pyramids	ancient temples	ancient mines	ancient palaces	ancient sculptured stones	ancient standing stones
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DistanceCer	-0.00279***	-0.000282	-0.000636**	-0.000210**	-0.000132**	-0.000232**	-0.0000152
	(0.000824)	(0.000187)	(0.000311)	(0.000106)	(0.0000550)	(0.000108)	(0.0000706)
DistanceAgr	0.000864	0.000105	0.000316	0.0000109	0.0000689	0.000166	0.00000243
	(0.000753)	(0.000146)	(0.000330)	(0.000144)	(0.0000487)	(0.000105)	(0.000119)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.0328	0.00451	0.0105	0.00294	0.00189	0.00930	0.0187
N	15927	15927	15927	15927	15927	15927	15927

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. The dependent variable is a dummy that takes the value of one if there is archaeological evidence of either ancient sites from the Stone Age (column 1), or ancient pyramids or mastaba (column 2), or ancient temples (column 3), or ancient mines or quarries (column 4), or ancient palaces (column 5), or ancient sculptured stones (column 6). or ancient standing stones (column 7). Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.30: The Origin of the Neolithic Transition and Archaeological Ruins.

	Dependent variable is the log (1+ number of archaeological ruin)						
	ancient archaeolog. sites	pyramids	ancient temples	ancient mines	ancient palaces	ancient sculputured stones	ancient standing stones
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DistanceCer	-0.00359***	-0.000264	-0.000556*	-0.000153**	-0.000109**	-0.000193**	-0.0000245
	(0.00114)	(0.000170)	(0.000306)	(0.0000740)	(0.0000446)	(0.0000932)	(0.0000652)
DistanceAgr	0.00120	0.0000880	0.000167	0.0000140	0.0000581	0.000152	0.0000752
	(0.000981)	(0.000133)	(0.000332)	(0.000101)	(0.0000382)	(0.0000977)	(0.000157)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.0266	0.00256	0.00784	0.00260	0.00159	0.00800	0.0176
N	15927	15927	15927	15927	15927	15927	15927

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. The dependent variable is log of one plus the number of either ancient sites from the Stone Age (column 1), or ancient pyramids or mastaba (column 2), or ancient temples (column 3), or ancient mines or quarries (column 4), or ancient palaces (column 5), or ancient sculptured stones (column 6), or ancient standing stones (column 7). Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.31: The Origin of the Neolithic Transition and Archaeological Ruins. Robustness Checks: Excluding Europe

	Dependent variable is a dummy that identifies evidence of						
	ancient archaeolog. sites	pyramids	ancient temples	ancient mines	ancient palaces	ancient sculputured stones	ancient standing stones
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DistanceCer	-0.00279***	-0.000281	-0.000637**	-0.000208*	-0.000132**	-0.000231**	-0.0000128
	(0.000826)	(0.000181)	(0.000311)	(0.000122)	(0.0000558)	(0.000104)	(0.0000637)
DistanceAgr	0.000776	0.0000616	0.000362	-0.0000761	0.0000597	0.000124	-0.0000923
	(0.000528)	(0.000109)	(0.000328)	(0.000141)	(0.0000476)	(0.0000768)	(0.0000663)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.0410	0.00524	0.0135	0.00405	0.00265	0.00520	0.00416
N	12052	12052	12052	12052	12052	12052	12052

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. The sample excludes Europe. The dependent variable is a dummy that takes the value of one if there is archeological evidence of either ancient sites from the Stone Age (column 1), or ancient pyramids or mastaba (column 2), or ancient temples (column 3), or ancient mines or quarries (column 4), or ancient palaces (column 5), or ancient sculptured stones (column 6), or ancient standing stones (column 7). Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Table E.32: Potential Crop Yields and Archaeological Ruins.

	Dependent variable is a dummy that identifies evidence of						
	ancient archaeolog. sites	pyramids	ancient temples	ancient mines	ancient palaces	ancient sculputured stones	ancient standing stones
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CerAdv	0.116***	0.00391	0.0193	0.0149**	0.00637**	0.0213**	0.0272**
	(0.0415)	(0.00331)	(0.0132)	(0.00637)	(0.00254)	(0.00829)	(0.0122)
LandProd	-0.0644**	-0.00187	-0.00927	-0.00809*	-0.00469**	-0.0108**	-0.0130*
	(0.0324)	(0.00340)	(0.0111)	(0.00486)	(0.00207)	(0.00521)	(0.00680)
CONTINENT FE	YES	YES	YES	YES	YES	YES	YES
r2	0.0578	0.00233	0.0143	0.00787	0.00312	0.0208	0.0328
N	15927	15927	15927	15927	15927	15927	15927

The table reports cross-sectional OLS estimates and the unit of observation is the 1x1 decimal degree square. The dependent variable is a dummy that takes the value of one if there is archeological evidence of either ancient sites from the Stone Age (column 1), or ancient pyramids or mastaba (column 2), or ancient temples (column 3), or ancient mines or quarries (column 4), or ancient palaces (column 5), or ancient sculptured stones (column 6), or ancient standing stones (column 7). Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Figure E.1: Farming surplus in pre-colonial societies

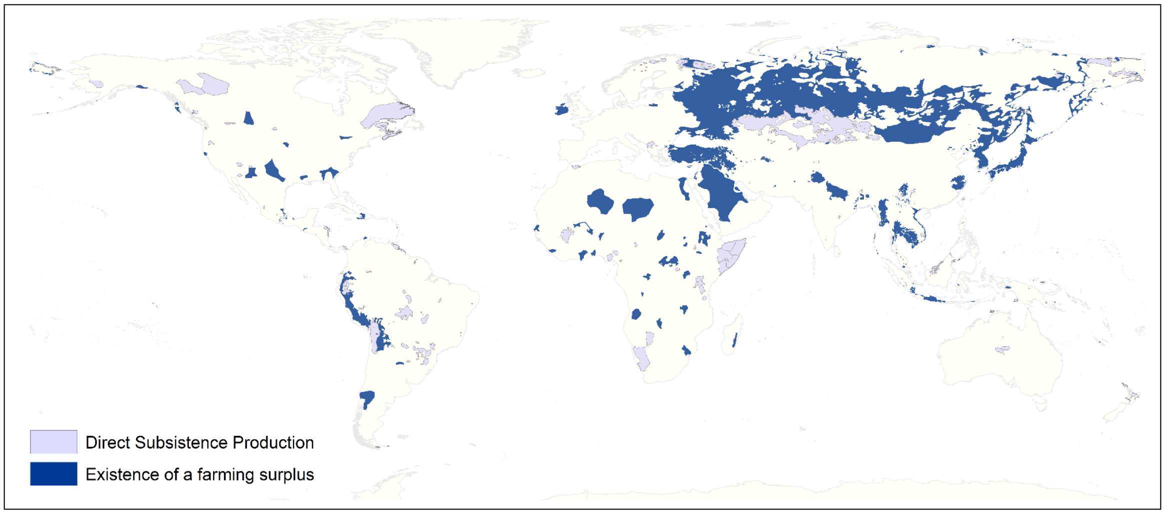


Figure E.2: Potential yields (calories per hectare) from cereal grains.

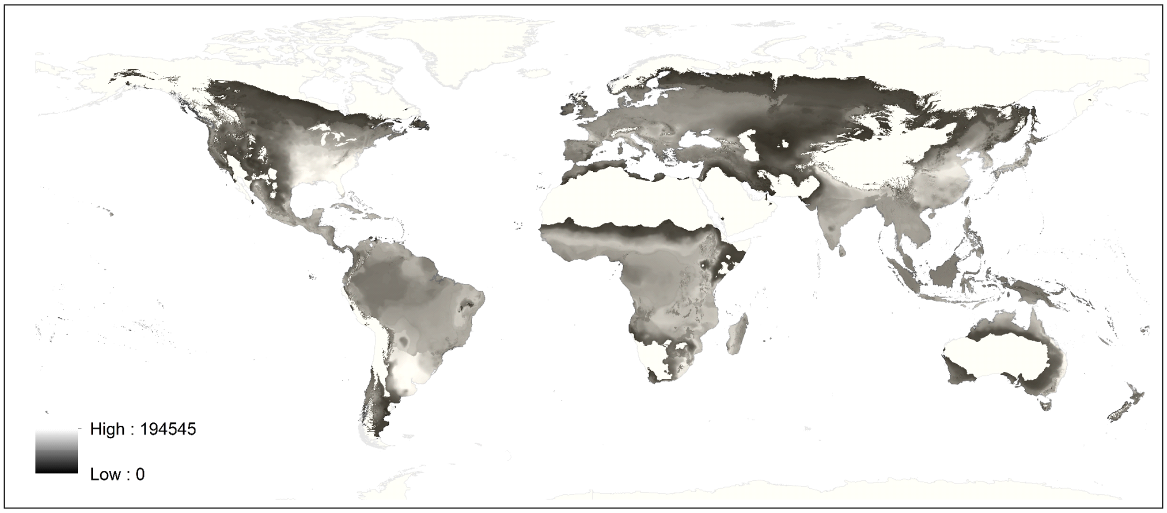


Figure E.3: Potential yields (calories per hectare) from roots and tubers

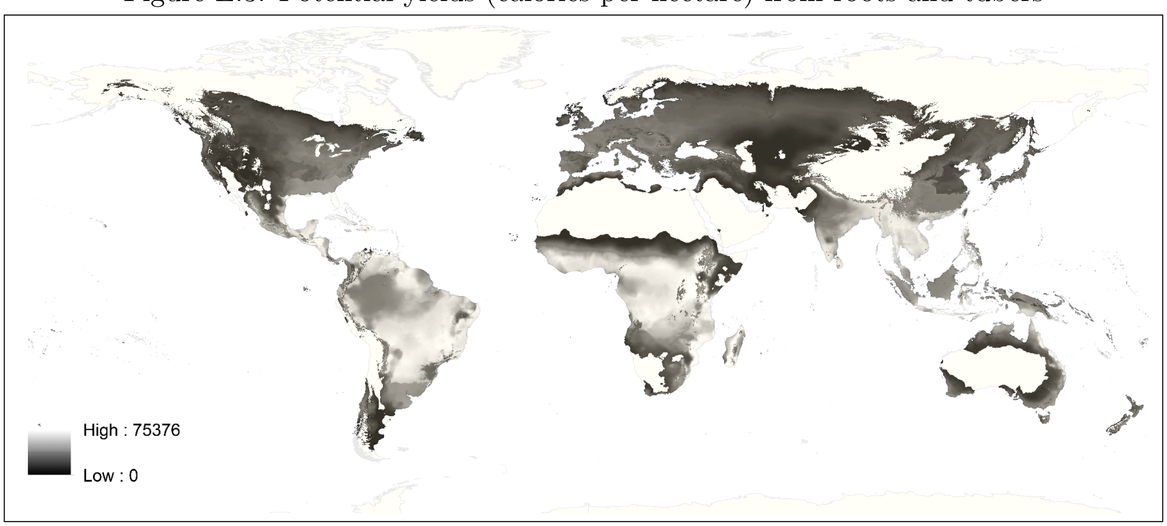


Figure E.4: Optimal crop in terms of caloric yields among cereals, roots and tubers

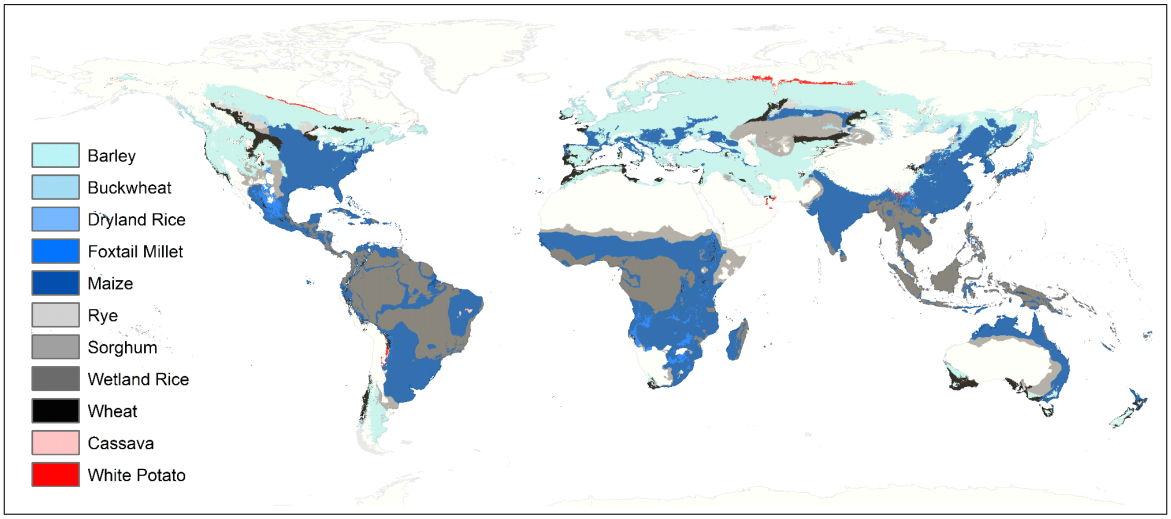


Figure E.5: Ancient cities and centers of independent domestication