Cereals, Appropriability and Hierarchy*

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

March 21, 2016

Abstract

We propose that the development of social hierarchy following the Neolithic Revolution was due to the ability of the emergent elite to appropriate crops from farmers and not a result of increased productivity, as conventional theory has it. We argue that since cereals are easier to appropriate than roots and tubers, regional variations in the suitability of land for the cultivation of different crops can account for differences in the formation of hierarchy and states. In support of this theory, our empirical investigation demonstrates that the productivity advantage of cereals over roots and tubers explains the formation and viability of hierarchical institutions, whereas absolute land productivity does not.

KEYWORDS: *Geography, Hierarchy, Institutions, State Capacity*

JEL CLASSIFICATION NUMBERS: *D02, D82, H10, O43*

*We would like to thank Daron Acemoglu, Alberto Alesina, 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, and John Wallis for useful comments. We are also grateful for comments from participants in seminars: Berlin SFB-seminar ,Birmingham, Boston University, Columbia, Harvard, Hebrew University, IDC, Manchester, Pompeu Fabra University, Tilburg, University College of London/London School of Economics, and Yale; and conferences: INET mini-conference, Cambridge, June 2014, CEPR Public Economics Annual Symposium, Warwick, December 2014, Deep-rooted Factors in Comparative Economic Development, Brown University, May 2015, Summer Workshop in Economic Growth, Warwick, July 2015, BREAD/NBER conference, Cambridge MA, September 2015, XII annual conference of the Colombian Central Bank's journal, Bogota, October 2015, CEPR-Development Workshop, Barcelona, October 2015, and Institutions Culture and Long-run Development, LMU Munich, November 2015.

†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, University of Warwick and CAGE. Email: luigi.pascali@upf.edu

1 Introduction

The transition to agriculture (known as the Neolithic Revolution) led to profound social changes. Hierarchies and city-states emerged, leading ultimately to the development of the great civilizations of antiquity. In Egypt, for example, state hierarchy grew rapidly following the adoption of farming in the Nile Valley and enabled the construction of the great pyramids in the third millennium BCE. Other regions of the world, however, followed a very different path: no state institutions emerged in New Guinea, even though agriculture was adopted there at about the same time as in Egypt.[1]

We offer an explanation for this disparity. The key factor for the emergence of hierarchy following the Neolithic Revolution, we posit, was the ability of the elite to appropriate food crops from farmers. Thus, regional variations in the suitability of land for roots and tubers, which are less appropriable, or cereals, which are more appropriable, can account for differences in the formation of hierarchy and related state institutions. We argue further that the conventional theory that higher productivity and food surplus were the key causal factor for the emergence of hierarchy is flawed. Our empirical investigation supports both claims.

Ever since Montesquieu (1748, book 14) asserted that the tropics were backward because people in hot climates tend to be timid and lazy, there have been various attempts to develop an environmental theory that explains the differences between countries such as ancient Egypt and New Guinea.[2] Nowadays, two main features of the tropics are typically argued to have impeded its development: low agricultural productivity and high burden of disease.[3] According to Diamond (1997), the east-west orientation of Eurasia resulted in greater plant variety and thereby in higher productivity of cultivable crops, which created food surplus and led to more developed social institutions. In other parts of the world such as in New Guinea, he argues, low productivity prevented the generation of surplus and thus retarded the emergence of state institutions and impeded economic development.[4]

[1] According to Denham (2011), systematic cultivation of bananas, taro and yam in New Guinea occurred ca. 5000-4500 BCE.

[2] Spolaore and Wacziarg (2013) provide a detailed survey and references of the empirical literature that links income per capita across countries with geographic variables.

[3] Sachs Mellinger and Gallup, (2001) and Olsson and Hibbs (2005) provide further discussion and additional mechanisms.

[4] Olsson and Hibbs (2005) support Diamond's theory by showing a strong effect of geography and plant variety on economic development.

Acemoglu, Johnson and Robinson (2001) deemphasize the role of geography. They argue that it is not the tropical climate per se, nor the endemic disease inflicted on indigenous people, which are responsible for the underdevelopment of countries closer to the equator, but rather the institutions that colonizers established there.[5] Acemoglu and Robinson (2012) argue further that extractive institutions are the key detrimental factor for economic prosperity, that institutions are by and large determined by the vagaries of human history, and that geography has limited, if any, direct effect on economic growth.

In this paper, we seek to identify the drivers for the rise of hierarchy following the transition to cereal farming, and to contribute to the debate on the role of geography vs. institutions in explaining current economic disparities. In our proposed framework, hierarchical institutions are endogenous, determined principally by geography. Our contribution is in identifying the elite's ability to appropriate food-produce as a causal mechanism that connects the environment and social hierarchy. Since modern economies were based on agriculture until the onset of the Industrial Revolution, our proposed theory pertains not only to antiquity, but also to the institutions of the more recent, pre-industrial world. Moreover, since the transition from agricultural-based economies is protracted, and since social institutions exhibit significant inertia, our theory may contribute towards explaining persistent differences in economic development, and particularly towards understanding the root-causes for the underdevelopment of tropical regions.[6]

The conventional explanation for the emergence of hierarchy maintains that the higher productivity of agriculture relative to foraging created food surplus, and this surplus facilitated, through various channels, the emergence of a non-food producing elite. We argue, in contrast, that surplus was neither necessary nor a sufficient precondition for the emergence of hierarchy. To understand why, consider a community of farmers with annual output above subsistence, who cultivate cassava (also known as manioc or yuca), a perennial root that can be harvested year-round but is highly perishable upon harvest. Since the crop isn't stored and rots shortly after harvest, it is difficult to

[5] Easterly and Levine (2003) and Rodrik, Subramanian and Trebbi (2004) demonstrate empirically that the link between the tropics and underdevelopment is indirect, due to the growth-retarding social institutions in tropical countries. 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.

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

appropriate. The available surplus in this case might lead to a population increase, but would not facilitate the emergence of hierarchy. Consider now another farming community growing a cereal grain, such as wheat, rice or maize, with no surplus, where each family's annual produce equals its subsistence needs. Since the grain has to be harvested within a short period and then stored, a visiting tax collector could readily confiscate part of the stored produce. Such ongoing confiscation may be expected to reduce population size; but this scenario demonstrates that surplus isn't a necessary precondition for taxation by the elite.[7]

In our alternative to the productivity-and-surplus explanation, we posit that the crucial element for understanding why the transition to agriculture only enabled the emergence of a hierarchy in some regions of the world and not in others, is that the crops cultivated in the former regions were vulnerable to appropriation. In particular, the emergence of fiscal capacity and hierarchy was conditioned on the cultivation of appropriable cereals as the staple crops, in contrast to less appropriable starchy crops such as roots and tubers. According to this theory, the non-emergence of complex hierarchy among hunter-gatherers was not because they live at subsistence, but because hunter-gatherers essentially (though not always, see section 5 below) live hand-to-mouth, with little that can be expropriated to feed a would-be elite. A similar explanation applies also to regions where farming was adopted, but where the land was adequate for the cultivation of non-appropriable crops.

Thus, rather than surplus facilitating the emergence of the elite, we contend that the elite emerged only once the opportunity to expropriate arose, implying that, in effect, the emerging elite created the food surplus on which it flourished. Whereas the conventional theory suggests that surplus led to hierarchy, we claim a reverse direction of causation. Moreover, while according to conventional wisdom it is land productivity (regardless of crop type) that accounts for the existence of hierarchy, in our appropriability theory it is environmental conditions that provide a sufficient

[7]Without any surplus, the population size is presumably such that the average product of labor equals subsistence. By taxing away some of the food, the government will cause a population decline. Given limited supply of farm land, average and marginal product of labor will be diminishing. The decline in population will thus result in higher average output, until the total output net of taxation would enable the population to survive. This scenario presumes however that the farmers' stored grains are somehow protected. In reality, as in our model below, stored grains are vulnerable to predation not only by tax collectors but also by bandits. If the government employs some of the tax revenue to protect farmers from confiscation by bandits, it may in fact facilitate a larger population. This is likely as long as the elite (unlike the bandits) is sufficiently forward looking and chooses the tax rate to maximize its long run revenue, while realizing that population size and total output decline as a result of higher taxation. For a formal presentation of this argument see Appendix B.

productivity advantage to appropriable cereals over less appropriable roots and tubers. It is this difference between the two theories that we take to the data.

In a sense we suggest here a new version of Hobbes's theory that states emerged in order to protect individuals. Warfare among clans of hunter-gatherers over desirable land parcels, or for other reasons, was apparently pervasive. But, contrary to Hobbes's theory, it did not lead to inheritable social hierarchy among simple hunter-gatherers. Nor did the transition to farming of non-seasonal crops give rise to complex (multi-level) social hierarchies beyond chiefdoms. Such hierarchies emerged, we argue, only alongside the transition to reliance on appropriable food sources that left farmers vulnerable to expropriation. In adapting Hobbes, the greater incentives for thievery that were created by the cultivation of cereals induced farmers to seek protection for their food stockpiles.[8] Due to increasing returns to scale in the provision of protection, these early farmers had to aggregate and to cooperate in order to defend their stored grains.

It does not matter for our case whether this demand for protection was met by outside roving bandits who turned stationary, as Olson (1993) posited, or by leaders from within, as is more conventionally assumed. Food storage and the demand for protection led to population agglomeration in villages and to the creation of a non-food producing elite that oversaw the provision of protection. Once a group became larger than a few dozens of immediate kin, it is unlikely that those who sought security were as forthcoming in financing the protection services they desired. This public-good nature of protection was overcome by the ability that storage provided to the elite to appropriate the necessary means. The ability to tax contributed also to specialization by those charged with protection, and to hereditary leadership. In other words, it was the transformation of technology for appropriation, due to the transition to cereals, which created both the demand for protection and the means for providing it. This is how we propose to explain the emergence of complex and hereditary social hierarchy, and eventually the state.

Our focus is on the emergence of hierarchy, but our proposed distinction between appropriable and non-appropriable crops provides some insights also with regard to the adoption of farming. It implies that, controlling for productivity, the transition to farming is more likely in areas suitable for roots and tubers that provide in-built protection, rather than in areas suitable for cereals that

[8] In addition to the necessity of storing cereals which makes farmers vulnerable to appropriation, cereal producers are vulnerable also to extortion. The long gestation of cereals in the field exposes farmers to the threat of arson, and to the possibility that violence could be used to deny access to land or water.

are less desirable due to their susceptibility to expropriation.

To indicate the applicability of our theory, consider again the case of New Guinea. We suggest that the basic cause for its low level of social complexity and its economic underdevelopment is that farming there relied mostly on tubers (yam and taro, and more recently sweet potatoes), where long-term storage is neither feasible (due to perishability) nor necessary (because harvesting was essentially non-seasonal).[9] These crop types provide farmers with sufficient immunity against bandits, as well as against potential tax collectors. Farmers' ability to cultivate highly productive, non-appropriable crops inhibits both the demand for socially-provided protection and the emergence of protection-providing elite. It is a curse of plenty.

In our formal model presented in section 3, a fixed number of farmers can allocate their land between only two crops, which we label cereals and tubers.[10] The productivity of the two crops is presumed to differ across geographic locations. We assume that it is possible to tax cereals, though at some costs. For simplicity, we assume that tubers, whose post-harvest storage is to a major extent unnecessary and unfeasible, are entirely immune from confiscation and taxation. As a result, cereals will only be cultivated if their productivity advantage over tubers is sufficiently high to compensate for taxation by the state or the risk of loss to bandits.

In a regime identified as 'anarchy,' we assume that non-farmers can engage in either banditry or foraging, where the income from foraging is exogenous and constant. Since we also assume that bandits are unorganized and cannot credibly commit, their number is determined endogenously, such that the average revenue from theft is equal to the alternative productivity in foraging. In turn, this number of bandits determines the probability that cereals would be stolen.

In the alternative regime of 'hierarchy' a fraction of the crop is taxed by the revenue maximizing

[9] In appendix A we support 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.

[10] For brevity, we often refer to tubers only, even if implying also roots – the biological distinction between roots (such as cassava) and tubers (such as potatoes or yam) relates mostly to regenerative functions. What we are really after is a distinction among crops according to their degree of appropriability. The dichotomous distinction between two types is evidently for the sake of simplicity. We are aware that potatoes for example (a tuber), have been freeze dried in ancient Peru, and are altogether somewhat storable and not a tropical crop. For tractability, we ignore nutritional needs other than energy, and other important food plants, such as fruits, vegetables and legumes, and also the role of livestock. To the extent that these are not easily appropriable, we would lump them with tubers. For the sake of simplicity, we do not include the decision to farm or forage in our model and ignore the possibility that farmers may partly engage in foraging.

elite, and cereal crops are protected from bandits. We assume that the hierarchical elite can commit to any feasible rate of taxation. Employing tax collectors to maximize the net revenue, the state will equalize the wage of a tax collector (the alternative income from foraging) to the marginal tax revenue. The state will therefore employ less tax collectors than the equilibrium number of bandits under anarchy. Finally, we also assume that in order to be viable and deter bandits, the state has to incur some fixed costs.

The main prediction of the model is that if tubers are sufficiently productive, then a state cannot exist, since the potential revenue is insufficient to cover the fixed cost of deterring bandits. This result illustrates our claim that it isn't low agricultural productivity per se that hinders the development of hierarchy: it is the relatively high productivity of less appropriable crops, which leads farmers to prefer the in-built protection that these crops provide from tax collection by the elite.

The model also illustrates that whenever hierarchy exists, it dominates anarchy in welfare terms. Anarchy is more distortionary than hierarchy for two reasons. First, the state can commit to a lower tax rate that encourages the cultivation of cereals, while bandits cannot. This may lead farmers under anarchy to grow more tubers, even when they are less productive. Second, since the state employs less tax collectors than the equilibrium number of bandits under anarchy, forgone productive foraging – the alternative cost of banditry/tax collection - is lower under hierarchy. Even if the elite is non-benevolent and seeks to serve only its own interest, the state enhances productivity by protecting farmers from bandits and by committing to predictable and relatively low taxation. As a result, cereal based farming, which renders farmers vulnerable to taxation, leads to the development of a state, and contributes to farmers' welfare.[11]

In the empirical section of the paper (section 4) we first show that controlling for productivity, the adoption of farming is more likely in areas suitable for roots and tubers rather than areas suitable for cereals, as consistent with our claim that the former provide in-built protection from confiscation. We then demonstrate how geography, through its effect on the type of cultivated crop, can explain differences in hierarchy and institutions, and that the data is consistent with our proposed appropriability mechanism, and is inconsistent with the productivity-and-surplus

[11] We do not consider 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 or the importation of better farming techniques.

theory. That is, we show that a sufficient productivity advantage of cereals over tubers generates hierarchical institutions and states, whereas absolute land productivity does not.

This result is obtained using two alternative data sets with information on social hierarchy: a cross section of societies and a panel of countries. For our cross section analysis we use Murdock's (1967) Ethnographic Atlas, which contains information on cultural, institutional and economic features of 1,267 societies from around the world, at an idealized time period of first contact with Europeans. In this cross-section analysis, our main outcome variable is "Jurisdictional Hierarchy beyond the Local Community." The Ethnographic Atlas provides information also on the major crop type for the societies that practice agriculture. However, since crop choice is endogenous, we instrument for crop type using data matched modern data on the suitability of soil for different crops from the FAO-GAEZ.

From the FAO data we calculate two measures for each region: the maximal caloric yield per hectare of the most productive cereal and of the most productive root or tuber. From these measures we construct our two main regressors for the analysis. First, we calculate for each region absolute soil fertility: the maximum caloric yield from either crop type. Second, we construct for each region a measure of productivity advantage of cereals over roots and tubers: the difference between the maximal caloric yield of cereals and that of roots and tubers.

The outline for our empirical analysis is as follows:

1. Using potential agricultural yields as an instrument for actual crop choice, we start by showing that the decision to cultivate cereals as a main crop depends positively on the productivity advantage of cereals over roots and tubers. We also demonstrate that, controlling for this productivity advantage of cereals, crop choice does not depend on the productivity of the land itself.

2. We investigate the reduced-form relationship. We find that, in line with our theory, when cereals are more productive than roots and tubers, societies have, on average, a more complex hierarchal structure. Moreover, we find that land productivity does not affect hierarchy, once we control for the productivity advantage of cereals. Thus, we challenge the standard argument that it is higher land productivity that leads to more hierarchical societies.

3. We report the 2SLS estimates, where we use the potential productivity advantage of cereals over roots and tubers as an instrument for the actual growing of cereals. Consistent with our theory,

we find that cultivating cereals has a considerable positive impact on the hierarchical complexity of the societies in our sample. This result holds when we control for land productivity. Moreover, when comparing how hierarchy is affected by agriculture in general and by the cultivation of cereals, we find that only the latter matters. That is, societies that practice agriculture are more hierarchical only where they cultivate cereals. This means that societies that cultivate roots or tubers have a similar level of hierarchy to that of non-farming, pastoral or foraging, societies.

4. Finally, we show that cereals are necessary to sustain more complex hierarchical structures. We use data covering a subset of the societies in the Ethnographic Atlas for which we also have information on the sources of political power, wealth and influence that contribute to the status of the politically dominant class in society (Tuden and Marshall, 1972). The 2SLS estimates confirm that it is only in areas where cereals are cultivated that the most prestigious members of society do not derive their income from their own subsistence activities.

These cross-sectional findings support our theory that it is not agricultural productivity and surplus that explain more complex hierarchical societies. It is the productivity advantage of cereals over roots and tubers, the type of crop which is cultivated as a result, and the appropriability of the crop.

Although this analysis accounts for a wide range of confounding factors, we cannot rule out completely that omitted variables may bias the results. To overcome this concern, we employ a panel dataset compiled by Borcan, Olsson and Putterman (2014). This dataset is based on present-day boundaries of 159 countries, with institutional information every five decades over the last millennium. With this data we can 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 the productivity of land and the productivity advantage of cereals over roots and tubers in the majority of the countries in our sample. Consistent with our theory, the panel regressions confirm that an increase in the productivity advantage of cereals over roots and tubers has a positive impact on hierarchical complexity, and an increase in land productivity does not.

Since the related literature is vast, we present our literature survey in three parts. In section 2 we examine selected recent studies that relate various aspects of farming to social outcomes. In section 5 (after presenting our empirical findings in section 4) we discuss qualitative anthropological and archaeological evidence on the role of storage in facilitating hierarchy and other pertinent evidence for our theory. In section 6 we seek to demonstrate that the productivity and surplus theory is the leading explanation for the emergence hierarchy and not a straw-man of our own creation; we then survey a number of alternative theories for explaining this phenomenon.

2 Related literature

Scott's The Art of Not Being Governed (2009) makes closely related claims. He asks why states in South East Asia emerged only in the river valleys and not in the highlands. His theory is that the tribal hill people could resist being subjugated by states by adopting foraging and shifting agriculture – an option that was not open in the lowlands, where the cultivation of (appropriable) rice is the norm.[12]

According to Dow and Reed (2013), inequality and hierarchy emerged after the adoption of agriculture, when the individuals who gained control over the most fertile land organized in order to exclude outsiders, and employ these outsiders as workers. Boix (2015) offers another variant of the conflict theory, in which exogenous technological shocks related to the transition to agriculture caused inequality between insiders, who were able to benefit from the new technologies, and outsiders. This inequality broke down the social order of statelessness and cooperation and led outsiders to raid the more productive insiders, leading to the emergence of two types of states to protect the insiders: dictatorships formed by outside bandits who turned stationary, or republics managed by the insiders.[13]

Our contribution relates closely to the analysis by Besley and Persson (2009, 2014) on the high

[12] Alongside the similarities, there are many differences between Scott's approach and ours. To begin with, his interest is not in how hierarchy and states emerged, but rather in how highland tribes were able to evade nearby states. His key distinctions relate to geographic elevation: between hill and valley people, and to the form of cultivation: sedentary vs. shifting, rather than to the choice of the cultivated crop: cereals vs. roots and tubers. While his theory may be applied to regions other than South East Asia, it fails totally in the case of South America, where the Incas had a powerful state in the mountains, while no states emerged in the Amazon valleys (see section 6 below). Moreover, in our perception the essential conflict is between foragers/bandits who pose a threat to farmers. In Scott's framework, it is the state that encroaches on highland tribes. Furthermore, while we stick to a materialistic understanding of what drives individual farmers, Scott presumes that what drives people to the highlands, is an overriding anarchic motivation to escape the state and retain freedom.

[13] We note that in both of these theories, it is an exogenous increase in productivity and inequality that generates conflict, whereas in our theory, it is simply the vulnerability of cereal-cultivating farmers that triggers potential conflict.

correlation between underdevelopment and low fiscal capacity.[14] They view this correlation as a vicious cycle that could be broken by investment in the ability to tax, which would create a more effective government and alleviate underdevelopment. Our theory proposes that low fiscal capacity may be an inherent problem. In a closely related study, Mayshar, Moav and Neeman (2014) claim that a state's capacity to tax and land tenure institutions depend on the transparency of production, suggesting an alternative link between geographic attributes and fiscal capacity.

Alesina, Giuliano and Nunn (2013) employ similar data to that used here on pre-industrial agricultural patterns (from the Ethnographic Atlas) and on land suitability (from the FAO). In analogy to our distinction between the cultivation of cereal or roots and tubers, they distinguish between regions where farmers prepare the soil by use of the plough, and regions of shifting cultivation where farmers use handheld tools like the hoe and the digging stick. The social phenomenon that they study is quite different, for they study beliefs about gender equality among descendants of societies that practiced these two types of agriculture. But if one considers gender inequality as analogous to political hierarchy, then our thesis and findings are quite comparable.[15]

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, and that tubers are non-seasonal and thus harder to expropriate than cereals. To simplify, we take the latter distinction to an extreme by positing that tubers cannot be expropriated, and in this sense are just like the product of hunting and gathering. We model farmers' choices of what crop to grow in two different regimes: anarchy and hierarchy, and derive conclusions regarding the circumstances under which hierarchy can emerge.

The economy is populated by a measure one of farmers and a measure N of non-farmers. We consider the productivity of cereals as constant, and, with crop units measured by their nutritional

[14] Gennaioli and Voth (2015) emphasize how investment in state capacity since the Middle Ages responds 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.

[15] In another related study, Nunn and Qian (2011) show that the adoption of the potato in Europe following the Columbian exchange led to population growth. Their interpretation is that this was due to the higher caloric yield of the potato in regions highly suitable for its cultivation. We suggest that farmers may have adopted the potato because it provided them with some immunity against taxation/theft. The higher net produce for farmers' own consumption may have served as another channel to explain the observed population growth.

value, normalize it to unity. Our main exogenous variable, δ, measures the productivity advantage of cereals over tubers, or the productivity disadvantage of tubers.[16] Thus, farmers can grow one unit of cereals, or $1-\delta$ units of tubers, or any linear combination thereof. Hence, a farmer who allocates a fraction $\beta \in [0,1]$ of his land to cereals and a fraction $1-\beta$ to tubers produces $\beta + (1-\beta)(1-\delta) = (1-\delta) + \beta\delta$ units of output.

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 also to be bandits, in which case their income would consist of the output they expropriate from farmers. In a state of hierarchy, we assume that some non-farmers are hired by the state (or by the non-working elite) to serve as tax collectors, and are paid the wage s. We denote by λ the measure of bandits or tax collectors. This is then also the number of bandits, or tax collectors, per farmer. N is assumed to be large enough so that the measure of foragers, $N-\lambda$, is positive.

3.1 Anarchy

Under anarchy, farmers face a risk of a raid by bandits. We assume that a raided farmer loses his entire cereal crop, but none of his crop of tubers. Farmers are assumed here to be risk neutral.[17] A farmer who is facing a raid with probability τ thus chooses β to maximize his expected income I:

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

The probability τ, also referred to as the rate of expropriation, is a function of the measure of bandits λ, $\tau = \tau(\lambda)$. We assume that 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 $\lim_{\lambda \nearrow \infty} \tau(\lambda) = \overline{\tau} \leq 1$.[18] The

[16] The variable δ can be either positive or negative. However, if it is negative, then tubers dominate cereals in the sense of providing both protection and higher productivity, so that farmers would only grow tubers in equilibrium. As a result, the equilibrium would be a simple one: hierarchy would be impossible, and also anarchy would be degenerate. Given our simplifying assumption that tubers are immune from confiscation, this situation generates a third type of regime, which presumably prevails among hunter-gatherers, where there is nothing to steal from one another. We realize, though, that to the extent that farmers who cultivate tubers are sedentary whereas foragers are not, there is likely to be a difference in the social institutions also between bands of foragers and farming groups that rely on non-appropriable crops. However, we choose to focus the analysis on the more interesting case where δ is positive, and where cereals may be cultivated, which creates a tradeoff between protection and productivity considerations.

[17] In section 3.5 we show that our results are robust to the introduction of risk aversion.

[18] Micro-foundation for the shape of $\tau(\lambda)$ can be obtained by assuming that banditry is time consuming, that bandits are not coordinated, and thus that as their number increases their marginal theft declines due to increased probability of raiding the same farmers.

inverse function of $\tau(\lambda)$ is denoted by $\lambda(\tau)$. Our assumptions imply that $\lambda(\cdot)$ is strictly increasing, strictly convex, with $\lambda(0) = 0$, $\lim_{\tau \searrow 0} \lambda'(\tau) = 0$ and $\lim_{\tau \nearrow \overline{\tau}} \lambda'(\tau) = \infty$.

As formulated in (1), it is evident that in selecting how much cereal to cultivate, farmers weigh the productivity advantages δ of cereals over tubers, against the disadvantage, as measured by the expropriation rate τ. Bandits are identical and uncoordinated. Thus a bandit's expected income π is given by the total amount of cereals they confiscate from farmers divided by the measure of bandits:

$$\pi = \frac{\tau(\lambda)\beta}{\lambda}.$$

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 $\pi = s$.*[19]

The last condition can be restated as requiring:

$$\frac{\tau\beta}{\lambda(\tau)} = s.$$

Define now a threshold rate δ_A by the implicit relationship:[20]

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

δ_A provides the lower bound for the productivity advantage of cereals above which only cereals are grown.[21]

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

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

[19] Our assumptions that $\lim_{\lambda \searrow 0} \tau'(\lambda) = \infty$ and that that N is "large enough" (in particular, for the δ_A defined below, we require that $N \geq \lambda(\delta_A)$) guarantee a solution with $\tau > 0$.

[20] 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.

[21] Our assumptions on $\tau(\cdot)$ imply that that δ_A is well defined for every $s > 0$. δ_A captures the confiscation rate that will exist in equilibrium if the option to grow tubers is relevant. Thus, tubers are not grown if $\delta \geq \delta_A$.

Proof. If $\delta > 0$, an equilibrium with no cereals ($\beta_A = 0$) can be ruled out. This is 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 equilibrium can only be either mixed ($0 < \beta_A < 1$), where both crops are cultivated; or one with cereals only ($\beta_A = 1$).

If $\delta \geq \delta_A$, so that the productivity disadvantage of tubers is sufficiently high, 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, $\lambda(\delta_A)$.

In the alternative case $0 < \delta < \delta_A$, the productivity disadvantage of tubers is low. Our assumptions on $\tau(\cdot)$ imply that the confiscation rate, $\tau(\lambda)/\lambda$, or $\tau/\lambda(\tau)$, is monotonically decreasing in τ, from infinity towards zero. Thus, when $\delta < \delta_A$, 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 combination (β_A, τ_A) in the mixed equilibrium. ■

Income distribution. It follows from Proposition 1 that if cereals' productivity advantage is low ($\delta < \delta_A$) and the equilibrium is therefore mixed, the values of β_A, τ_A and $\lambda_A = \lambda(\tau_A)$ tend to zero when δ tends to zero, and are all strictly increasing in δ. As a result, also the total expected amount of cereals confiscated by bandits, $\tau_A\beta_A$, strictly increase in δ. As (1) reveals, farmers' income in that range is $1 - \delta$, thus decreasing in δ. On the other hand, when the productivity advantage of cereals exceeds the threshold δ_A, all theses variables become independent of the value of δ, with farmers income equaling $1 - \delta_A$. In these two ranges combined, the proposition thus implies that $\tau_A\beta_A$, τ_A and λ_A are all weakly increasing 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 s is the larger is 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, on τ and therefore on λ; it will rather reduce 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: $Y_0 = 1 + Ns$.

The equilibrium (β_A, τ_A) introduces two deviations from this maximal level of output: the first is due to the possibility that farmers may grow tubers (if their productivity disadvantage is sufficiently small: $\delta < \delta_A$); and the other is due to the forgone output by banditry. This means that equilibrium output is given by:

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

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, the mixed equilibrium implies $\tau_A = \delta$, which makes farmers indifferent between the two crops. It follows from the fact that expected revenue per-bandit is equal to $\tau_A \beta_A / \lambda(\tau_A) = s$ that $s\lambda(\tau_A) = \tau_A \beta_A$, and thus it follows that:

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 its revenue net of the cost of tax collection. In order to facilitate comparison between the regimes of hierarchy and anarchy, 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 at cost s per tax collector, it can generate revenue of $\tau(\lambda)\beta$ from the farming sector. In adopting Weber's definition, we also assume that a state has to be able to deter bandits, and thus has to have monopoly power over the use of force. In recognizing economies of scale in the use of force, we simplify by assuming that the army required to possess such a monopoly over the use of force entails fixed cost $G_0 > 0$.[22]

[22] We distinguish here between the cost of maintaining an army and the costs of employing tax collectors. To the extent that these functions overlap, G_0 may be thought of as "small."

A key advantage that a state has, in comparison to anarchy, is that it is farsighted and organized, and can thus commit not to expropriate farmers beyond a certain tax rate.[23] That is, the state selects the number of tax collectors to maximize its net revenue, taking into account farmers' response to the implied tax rate. Farmers' freedom to choose to avoid taxation completely by cultivating tubers, implies that the state cannot gain from setting a tax rate higher than δ. Thus, the objective of the state is to choose a tax rate τ, and thus to hire $\lambda(\tau)$ tax-collectors at cost $s\lambda(\tau)$, to maximize its net revenue, subject to the constraint that farmers respond optimally to the tax rate:

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

subject to

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

Since it is evident that $\beta = 0$ if $\tau > \delta$ and $\beta = 1$ if $\tau < \delta$, we assume that $\beta = 1$ if $\tau \leq \delta$, and note that the state's problem is in fact to choose τ to maximize $\tau - s\lambda(\tau)$, subject to $\tau \leq \delta$. The optimal tax rate under hierarchy is thus: $\tau_H(\delta) = \min\{\delta, \delta_H\}$, where δ_H is the parameter that solves $s\lambda'(\delta_H) = 1$. At a very low range of tubers' productivity disadvantage, where, $\delta < \delta_H$, $\tau_H = \delta$ and $R(\tau_H(\delta)) = \delta - s\lambda(\delta)$, increases in δ. Our assumption that the state is viable only if it sustains an army at a fixed cost $G_0 > 0$ sets a lower limit on net revenue. Thus we assume that these fixed costs are low enough to satisfy: $R(\tau_H(\delta_H)) > G_0$. We also define then the viability threshold $\underline{\delta} < \delta_H$, such that: $R(\tau_H(\underline{\delta})) = G_0$.

We have thus established:

Proposition 2. (i) *If δ is small ($\delta < \underline{\delta}$), then a state cannot exist. (ii) If δ has an intermediate value ($\underline{\delta} \leq \delta < \delta_H$) then the optimal tax rate set by the state is given by $\tau_H = \delta$. (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

[23] 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. Below we show that our qualitative results hold also when farmers are risk averse.

tubers δ. Total tax receipts equals τ_H, and the net tax revenue received by the elite, after paying the tax collectors and covering the cost of the army is:

$\tau_H - s\lambda(\tau_H) - G_0$. Both the gross and net tax receipts strictly increase in δ up to the threshold δ_H, where they remain constant.

Output Loss. Analogously to the case of anarchy, we define the efficiency loss for hierarchy as the deviation of total output from the maximal potential:

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

Corollary 2. The output loss $(Y_0 - Y)$ due to hierarchy is:

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

.

3.3 Anarchy vs. hierarchy

As explained in the previous section, a state can only exist if tubers are sufficiently unattractive to farmers, that is, if their productivity disadvantage 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 $\underline{\delta}$.

Proposition 3. *If δ is small ($\delta < \underline{\delta}$), then only anarchy is possible, with a mixed equilibrium in which $\tau_A = \delta$ and where both cereals and tubers are grown. If δ is high enough for the state to be viable ($\delta \geq \underline{\delta}$), then a hierarchy weakly Pareto dominates anarchy.*

Proof. 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: $\tau - s\lambda(\tau) - G_0 > 0$. ■

Proposition 4. In the range where hierarchy is viable, the economy is more productive under hierarchy than under anarchy.

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

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

By the definition of $\underline{\delta}$, $R(\underline{\delta}) = \underline{\delta} - s\lambda(\underline{\delta}) = G_0$ so that the output gap between the two regimes is zero when $\delta = \underline{\delta}$. 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 (when $\underline{\delta} > \delta$), 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 taxes less, since it sets the scale of tax collectors so that their marginal product equals their cost s, whereas under anarchy it is the average product of bandits that equal s. As a result, (weakly) fewer non-farmers are engaged in non-productive appropriation.

The main predictions of the analysis

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

2. If tubers are sufficiently productive in comparison to cereals ($\delta < \underline{\delta}$), then a state cannot exist. This result illustrates our claim that it isn't low productivity that restrains the development of hierarchy and related institutions, but rather high productivity of crops that are hard to

expropriate. If, however, the reverse is true ($\delta > \underline{\delta}$) hierarchy could emerge and farmers would produce food surplus that would be taxed by the elite.

3. Whenever it exists, even a non-benevolent state that monopolizes coercive force dominates anarchy efficiency-wise (Propositions 3 and 4). This is a result of our assumption that the state can commit to a tax rate that maximizes its revenue net of collection costs, and that consequently farmers cultivate only the more efficient cereals.

We test predictions 1 and 2 in the empirical section below. Before turning to that section, we analyze a simple example that enables us to present the model's predictions diagrammatically and to examine also the case of risk aversion.

3.4 Example

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 the equilibrium under anarchy is given by

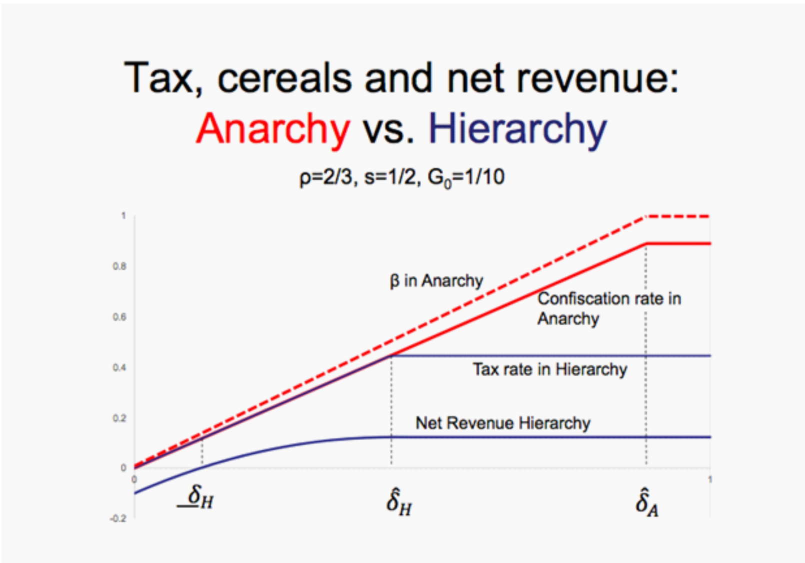
$$(\beta_A, \tau_A) = \begin{cases} \left(\frac{s\delta}{\rho^2}, \delta\right) & if \quad \delta < \delta_A \\ \\ \left(1, \frac{\rho^2}{s}\right) & if \quad \delta \geq \delta_A \end{cases} .$$

Under hierarchy, $\delta_H = \alpha\rho^2/s$ and the lower limit for state existence, $\underline{\delta} > 0$, is implicitly defined by the quadratic equation: $\underline{\delta} - s\left(\frac{\underline{\delta}}{\rho}\right)^2 = G_0$.[24]

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 $\delta = \delta_H$ upon 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 2 illustrates the efficiency advantage of hierarchy in comparison the anarchy.

[24] The existence of such a positive solution is conditioned on: $G_0 \leq \rho^2/4s$.

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



Risk-averse farmers

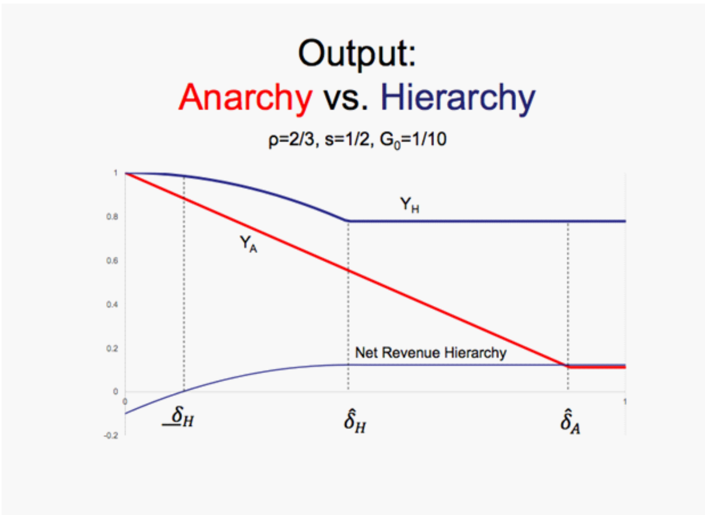
In this subsection we illustrate the robustness of the model's qualitative predictions when farmers are risk averse. The results are in a sense even stronger, given that risk-averse farmers under anarchy seek more protection by choosing a smaller share of cereals. Farmers' risk aversion does not affect 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 above specification of the expropriation function, $\tau(\lambda) = \rho\sqrt{\lambda}$, 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:

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

The solution is

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

Figure 2: Output: Anarchy vs. Hierarchy



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 $\beta_A > 0$):

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

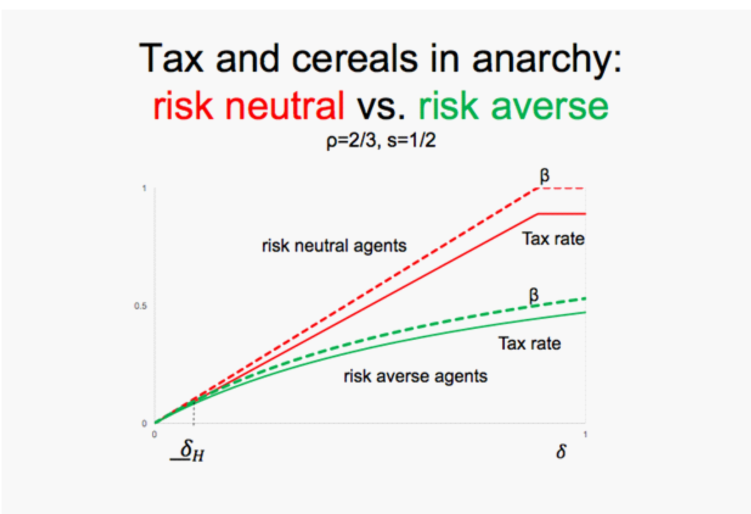
Inspection of the equilibrium values of (β_A, τ_A) reveals that as δ tends to zero, both β_A and τ_A tend to zero. As δ increases towards one, τ_A approaches $\rho^2/(\rho^2 + s)$ and β_A approaches $s/(\rho^2 + s)$. This implies that even in the limit, when the productivity of tubers approaches zero, 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.[25] Figure 3 illustrates the difference between the two types of equilibrium: the case of risk neutral farmers and risk averse farmers.

Figure 3: Output: Anarchy vs. Hierarchy



[25] 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.

4 Evidence

In this section, we provide supportive evidence for our main theoretical predictions. We employ two alternative datasets with information on social hierarchy: a cross section of societies and a panel of countries. Our main regressors are two measures of agricultural productivity: the productivity of the soil (the maximum caloric yield that can be obtained from a given unit of land) and the productivity advantage of cereals over roots and tubers (the difference between the maximum calorie yield that can be obtained from cereals and roots or 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.

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 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).[26]

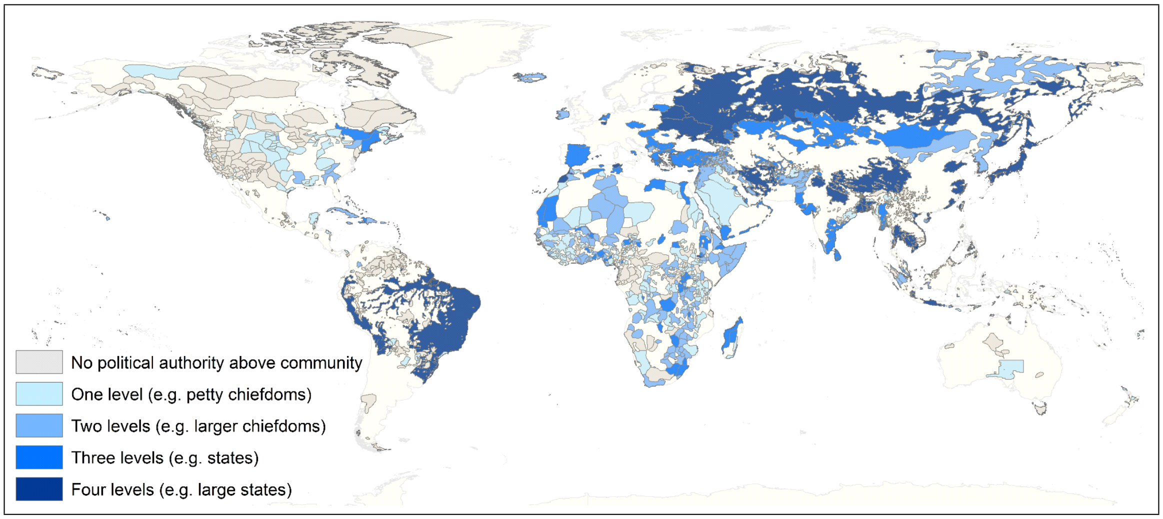
We measure pre-colonial hierarchical complexity using the variable "Jurisdictional Hierarchy beyond the Local Community."[27] This is an ordered variable with five possible levels: (i) no political authority beyond community, (ii) petty chiefdoms, (iii) larger chiefdoms, (iv) states, and (v) large states. We plot this measure of hierarchy in Figure 4 and present the summary statistics in the first row of Table 1. The majority of our sample is composed of societies lacking any political integration above the local community, and groups where petty chiefs rule over very small districts.

[26] 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).

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

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 tend to prevail.

Figure 4: Jurisdictional hierarchy beyond the local community in pre-colonial societies

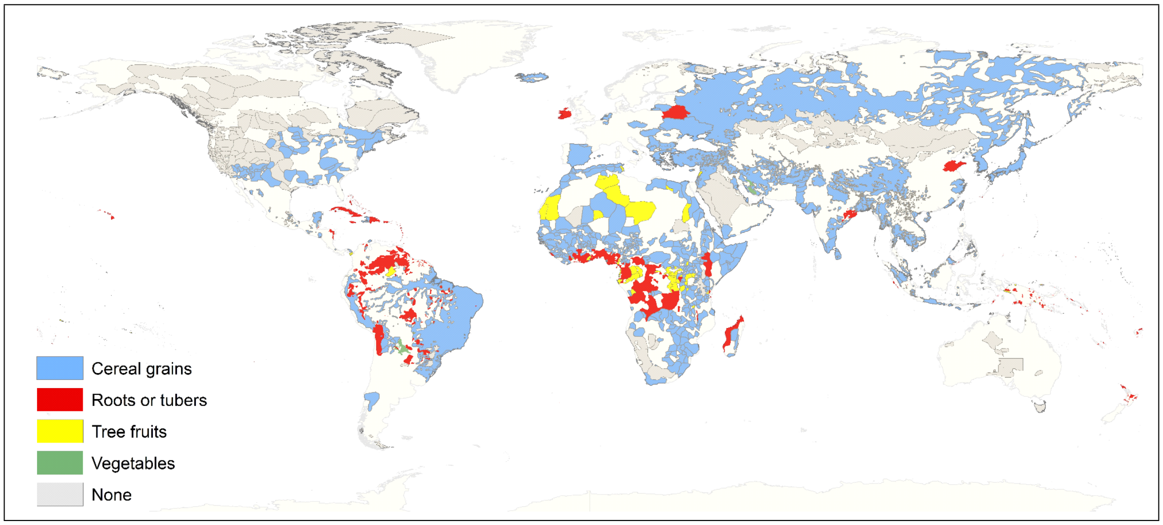


The Ethnoatlas also provides information on the reliance of these societies on agriculture for their diet, and on the major crop type of societies that practice agriculture. These two variables are plotted in Figure 5. with summary data in rows 2 and 3 of Table 1. As can be seen from Figure 5, approximately 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% 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.[28] 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 1.

The second source of ethnographic information is provided by the Standard Cross-Cultural Sample (SCCS), which is a derivative of the Ethnographic Atlas. This data is based on a representative sample, defined by Murdock and White (1969), of 186 societies taken from the Ethnoatlas. A large

[28] 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 5: Major crop in pre-colonial societies



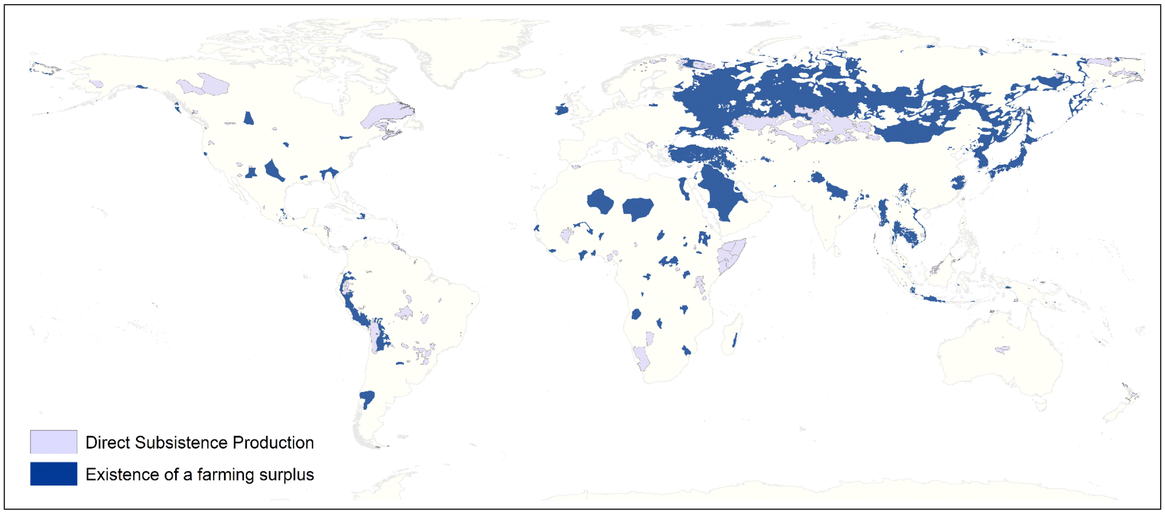
number of publications by diverse authors coded the SCCS societies for many different types of societal characteristics. Cumulative ethnographic codes and codebooks are published in the World Cultures electronic journal.

We use two variables from the SCCS (rows 4 and 5 in Table 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 support from their own subsistence activities and one otherwise. This dummy is plotted in figure 6. 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, while the last one societies with 500+ persons per square miles).

4.1.2 Country-level data

At the country level, we construct a hierarchy index using data from Borcan, Olsson and Putterman (2014). This data cover 159 modern-day countries for every half century from 50 CE to 2000 CE. 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 points if the organization of the state can be at best be described as a paramount chiefdom, and 0 points if the answer is no. This data

Figure 6: Farming surplus in pre-colonial societies



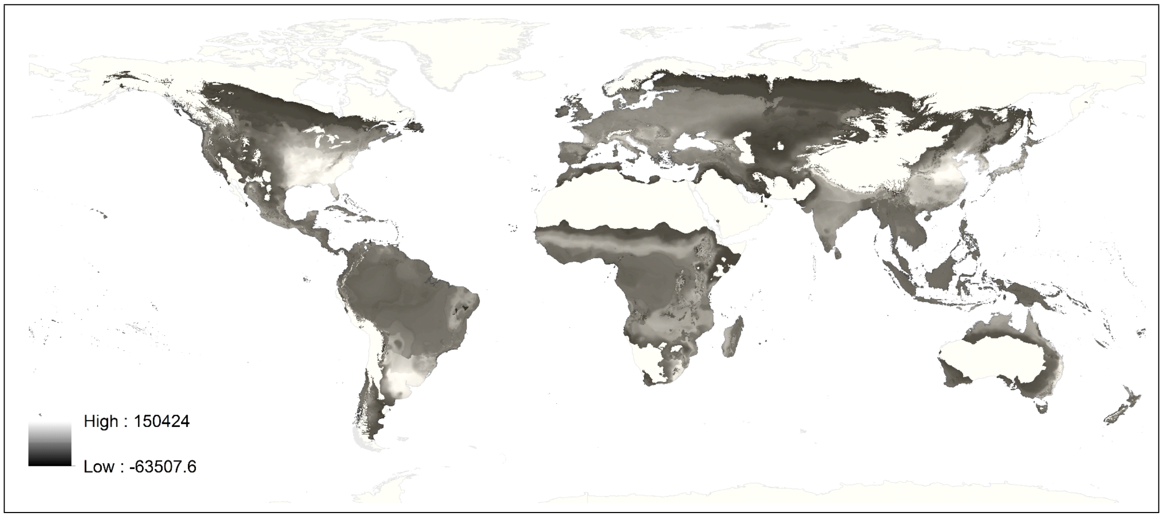
is merged with data on: the legal origin of the country (from La Porta et al., 1999); population density in 1500 (Acemoglu, Johnson and Robinson, 2002); mortality of early settlers (Acemoglu, Johnson and Robinson, 2001); and the number of exported slaves (Nunn, 2008).

4.1.3 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 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 capture the conditions that were prevalent before the first significant contact of the societies in the Ethnoatlas with Europeans, and to exclude problems of reverse causality, we consider potential yields based on agro-climatic conditions under rain-fed low-input agriculture.

GAEZ provides data on potential yields, in terms of tons per hectare per year, for 11 cereal

Figure 7: Difference in potential yields (calories per hectare) of cereals versus roots and tubers.



grains and 4 roots and tubers. Following the same procedure as in Galor and Ozak (2014), 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. We then find the crop with the highest potential caloric yields for each raster point. The results are illustrated in figure C.3 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.

On the basis of this data set we construct two measures: a measure of the productivity of land, measured as the maximum potential caloric yield per hectare; and a measure of the productivity advantage of cereals over roots and tubers, which equals the difference between the maximum caloric yield of cereals and the maximum caloric yield of roots or tubers. These measures 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.[29] The two measures are attributed to the different countries by using the FAO country boundaries.

[29] 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).

4.1.4 Other demographic and geographic data

The History Database of the Global Environment (HYDE) supplies global estimates on population density at the raster level between 1500 and 2000 with cell size of 5'x5'. To each society in the Ethnoatlas, we assign a value that is equal to the average population density across the raster points within its territories for the year of observation recorded in the atlas. The median community had historical population density of 41 inhabitants per square mile, while the community with the highest population density was the Okinawans in Japan with 3627 inhabitants per square mile. Data on population density for 1995 is provided by GAEZ and is similarly averaged within the territory of each society. Finally, we employ data on distance to major rivers or to the coast, precipitation, temperature, elevation, ruggedness, absolute latitude, incidence of malaria both at the society and the country level. Sources are detailed in Table 1.

Table 1: Descriptive Statistics

	SOURCE	Mean	p50	SDev	Min	Max	N
PANEL A: Societies in Ethnoatlas							
Hierarchy beyond Local Community	Ethnoatlas	1.89	2.00	1.04	1.00	5.00	1,059
Major Crop: Cereals	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
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
Cal/ha Best Crop (std)	authors	0.00	0.23	1.00	-1.92	2.66	1,179
Cal/ha Cereals- Cal/ha Tubers (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)	FAO-GAEZ	0.00	0.17	1.00	-9.24	3.58	1,179
Ruggedness (std)	FAO-GAEZ	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
PANEL B: Countries X 50 years							
Hierarchy index	Borcan et al. (2014)	0.72	1.00	0.45	0.00	1.00	2,869
Cal/ha Best Crop (std)	authors	0.00	0.35	1.00	-1.64	2.69	2,959
Cal/ha Cereals- Cal/ha Tubers (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)	FAO-GAEZ	0.00	-0.33	1.00	-1.10	4.65	2,845
Ruggedness (std)	Nunn and Puga (2012)	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
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	Nunn and Puga (2012)	0.35	0.00	0.43	0.00	1.00	2,959

std - a standardized variable that has been rescaled to have a mean of zero and a standard deviation of one.

4.2 Empirical results

4.2.1 The choice of crop

We start our empirical analysis by studying the geographical factors that influence the choice of cultivating cereals rather than alternative crops or non-farming. Our theory suggests that farmers make this choice on the basis of comparing the net caloric yield of cereals to that of the alternatives crops (in which we focus on roots and tubers).

The first three columns of Table 2 presents the results of the following regression:

$$Cer_i = \alpha CalDiff_i + X_i'\beta + \varepsilon_i.$$

Cer_i is a dummy variable that identifies that society i cultivates a cereal grain as its main crop; $CalDiff_i$ is the caloric advantage of cereals in the land of society i (the difference between the maximum potential calorie yield of cereals and of roots or tubers); and X_i is a set of control variables. Column 1 reports the bivariate relationship 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 planting cereals as main crop in the order of 20 percent. Moreover, variation in this regressor alone is able to explain 13 percent of the entire variation in the dependent variable. The first concern is that the productivity advantage of cereals might reflect the potential caloric yield of the soil, since cereals grains are the most productive crops in most of the world. Column 2 reports the results when adding as a control variable the productivity of the soil (when cultivating the crop with the highest potential caloric yield). This variable does not produce any significant impact on the decision on whether to plant cereals or not, while the impact of the productivity advantage of cereals is unchanged. Adding this control leaves the R^2 of the regression practically unchanged, which suggests that soil productivity isn't relevant to explain the decision to cultivate cereals. Column 3 shows that the results are unchanged when only exploiting within-continent variation. Moreover, the results of the first three columns of Table 2 survive a battery or robustness checks that are detailed in the appendix of the paper. In table C.1, we control sequentially for precipitation, temperature, elevation ruggedness and absolute latitude, which are the main factors affecting crop productivity in the GAEZ dataset. In table C.2, we control for geographical isolation (proxied as the distance to the nearest major river or coast), malaria endemicity and actual and historical population density. In all cases, the qualitative results on the effect of the productivity advantage of cereals over roots and tubers are almost unaffected (coefficients vary from 0.139 to 0.276 and are always statistically significant at the 1 percent confidence level).

Table 2: Potential Crop Yields, Choice of Crops and Reliance on Agriculture

	Dependent variable is:					
	Major crop is cereal grains (dummy)			Reliance on agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)
CALORIC DIFF	0.205***	0.210***	0.253***	0.0812***	-0.0978***	-0.0464***
(CER - TUB)	(0.0168)	(0.0310)	(0.0329)	(0.00945)	(0.0134)	(0.0136)
MAX CALORIES		-0.00664	-0.137***		0.230***	0.128***
(ALL CROPS)		(0.0338)	(0.0386)		(0.0153)	(0.0178)
CONTINENT FE	NO	NO	YES	NO	NO	YES
r2	0.132	0.132	0.359	0.0733	0.235	0.387
N	982	982	982	1063	1063	1063

The table reports cross-sectional OLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is either a dummy that identifies societies that cultivate cereal grains as main crop (columns 1-3) or the reliance of the these societies on agriculture (columns 4-6). CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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.

The last three columns of Table 2 repeat the analysis of the first three columns, but with the reliance of the society on agriculture as the dependent variable. The results are striking. First, as expected, land productivity increases the probability of reliance on farming. Second, in line with our theory, the productivity advantage of cereals has a negative effect on practicing agriculture, when controlling for soil productivity.

Thus, in this subsection, we showed that while the productivity advantage of cereals has a positive impact on the probability of cultivating cereals as the main crop, it has a negative impact on the reliance of societies in our sample on agriculture. Furthermore, the absolute productivity of land has a positive impact on reliance on agriculture, but no significant impact on the probability of cultivating cereals.

4.2.2 Cereals and hierarchy

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, we estimate a regression of the form:

$$Y_i = \alpha Cer_i + X_i' \beta + u_i, \tag{2}$$

where Y_i is either a measure of hierarchy or an indicator for the presence of farming surplus in the society i; Cer_i, is, as mentioned above, a dummy variable that identifies societies that rely mainly on cereals for their subsistence; 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. In particular, according to our theory it is riskier to cultivate cereals in societies characterized by low state capacity, and thereby by low protection against bandits, since cereals render farmers more vulnerable to theft. 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 will run IV regressions, where we will instrument for Cer_i by using the productivity advantage of cereals, $CalDiff_i$.

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, we exploit the Columbian exchange and the effects it had on the productivity potential crops, to conduct panel regressions at the country-level that will rule out potential time-invariant omitted variables.

Before presenting the 2SLS regressions that estimate the effect of cereals on hierarchy and surplus, we report in Table 3 OLS estimates of the reduced form of the analysis. Columns 1 illustrates that the higher the productivity advantage of cereals, the higher is the level of jurisdictional hierarchy that is reached by the societies in the Ethnoatlas. This result is unchanged when controlling for the productivity of the soil (column 2). More specifically, while one standard deviation increase

Table 3: Cereals, Surplus and Hierarchy - Reduced Form

	Dependent variable is:					
	Jurisdictional Hierarchy Beyond Local Community			Existence of farming surplus		
	(1)	(2)	(3)	(4)	(5)	(6)
CALORIC DIFF	0.244***	0.179**	0.274***	0.141***	0.241***	0.202***
(CER - TUB)	(0.0394)	(0.0732)	(0.0758)	(0.0319)	(0.0681)	(0.0742)
MAX CALORIES		0.0825	-0.188**		-0.132	-0.0985
(ALL CROPS)		(0.0713)	(0.0886)		(0.0870)	(0.0985)
CONTINENT FE	NO	NO	YES	NO	NO	YES
r2	0.0416	0.0429	0.249	0.0757	0.0911	0.157
N	952	952	952	140	140	140

The table reports cross-sectional OLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is either a dummy that identifies societies that produce a farming surplus or Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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.

in the relative productivity of cereals increases the hierarchy index by 0.18 (0.27 in the specification with continent fixed effects), an increase of soil productivity does not produce any significant impact on the dependent variable. In column 3, we control for continent fixed effects. The impact of the relative productivity of cereals becomes larger, while the impact of the soil productivity becomes negative. Columns 4-6 provide further support for the appropriability hypothesis versus the productivity-surplus hypothesis. In fact, the higher the productivity advantage of cereals, the higher is the probability of having an economy that produces a farming surplus – elite consumption isn't based on direct subsistence (column 4). When we run a horse race between the productivity advantage of cereals and the absolute productivity of the soil (columns 5 and 6), we find that only the former has a significant impact on surplus, independently on whether we control for continent fixed effects or not.

Table 4 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

Table 4: 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
MAIN CROP:	0.707***	1.170***	0.863**	1.040***	0.304***	0.892***	1.064***	0.993***
CEREALS	(0.0630)	(0.195)	(0.364)	(0.245)	(0.0762)	(0.261)	(0.332)	(0.277)
MAX CALORIES			0.0811				-0.0368	
(ALL CROPS)			(0.0714)				(0.0564)	
DEPENDENCE ON				0.334				-0.419
AGRICULTURE				(0.298)				(0.644)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	952	952	952	952	952	952	952	952
F excl instrum.		145.6	42.53	63.39		95.00	58.58	22.37
A-R Test (p-val)		0.000	0.0147	0.000		0.000	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. The dependent variable is Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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.

increase of 0.70 in the hierarchy measure. Clearly, this positive association cannot be interpreted as causal. In order 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 the following two columns, we run a horse race between our appropriability hypothesis and the land productivity-surplus hypothesis. In column 3, we add the productivity of land as a control variable. As can be seen, it does not have any significant effect on hierarchical complexity. In column 4, we add the dependence of the society on agriculture as a second endogenous variable. The instruments now are both the caloric advantage of cereals and absolute soil productivity; where the intuition is that the latter influences only the decision whether to become farmers, but not the choice of the crop. The results are once gain striking: 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.

Table 5: 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
MAIN CROP:	0.359***	0.940***	0.846***	0.846***	0.299***	1.005***	0.797**	0.799**
CEREALS	(0.0791)	(0.260)	(0.273)	(0.275)	(0.0901)	(0.316)	(0.314)	(0.317)
MAX CALORIES			0.0186				0.0361	
(ALL CROPS)			(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
A-R Test (p-val)		0.000	0.000	0.000		0.000	0.00878	0.000

The table reports cross-sectional OLS and 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is a dummy that identifies societies that produce a farming surplus. The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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.

The results of Table 4 survive a battery or robustness checks that are detailed in the appendix of the paper. In Table C.3, we control sequentially for precipitation, temperature, elevation ruggedness and absolute latitude, the main factors affecting crop productivities. In Table C.4, we control for geographical isolation, malaria endemicity and actual and historical population density. In 9 out of 10 cases, cultivating cereals as main crop exerts a statistically significant impact on hierarchical complexity. The results are also practically unaffected when using ethnic boundaries as defined by Fenske (2013) to extract data on crop productivities (Table C.5), or when the sample includes societies living in desertic soils (Table C.6). In all cases, the qualitative results on the effect of cultivating cereals as main crops are almost unaffected (the coefficients vary from 0.475 to 0.900).

Table 5 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 (column 1). The coefficient more than doubles when turning to the 2SLS estimates (column 2). As in the previous table, also in this case absolute productivity of soil and reliance on agriculture do not affect the dependent variable (columns 3 and 4); and the results are robust when adding continent fixed effects in the specification (columns 5 to 8). Also in this case, the empirical results survive a long list of robustness checks reported in the appendix (Tables C.7-C.10).

These results provide evidence in support of our theory, as they indicate that the decision to cultivate cereals is crucial to develop complex hierarchical institutions and a farming surplus. This analysis accounts for a large set of possible confounding geographical characteristics. But still, we cannot rule out that unobservable characteristics, systematically correlated with the productivity of different crops, might be driving our results. In order to overcome this potential concern, we exploit an exogenous variation in the available crops in different locations induced by the Columbian exchange.

In the New World, the only available roots and tubers before 1500 were cassava, white potatoes and sweet potatoes; while the only available cereal grain was maize. In the Old World, the only available crop among roots and tubers were yams; while the available cereal grains were barley, buckwheat, foxtail millet, indigo rice, oat, pearl millet, rye, sorghum, wetland rice, and wheat – but not maze. Thus, for each raster point of the world we define the highest yielding crop among cereals and among roots and tubers both before and after the Columbian exchange. We then compute for each location the productivity advantage of cereals over roots and tubers and the absolute productivity of land before the Columbian exchange (prior to 1500) and after the Columbian exchange (in the years after 1600).[30]

Since the data in the Ethnographic Atlas pertains only to societies after the Columbian exchange, we exploit a different country-level panel dataset that reports on hierarchical complexity for the majority of the world over the last millennium. The unit of observation is the territory

[30] We exclude the years from 1500 to 1600 as the historical evidence points out that the New World's crops were adopted in Europe and Africa in the seventeeth century. For instance, the adoption of the potato in the Old World began in the late seventeeth 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). In the appendix, we show that our results are robuts when excluding the years between 1500 and 1750 (see Table C.12).

delimited by modern-day country borders for 159 countries every 50 years. Since we lack observations on the major crop cultivated in these territories for the period of analysis, we can only run the reduced form version of our empirical analysis where we regress the hierarchy index on the productivity advantage of cereals (and on the productivity of the soil):

$$Hier_{it} = \alpha CalDiff_{it} + X_{it}'\beta + \eta_i + \eta_t + u_{it}, \tag{3}$$

Here $Hier_{it}$ is the hierarchy index of country i in year t and $CalDiff_{it} = CalDiff_{i,BeforeExchange}$ (the caloric advantage of cereals over roots and tubers before the Columbian exchange) if $t \leq 1500$ and $CalDiff_{it} = CalDiff_{i,AfterExchange}$ (the caloric advantage of cereals over roots and tubers after the Columbian exchange) if $t \geq 1600$. X_i is a set of control variables, which includes the productivity of the soil. Country fixed effects control for all time invariant factors that differ between countries, while time period fixed effects control for any time patterns of hierarchical complexity that affects all countries. The identification assumption is that there are no events that occurred in the sixteenth century and are systematically correlated with the changes that were induced by the Columbian exchange in the productivity advantage of cereals and in the productivity of land.

The results are illustrated in Table 6. Column 1 confirms that the higher the productivity advantage of cereals, the higher is the country's hierarchy index. This result is unchanged when controlling for soil productivity (column 2). More specifically, while a one standard deviation increase in the productivity advantage of cereals increases the hierarchy index by 0.19, soil productivity does not have any significant impact on the dependent variable. In the next five columns, we show that the results are robust when controlling for precipitation, temperature, elevation ruggedness and absolute latitude (interacted with the time-period fixed effects). In Table 7, we consider a host of additional factors (each interacted with time-period fixed effects) that might have affected hierarchical complexity. Our choice of controls is driven by the determinants of long-term economic development that have been emphasized in the literature. Sequentially, we control for: legal origin of the country; population density in 1500; settlers mortality; the number of exported slaves; distance to rivers and coast; endemicity of malaria; and the percentage of tropical land. Once again, our results are unaffected.

Table 6: Cereals and Hierarchy - Panel Regressions

	Dep. Variable: Hierarchy Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CALORIC DIFF	0.189***	0.272***	0.282***	0.240***	0.255***	0.261***	0.197**
(CER - TUB)	(0.0683)	(0.0834)	(0.0760)	(0.0857)	(0.0889)	(0.0839)	(0.0795)
MAX CALORIES		-0.163	-0.193	-0.152	-0.115	-0.148	-0.165
(ALL CROPS)		(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. The dependent variable is an hierarchy index: it equals 0 if there is not a government above tribal level, 0.75 if the political organization can be at best described as a paramount chiefdom and 1 otherwise. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

Equation (3) examines the average effect on the hierarchy index following the change in the productivity advantage of cereals over roots and tubers due to the Columbian exchange. This estimation requires that we choose a date in which the Columbian exchange was completed. The historical evidence suggests that it took at least a century for the Columbian exchange to happen and therefore we took the entire sixteenth century as the relevant cutoff. In order to examine whether patterns of data are consistent with this assumption, we also estimate a more flexible equation that takes the following form:

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

It is important to note that in this specification we are not particularly interested in the individual magnitudes of the point estimates but in their pattern over time. Because the main regressor is time invariant and equation (4) includes country- and time-period fixed effects, the estimated αs

Table 7: Cereals and Hierarchy - Panel Regressions - Robustness Checks

	Dep. Variable: Hierarchy Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CALORIC DIFF	0.160*	0.127	0.206*	0.274***	0.245***	0.258***	0.273***	0.254***
(CER - TUB)	(0.0892)	(0.0843)	(0.116)	(0.0833)	(0.0928)	(0.0957)	(0.0840)	(0.0675)
MAX CALORIES	-0.0507	0.0471	-0.261	-0.176	-0.121	-0.133	-0.199	-0.211**
(ALL CROPS)	(0.133)	(0.132)	(0.192)	(0.143)	(0.151)	(0.151)	(0.145)	(0.102)
Controls (x Year FE):								
Legal Origin	YES	NO	NO	NO	NO	NO	NO	NO
Pop Density 1500	NO	YES	NO	NO	NO	NO	NO	NO
Settlers Mortality	NO	NO	YES	NO	NO	NO	NO	NO
Slave Exports	NO	NO	NO	YES	NO	NO	NO	NO
Distance River	NO	NO	NO	NO	YES	NO	NO	NO
Distance Coast	NO	NO	NO	NO	NO	YES	NO	NO
Pct Malaria	NO	NO	NO	NO	NO	NO	YES	NO
Tropical Land	NO	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	0.699	0.714	0.707	0.683	0.678	0.679	0.681	0.744
N	2869	2869	1501	2869	2755	2755	2793	2869

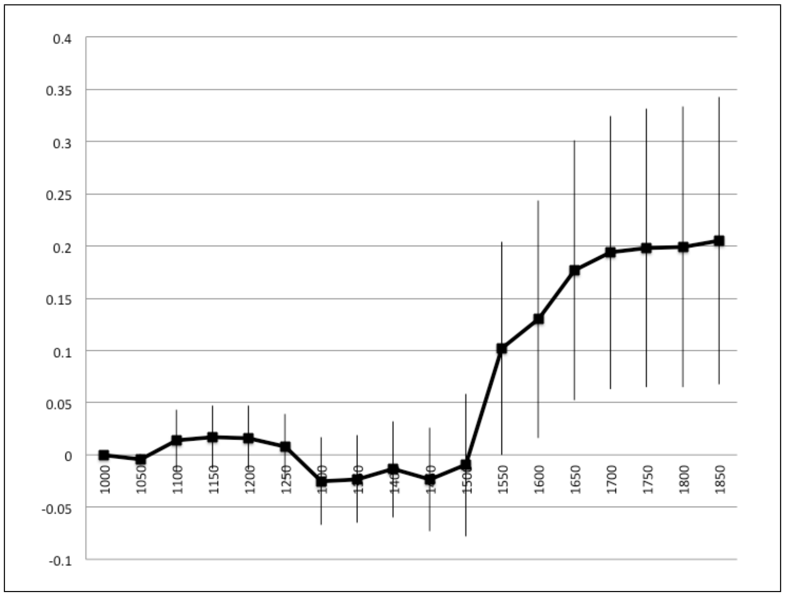
The table reports panel OLS estimates and the unit of observation is the territory delimited by modern-country borders every 50 years. The dependent variable is an hierarchy index: it equals 0 if there is not a government above tribal level, 0.75 if the political organization can be at best described as a paramount chiefdom and 1 otherwise. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. Robust standard errors, clustered at the country-level, in parentheses *** significant at less than 1 percent; ** significant at 5 percent; * significant at 10 percent.

must be measured relative to a baseline time-period, which we take to be 1000. The estimated coefficients and their 10 percent confidence intervals are reported in Figure (8).[31]

The impact of the change in the productivity advantage of cereals over tubers due to the Columbian exchange is constant over time between 1000 and 1500; it increases steadily during the sixteenth century; it continues to increase but a lower rate until 1700; after which it stabilizes. This confirms the story that the Columbian exchange produced a differential increase in hierarchy in those countries for which it also caused a larger increase of the productivity advantage of cereals over roots and tubers and that the great majority of the full impact happened in the sixteenth

[31] The 17 coefficients reported in Figure (8), 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.

Figure 8: 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.



century.

In conclusion, our empirical analysis strongly supports our appropriability theory, and does not support the alternative land productivity-surplus hypothesis. We show that the cultivation of cereals is crucial for the development of complex hierarchical institutions and for the existence of a farming surplus. On the other hand, both soil productivity and the reliance on non-cereal agriculture do not exert any effect on hierarchy and surplus.

5 Additional evidence

In the earliest phases of the transition to agriculture in the ancient Near East, known as the (late) Natufian Age (ca. 12,500-9,500 BCE), foragers adopted semi-sedentary living and relied at least in part on collecting plants, including wild cereals. However, they did not yet sow or domesticate cereals that required storage.[32] Kuijt and Finlayson (2009) report an important archeological

[32] Kuijt (2008) and Price and Bar-Yosef (2010) point out the limited archeological evidence for storage facilities from the Natufian age; and Goring-Morris and Belfer-Cohen (2011, p. S200) note that there is no evidence that the Natufians engaged in intentional plant cultivation. They apparently did not attempt to store grain over the winter, which would have necessitated more permanent, covered and plastered storage facilities to protect the stored grain

discovery in the Jordan Valley of a large and elaborate communal storage pit from about 9,000 BCE of farmer-foragers who apparently collected and possibly cultivated wild cereals. This finding reveals that storage was an integral part of the earliest phase of the transition to cereal farming, and that it involved social cooperation and organization. It indicates that, consistent with our theory, hierarchy developed alongside the gradual intensification of cereal farming and its concomitant storage, and that it was not necessarily preceded by the existence of surplus.

Our appropriability theory is, indeed, closely related to the idea that storage played an important role in the Neolithic Revolution. Since evidence of early dwellings, grinding stones and storage facilities are some of the most distinctive archeological indicators for the early phases of the Neolithic Age, it is no surprise that sedentism and storage are often cited as causal mechanisms in theories that purport to relate the transition to agriculture to the rise of social hierarchy. In this literature storage is often perceived as a marker of surplus, whereas we argue, that it is a feature that renders crops vulnerable to appropriation.

An important strand of the literature proposes that storage served for distributive purposes and the emergence of hierarchy is a result of the need for an elite to supervise this distribution. Halstead (1989) suggests that early farmers generated "normal surplus" above their subsistence needs in average years, as a precaution against years of shortage. The elite, he contends, emerged as a "social storage agent" to coordinate between surplus and deficit households. Building on Polanyi (1944), this theory posits that early agricultural societies were "redistributive," where surplus output was (voluntarily) transferred to a central authority, stored and then "redistributed," with the elite effectively serving as an agency for mutual insurance.[33] Leaving aside the plausibility of this functionalist theory, and the benevolence it attributes to the elite, we note that this interpretation misses the point that storage of cereals was primarily and necessarily intra-annual, due to the seasonality of cereals.[34] While storage of cereals as an inter-annual buffer is possible, it is unlikely that such longer-term storage was significant among the early farmers, who probably continued to

from spoilage by moisture, insects or rodents.

[33] See Johnson and Earle 2000, pp. 251-256, 301-302.

[34] Hayden (2001) questions the validity of Polanyi's presumption that the elite altruistically benefitted the periphery through distributive, insurance services, remarking (p. 247) that he "was completely astonished ... that local elites provided essentially no help to other members of the community in times of crisis." On a practical side, we conjecture that idiosyncratic shortfalls to individuals were taken care of within the small kin group (as was the case also among hunter-gatherers) and did not need proto-state institutions.

forage.

Some of the most important studies to indicate the ubiquity of storage in hierarchical societies concern complex hunter-gatherers, not farmers. In a survey of Native Americans in the northwestern coast of America, Testart (1982) concludes that hunter-gatherer societies that relied on seasonal and storable resources such as acorns or dried salmon became more complex, and acquired social features like those of the Neolithic societies that cultivated cereals. However, Testart's study only identified correlates of storage, not the causal mechanism that relates storage to inequality.[35] Even so, in conjunction with the observations of limited social complexity among farming societies in the tropics, we note that these complementary observations confirm our claim that agriculture was neither necessary nor sufficient to explain social complexity.

Tushingham and Bettinger (2013) provide further evidence about storage and hierarchy among aboriginal Californians. They note that even though salmon is a better source of nutrition, earlier foragers preferred to store acorns. They propose to resolve this puzzle by applying Bettinger's (1999) distinction between back-loaded and front-loaded food resources.[36] According to his explanation, the procurement and storage of back-loaded acorns involves little effort, but preparation for consumption is costly. The opposite pattern pertains to salmon: it is costly to catch and preserve, but consumption is easy. Tushingham and Bettinger suggest that back-loaded resources, like acorns, which require significant post-storage processing, offered earlier foragers the advantage of a lower probability of loss of caches due to pillage. They identify the key detriment to reliance on stored salmon as: "the possibility that others will rob caches, which mobile foragers are not positioned to protect" (p. 533). In addition to the greater vulnerability of front-loaded foods to theft by outside groups, they recognize also loss to "freeloaders" from the inside (p. 534). Their explanation for the late and rather abrupt transition to salmon intensification among aboriginal Californians is that reliance on salmon became feasible once a community reached a threshold size that facilitated on-going protection of salmon caches. Indeed, that transition coincided with the aggregation of people into permanent villages and with the appearance of plank houses that also

[35] In fact, Testart referred to all the "usual suspects": sedentary living, high population density, trading activity, prestige, and altered ideology. At one point, though, he referred also to the mechanism that we emphasize here: "stored food is the primary object of raids, and it may be stolen, monopolized by men of high status, or made the subject of rent or tribute" (p. 527).

[36] Woodburn (1980) offered a related distinction.

functioned as storage facilities.[37]

Tushingham and Bettinger do not mention hierarchy, nor do they relate their observations to agriculture. But their analysis is perfectly consistent with our appropriability theory on the distinction between tubers and cereals. Reliance on a food resource that offers in-built protection against robbers enables living in small groups, and does not require hierarchy to provide protection. In contrast, reliance on stored food sources that are attractive to thieves requires substantive social innovations. This explains the rise of villages and the emergence of hierarchy. Thus, we posit that our theory provides the causal mechanism of how vulnerability to appropriation leads to social hierarchy among complex pre-agricultural hunter-gatherers and among cereal-dependent farming societies – but not among other foragers or among farming communities that rely on roots and tubers.[38]

6 Alternative theories on the emergence of hierarchy: a literature survey and critique

Anthropologists and archaeologists have long concluded that hunter-gatherer societies were fairly egalitarian and ostensibly leaderless – in sharp distinction to the hierarchical nature of apes. We shall not review the literature about this very early transition to egalitarianism (Boehm 1999). Neither shall we discuss the theories proposed to explain the transition of foragers to agriculture.[39] In this section we focus on surveying the extensive literature that links the transition to agriculture

[37] A vivid eye-witness description of these villages is available in Cook's account of his voyages in the eastern shores of the Pacific Ocean (1784, volume II, book IV).

[38] Chiwona-Karltun et al. (2002) provide a vivid example that illustrates our case, even if it does not involve cereals. They report that women in modern Malawi, and particularly single women, prefer to grow bitter and toxic cassava variants even though they require significantly more post-harvesting processing. This pattern is explained by the resiliency against pests that this variant offers (which applies irrespective of farmer's gender), and mostly due to the advantages of this extra post-harvest drudgery. First, it protects these women against thievery, since thieves prefer to steal the non-bitter variant that requires less processing. Second, the extra processing reduces the pressure on these women to share their cassava crop with neighbors.

[39] The conventional 'resource availability' theory attributes the timing and location of the initial transition to agriculture in the fertile crescent to climatic changes that led to evolutionary modifications in plant species (and in particular to grasses with larger seeds, to cope with the extended summer drought) which facilitated the adoption of these grasses by humans foragers (see Bar-Yosef and Meadow, 1995, and Diamond 1997). An alternative explanation for that transition contends that it was food shortage due to 'population pressure' 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 we use below against the idea that it was population pressure that led to the rise of hierarchy.

and the shift from egalitarianism toward hierarchy.

6.1 The surplus theory for explaining the emergence of hierarchy

Our understanding of the role of environmental conditions in generating social hierarchy brings us into line with recent scholars who reject theories that identify multiple stages of social organization and posit that chiefdoms are a necessary precursor stage to statehood.[40] Likewise, our distinction between crop types leads us to question the idea that horticulture (typically involving the cultivation of roots and tubers), which is often associated with a low level of social complexity, is a preliminary stage in the transition to agriculture (typically related to the cultivation of cereals) and to more complex society.[41]

These stage theories are closely related to the idea that it is agriculture that generates surplus and hierarchy, and that more extensive forms of agriculture generate more surplus and more complex hierarchy. The theory that agriculture is a prerequisite for surplus and that the availability of surplus is a prerequisite for hierarchy can be traced to Adam Smith and to earlier seventeenth century social thinkers.[42] According to Smith, government and property protection first emerged with the transition to pastoralism and the need to protect herds from theft (Smith 1978, p. 16); but only the subsequent transition to agriculture generated surplus, division of labor, production by artisans, and exchange, and thus extended significantly the role of government (1978, p. 409).[43]

For Smith, as well as for Marx and Engels, surplus had to be produced before the landlord, the capitalist or the ruler could seize it. Engels stated that in the earliest pre-agricultural stages of social evolution "Food had to be won anew day by day" and "Human labor power… yielded no noticeable surplus as yet over the cost of its maintenance" (1972, p. 65). It was the adoption of agriculture and the surplus that it generated that triggered the transition from classless society to a class society in which the usurpation of labor surplus was the essential source of class division.

[40] See Price and Feinman (2010).

[41] On the perception of horticulture as a preliminary, more primitive form of agriculture which employs extensive slash and burn, rather than intensive land use, and simpler techniques like using the digging stick or the hoe, instead of the plow, see Lensky (1966) and Johnson and Earle (2000).

[42] The surplus theories up to Adam Smith and Karl Marx are surveyed by Meek (1976) and Aspromourgos (1996).

[43] In emphasizing the role of appropriable property among pastoralists, Adam Smith can be considered to have anticipated our proposed theory. However, in the case of pastoralists he adopted a functionalist approach and emphasized that government became "necessary," rather than that it became possible. Moreover, when he turned to agriculture, he reverted to emphasize the role of surplus.

Marx's ideas were highly influential in subsequent theories of social evolution. Childe (1936), who was the first scholar to theorize about social evolution on the basis of archaeological evidence, contended that the transition to agriculture resulted in food surplus which enabled individuals to specialize in non-farming activities.[44] This surplus, and the concomitant resort to trade, helped create "political integration," and led eventually to the formation of city-states under a state bureaucracy. Agricultural surplus was thus a precondition for the emergence of the elite and artisans in urban centers.

In an influential similar theory of human development, Lenski (1966) views societies as advancing due to technological progress, from hunting-gathering to horticulture, to agriculture, to industry. Integrating functionalist and conflict approaches, he contends that the egalitarian hunter-gatherers could not produce a surplus, but progressively more advanced technologies generated surplus of goods and services. Social "power" then emerged to "determine the distribution of nearly all of the surplus possessed by a society" (p. 44).

In a more recent survey of the anthropological and archaeological literature on the emergence of inequality in the Ancient Near East, Price and Bar-Yosef (2010) reach a similar conclusion: "The success of early cultivation and the advantages afforded by the genetic mutations among plants and animals, allows for rapid increase in human population ... Cultivation also supported a stable economy with surplus that resulted in the formation of elite groups as predicted by Lenski" (1966, p. 160).[45]

Diamond's (1997) theory aligns with the conventional view. He illustrates his environmental explanation for current income disparities by comparing two groups of seafaring migrants in the Pacific whose ancestors were farmers. One group settled on an island whose environment forced them to revert to hunting-gathering. As a result: "Since as hunter-gatherers they did not produce

[44] Childe states (1936 pp. 82-83): "food production, even in its simplest form, provides an opportunity and a motive for the accumulation of a surplus. A crop must not be consumed as soon as it is reaped. The grains must be conserved and eked out so as to last till the next harvest, for a whole year. And a proportion of every crop must be set aside for seed. The conservation is easy. But it means on the one hand forethought and thrift, on the other receptacles for storage."

[45] In a review of the literature on "transegalitarian" North American societies of hunter-gatherers, Hayden (2001, p. 242) reaches a similar assessment: "With food production, in some favorable (productive) locations in the world, even greater levels of surplus production became possible. In these situations, social inequality could develop into even more extreme forms resulting in chiefdoms, states and empires." He associated the emergence of elite with "aggrandizing" activities such as feasting and states: "nearly all, if not all, aggrandizing activities involve the use of surpluses" (p. 247), leading him to conclude: "the surplus-based political models have proved to be far more insightful and rich with more interesting explanations [of complexity and inequality] than other approaches" (p. 265).

crop surpluses available for redistribution or storage, they could not support and feed nonhunting craft specialists, armies, bureaucrats, and chiefs." The other group landed on an island that was suitable for agriculture, and "With the crop surpluses that they could grow and store, they fed craft specialists, chiefs, and part-time soldiers." Diamond summarizes his theory by stating (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." He then applies this logic to attribute the economic advantage of Eurasia over Africa, America and the Pacific to a geographical factor, claiming that to Eurasia's east-west orientation enabled the exploitation of a greater variety of domesticated plants and animals and thus more productive agriculture.

Our main critique of the productivity-surplus theory for the emergence of hierarchy is based on our claim, discussed in the introduction, that surplus was neither necessary nor sufficient as a precondition for appropriation. We argue moreover, along Malthusian lines, that surplus was unlikely to emerge following the transition to farming.[46] We note that there was also a gradual increase in productivity among hunter-gatherers, due to improved hunting techniques and learning by doing. Yet, this increase in productivity was apparently translated in its entirety to an increase in population density, without leading to surplus or to hierarchy. Since the Neolithic Revolution was protracted and took place over several millennia, one could expect that also this gradual increase in productivity would have been dissipated through increased population.[47] These considerations, in conjunction with the observation about rudimentary hierarchy among farmers who rely on roots and tubers, lead us to conclude that it could not have been an increase in productivity per se that led

[46] 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 income per capita in the long run. Bellwood (2006:14-19) summarizes evidence of phenomenal population growth rates through reproduction (including one historical case of a seven-fold increase in 66 years) as a result of the introduction of mature agriculture to some frontier environments. Yet, the rate of population growth during the Neolithic Period is estimated to have been about 0.01 per cent. Comparing this rate to the potential growth rates found by Bellwood suggests that natural population increase could have easily outpaced increases in productivity so as to leave no surplus.

[47] The pace of the transition to agriculture has been debated in recent years by archaeologists and botanists. Purugganan and Fuller (2010) conclude that it must have lasted several thousand years. Childe was aware of the Malthusian argument, but apparently failed to realize its applicability by underestimating how protracted the increase in productivity was. Thus he stated (1929, p. 141): "In the long run no doubt the population would adjust itself to the means of subsistence, but the immediate result of sedentary life in a congenial environment is a surplus that must overflow." It should be noted that we do not contend that temporary random shocks to productivity or to population cannot give rise to a temporary surplus (or shortage). But foragers were most likely subject to similar shocks.

to the emergence of hierarchy.[48] Indeed, Wiesner and Tumu (1998) report that the introduction of sweet potatoes to the rugged highlands of New Guinea some 300 years ago increased land productivity and generated surplus. This surplus was rather quickly transformed into prestige goods (like pigs, slaughtered in communal festivals) and substantial population growth. But, significantly, the highland population remained fractured, characterized by endemic tribal warfare, without any consolidation of power or increase in social complexity.

6.2 Other theories about the emergence of hierarchy

We are not the first to find fault with the surplus theory for the emergence of hierarchy. Others have already pointed out that an increase in productivity may be dissipated in various ways, without leading to any surplus. Yet we find that the proposed alternatives to the surplus theory fall short of adequately accounting for the mechanism behind the correlation between the transition to agriculture and the rise of complex hierarchies.

Pearson (1957) contended that the surplus theory confines attention to food necessities and suggested that cultural needs would evolve to eliminate surplus. Sahlins (1972) insists that hunter-gatherers could procure food beyond their immediate needs but deliberately refrain from doing so. He infers that also the first farmers could have responded to increased productivity simply by working less hard, without producing any surplus.[49] In contrast to the conventional attribution of chieftainship to the production of surplus, Sahlins thus concludes (p. 140): “in the functioning of primitive society it is rather the other way round. Leadership continually generates domestic surplus. The development of rank and chieftainship becomes, pari passu, development of the productive forces.” This explanation, which claims like us that it was hierarchy that generated surplus and not vice versa, doesn’t answer the key questions: what accounts for the rise of leadership and why did its emergence correlate with agriculture?

In a survey of more than a dozen theories that purport to explain the linkage between agriculture

[48] In Appendix B we consider the case of endogenous population. We study the effect of increased productivity of an appropriable crop by comparing two variants: one with constant population; the other where population adjusts in a Malthusian fashion. To capture the idea that it is easier to tax surplus, we assume that farmers would resist taxation less, the higher their surplus is. Under the assumption that the elite choose the tax policy to maximize revenue, we show that when the population is constant, technological progress indeed increases the relative scale of the state. However, once the population adjusts endogenously to absorb any surplus, technological progress results in a proportional increase of output and tax revenue, but leaves the relative scale of the state constant.

[49] Carneiro (1970, p. 734) provides a similar argument with respect to agriculture among Amazonian Indians.

and hierarchy, Ames (2008, pp. 493-494) lists a number of explanatory factors. These include: increased population density; shift to sedentary living; storage; specialization and exchange; trade to exploit environmental heterogeneity; increased importance of property ownership; competition and warfare; aggrandizing ideology; religion, and more. These factors, though, are typically viewed as complementing, rather than supplanting, the conventional wisdom on the paramount role of productivity and surplus. Moreover, this long list of factors is almost identical to the list of correlate attributes that anthropologists advance to distinguish among pre-agricultural hunter-gatherer societies between 'simple' egalitarian societies and 'complex' hierarchical societies (see Kelly 1995, p. 294). We find that the theories that emphasize the above factors often confuse correlates of the transition to hierarchy with causal relations, and are usually vague in identifying the mechanism by which any of these factors may have contributed to increased complexity, or hierarchy.

Of the theories listed above and not dealt with thus far, the most influential is the one invoking population pressure. Anthropologists often argue that the increased productivity of agriculture led to increased population density, which led to the deterioration of living conditions through fiercer competition over resources, violence and warfare. These adverse social developments are claimed to have necessitated the reorganization of society into ever more complex social forms, leading ultimately to the formation of the central-state.[50]

Motivated by comparing the political structures that evolved in the valleys of Peru with those in the Amazon Basin, Carneiro (1970) offers a variant of the pressure and conflict argument. He contends (p. 735) that states could not emerge in the Amazon Basin, "where extensive and unbroken forests provided almost unlimited agricultural land," because "in Amazonia ... the vanquished could flee to a new locale, subsisting there about as well as they had subsisted before, and retaining their independence." In contrast, "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." According to his influential 'circumscription theory,' states arise as a result of conflict among autonomous farming villages, when the winner is able to extract ongoing surplus from the losers,

[50] Johnson and Earle (2000) are leading advocates of this theory. We readily accept that the adoption of productive agriculture and sedentary living increased population density. At issue is why this would lead to hierarchy. Rather than offering a functionalist explanation, from our point of view it may be that denser living made the early farmers more vulnerable to exploitation, by enhancing the opportunity for potential exploiters.

due to the losers' geographic entrapment, which thus permits the integration of villages into a viable state.

Carneiro's puzzlement over the limited social complexity in the Amazon Basin is analogous to Diamond's similar concern with respect to New Guinea and the Pacific Islands. Yet we note that the environmental theories that each of them offers are inconsistent with the geographical evidence that motivates the other. Diamond's theory about the advantage of an east-west orientation of a land-mass can hardly resolve Carneiro's comparison between Peru and Amazonia. And Carneiro's circumscription theory fails to resolve Diamond's concern about limited social complexity in the circumscribed Pacific tropical islands. Our appropriability theory, though, resolves the important puzzles that motivated both scholars. In particular, whereas agriculture in the tropical Amazon was based on tuber crops, farming in the western valleys of the Andes relied mostly on cereals (maize).[51]

Our theory sheds light also on another functionalist theory that sought to explain the coincidence between the emergence of early major civilizations (in Mesopotamia, Egypt, China and Peru) and riverine environments. Wittfogel (1957) argued that strong despotic hierarchies were required in these areas to realize their agricultural potential through the construction and management of large irrigation projects.[52] Mayshar, Moav, and Neeman (2014) argue, however, that the direction of causality may have been a reverse one: 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 – in analogy to our interpretation here that the need to store food facilitated confiscation.

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 increase the demand for hierarchy.[53] Fenske (2014) and Litina (2014) provide evidence for this theory. We interpret these findings as consistent with our general

[51] Counter to Diamond's insistence on the lack of mobility of plant species across the equator, maize was apparently domesticated in the northern hemisphere, before migrating to South America (see Piperno and Pearsall, 1998). The formation of Mayan state societies in the tropical lowlands of Mexico, where maize was apparently first domesticated and became the staple crop, provides additional support for our theory on the preponderant importance of cultivating cereals rather than tubers for the emergence of hierarchy.

[52] Wittfogel's many critics – see Adams (1966) with respect to Mesopotamia, and Butzer (1976:110-111) with respect to Egypt – pointed out that irrigation systems in the early civilizations were in fact constructed by local communities, prior to the emergence of a strong central state, and that even after the emergence of such central states the management of these irrigation systems remained with the local elites.

[53] Algaze (2008) proposes an analogous theory to explain the emergence of ancient Sumer.

appropriability approach, since trade also facilitates taxation.

Finally, we note recent claims that reverse the conventional causal direction between agriculture and hierarchy, positing that hierarchy preceded and led to agriculture. Challenging conventional materialistic socio-economic explanations, Cauvin (2000) argues that the willingness of hunter gatherers to abandon their traditional ways of life and engage in farming is explained by a prior change in collective psychology which he associates with the rise of religion ("the birth of the Gods").

Acemoglu and Robinson (2012, pp. 139-142) carry this idea further. Retrojecting their general perception of exogenous institutional innovations, they suggest that an institutional innovation among the semi-sedentary Natufians in the ancient Near East enabled a political elite to gain power and to extract resources from the rest of society. It is to this political elite that they attribute "the transition first to sedentary life and then to farming" (p. 140). This theory resembles ours in suggesting that hierarchy was the cause of surplus rather than its consequence. However, it is diametrically different from ours in that we take the transition to farming as given, and seek to explain hierarchical institutions by geographical factors. As noted already, this does not mean that hierarchy lagged behind agriculture, for even the earliest phases of the transition to reliance on cereals – which involved the collection and storage of natural grains, prior to cultivation and domestication – implied a fundamental shift in the ability to appropriate. It is this shift, we argue, which increased the efficacy of thievery and led to a rise of hierarchy, in parallel to the development of cultivation, domestication, increased productivity, and increased population.[54]

7 Concluding remarks

The prevailing view in the literature attributes the emergence of hierarchy to the increased productivity of agriculture. It is presumed that the increase in productivity generated food abundance (surplus), which, in turn, led to population increase, specialization in crafts, exchange, and the rise of elite. Without denying that an increase in productivity did occur, we contend that the logic behind the proposed mechanisms for the formation of hierarchy is flawed. In particular, surplus

[54] As noted in section 5, our approach, which treats political innovations as endogenous rather than exogenous, has merit in that it accounts also for the emergence of forms of hierarchy among some pre-agricultural complex foraging societies, as well as for the opposite cases in which the transition to (non-appropriable) farming was not accompanied by the emergence of complex hierarchy.

was neither necessary nor sufficient for the rise of hierarchy. Moreover, its existence is altogether inconsistent with the Malthusian theory of endogenous population, given the extremely protracted increase in productivity in the Neolithic transition. We do not challenge the prevailing perception that the transition away from egalitarianism towards hierarchy was correlated with the shift to agriculture, but we contend that the causality is more nuanced than is commonly perceived. Noting that states failed to develop in regions that farmed roots and tubers, we propose that the key feature of the Neolithic Revolution that brought about the rise of elite, and contributed to subsequent major developments in social hierarchical structures, is that cereal farmers are vulnerable to appropriation. Thus, we identify the shift to appropriable food sources as the key causal element that explains the emergence of hierarchy and accounts for differences in the scale of hierarchy across regions.

Our appropriability theory can be considered as neo-Hobbesian, in the sense of emphasizing the role of theft, banditry and expropriation. Rather than the common dichotomous view of early human societies as either hunter-gatherers or agricultural, our framework accounts for four prototypes of early societies. Within non-agricultural societies that procure food rather than produce it, our theory explains the existence of hierarchy among what anthropologists call complex hunter-gatherers. Within agricultural societies, our theory explains the substantially higher degree of social hierarchy among societies that obtain their carbohydrates mostly from cereals, as distinct from those that rely on roots and tubers. The greater vulnerability of those who rely on stored food to theft generates a demand for protection, and, simultaneously, facilitates the organized supply of protection by enabling a non-food producing elite to tax its subjects.

Our theoretical claims are illustrated with a simple model of agricultural societies that serves to motivate our empirical investigation. The main testable prediction of the model is that the key variable to explain the emergence of hierarchy is a sufficient productivity advantage of cereals over roots and tubers, and that given this advantage, absolute land productivity will have no effect on hierarchy. This implies that sufficiently high productivity of crops that are harder to tax, such as roots and tubers, in fact retards the emergence of hierarchy. Thus, whereas conventional theories suggest that it is low agricultural productivity and disease which retards the development of tropical regions, our theory and our empirical results suggest that the true hurdle that held back the development of hierarchical institutions in these regions in the pre-industrial age was the relatively high productivity of crops that provide farmers with substantial immunity against appropriation.

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

[5] Alesina, Alberto, Giuliano, Paola and Nunn Nathan (2013) "On the Origins of Gender Roles: Women and the Plough," The Quarterly Journal of Economics, 128, 155–194.

[6] Ames, Kenneth M. (2008) "The Archaeology of Rank," in Handbook of Archaeological Theories, edited by R. Alexander Bentley, Herbert D.G. Maschner and Christopher Chippindale, Altamira Press, 487-513.

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

[8] Aspromourgos, Tony (1996), On the origins of classical economics: distribution and value from William Petty to Adam Smith, Routledge.

[9] 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.

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

[11] Bellwood, Peter (2006) First Farmers: The Origins of Agricultural Societies, Blackwell Publishing.

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

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

[14] Bettinger, Robert L. (1999) "From Traveler to Processor: Regional Trajectories of Hunter-Gatherer Sedentism in the Inyo-Mono Region, California," in Settlement Pattern Studies in the Americas: Fifty Years since Virú, edited by Brian R. Billman and Gary M. Feinman, Smithsonian Institution Press, 39-55.

[15] Boehm, Christopher (1999) Hierarchy in the Forest: The Evolution of Egalitarian Behavior, Harvard University Press.

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

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

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

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

[21] 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.

[22] 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.

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

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

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

[26] Dincecco, Mark and Gabriel Katz (2014) "State Capacity and Long-Run Economic Performance," forthcoming, Economic Journal.

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

[28] Easterly, William and Ross Levine, (2003) "Tropics, Germs, and Crops: How Endowments Influence Economic Development," Journal of Monetary Economics, 50, 3–39.

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

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

[31] Galor, Oded. and Omer Ozak, (2014) "The agricultural origins of time preference," NBER WP 20438.

[32] Gennaioli, Nicola and Hans-Joachim Voth (forthcoming) "State Capacity and Military Conflict," Review of Economic Studies.

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

[34] Goring-Morris, Nigel and Anna Belfer-Cohen (2011) "Neolithization Processes in the Levant: The Outer Envelope," Current Anthropology, 52, S195-S208.

[35] Halstead, Paul, (1989), "The Economy has a normal Surplus: Economic Stability and Social Change among Early Farming Communities of Thessaly, Greece," in Bad Year Economics: Cultural Responses to Risk and Uncertainty, edited by Paul Halstead and John O'Shea, Cambridge University Press, 68-80.

[36] Hayden, Brian (2001) "Richman, Poorman, Beggarman, Chief: The Dynamics of Social Inequality," in Archaeology at the Millennium: A Sourcebook, edited by Gary M. Feinman and T. Douglas Price, Kluwer, 231-272.

[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] Kelly, Robert L. (1995) The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways, Smithsonian Institution Press.

[40] Kuijt Ian (2008) "Demography and Storage Systems During the Southern Levantine Neolithic Demographic Transition," in The Neolithic Demographic Transition and its Consequences, edited by Jean-Pierre Bocquet-Appel and Ofer Bar-Yosef, Springer, 287-313.

[41] 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.

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

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

[44] Litina, Anastasia (2014) "The Geographical Origins of Early State Formation," WP.

[45] Mayshar, Joram, Omer Moav, and Zvika Neeman, (2014) "Geography, Transparency and Institutions." Working paper.

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

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

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

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

[50] Montesquieu, Charles de Secondat, ([1748], 1989), The Spirit of the Laws, translated and edited by Anne M. Cohler et al., Cambridge University Press.

[51] Murdock, George P. (1959) Africa: its peoples and their culture history. New York, McGraw-Hill.

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

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

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

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

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

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

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

[59] 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.

[60] Piperno, Dolores R. and Deborah M. Pearsall (1998) The Origins of Agriculture in the Lowland Neotropics, Academic Press.

[61] Polanyi, Karl (1944) The Great Transformation: The Political and Economic Origins of Our Time, Beacon Press.

[62] Price, T. Douglas and Gary Feinman, (editors) (2010) Pathways to Power: New Perspectives on the Emergence of Social Inequality, Springer.

[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] 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.

[67] Rodrik, Dani, Arvind Subramanian and Francesco Trebbi (2004) "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development," Journal of Economic Growth, 9, 131-165.

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

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

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

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

[75] 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.

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

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

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

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

[80] Woodburn, James (1980) "Hunters and Gatherers Today and Reconstruction of the Past," in Soviet and Western Anthropology, edited by Ernest Gellner, Columbia University Press, 95-117.

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.

Table A.1: Staple crops in sub-Saharan Africa and in Eurasia in 1961

		World 1961		Sub-Sahara 1961		Nigeria 2013
	Energy Content (Kcal/100g)*	Average Caloric Yield (mil Kcal/ha)**	Total Energy Produced (10^{12} kcal)**	Average Caloric Yield (mil Kcal/ha)**	Total Energy Produced (10^{12} kcal)**	Total Energy Produced (10^{12} kcal)**
Rice	365	6.82	787	4.51	11	17
Maize	365	7.09	748	3.66	53	38
Wheat	327	3.56	727	2.25	6	
Barley	354	4.70	256	2.81	3	
Oats	389	5.04	193	4.52	1	
Rye	338	3.92	119	0.60	0	
Sorghum	329	2.93	135	2.46	28	22
Millet	378	2.24	97	2.17	24	19
Potatoes	77	9.41	208	5.14	1	1
Cassava	160	11.85	114	9.10	50	85
Sweet Potatoes	88	6.47	86	4.55	3	3
Yams	118	8.54	10	8.65	9	5
Total of above			3480		188	190
Population***			3083		223	174

* http://ndb.nal.usda.gov/ndb/, accessed Feb 2015. Rice: white, long-grain, regular, raw unenriched; maize: corn grain, yellow; wheat: hard red winter; Barley: hulled; oats; rye: grain; sorghum: grain; millet: raw; potatoes: flesh and skin, raw; cassava: raw; sweet potatoes: raw unprepared; Yams: raw; soybeans: green, raw; Bananas and plantain: raw. ** calculated on the basis of first column and FAO 1961 data on area and production in the world, in Africa and in northern Africa, and 2013 data for Nigeria. http://faostat3.fao.org/download/Q/*/E, accessed Feb 2015 *** http://faostat3.fao.org/download/O/*/E, accessed Feb 2015.

Table A1 presents summary data on the main staple crops in sub-Saharan Africa and in Eurasia in 1961 – the earliest year for which the Food and Agriculture Organization, FAO, provides that information.[55] Its last column presents comparable data for Nigeria in 2013. In relying on relatively recent data, our presumption is that the soil and climatic conditions have not changed significantly

[55] 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 the total energy produced by the listed twelve major crops could roughly feed the entire population.

since the Neolithic period. We recognize, of course, that the starchy 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.[56]

1. The data in Table A1 reveals that roots and tubers provided 33.5 percent of the total calories produced by the main staple crops in sub-Saharan Africa in 1961, and that cassava alone provided about 45 percent of the total calories produced by these crops in Nigeria in 2013.

2. 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).

3. The seasonality of cereals is well known. They have to be sown and reaped in a relatively fixed time 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 provides farmers with much flexibility as to 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."[57]

4. 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 rotting of these roots and tubers is often hastened by abrasions cause 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 moist root into dry flour.

[56]While varieties of yam were known in the entire tropical zone, including Asia, Africa and South America, and are believed to have been cultivated in New Guinea as early as eight millennia ago, maize, potato, sweet potato and cassava (also known as manioc, yuca or tapioca) were introduced to the Old World from America.

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

5. Finally, 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.

According to these considerations, even though the potato is biologically a tuber, for our purposes here that concern the degree of appropriability, it may as well 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.

Appendix B: Surplus and appropriation − the role of population

We develop here another 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:

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

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

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

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

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

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

Appendix C: Additional Evidence

Table C.1: Potential Crop Yields and Choice of Crops - Robustness Checks 1

	Dep. Variable: Major crop is cereal grains (dummy)				
	(1)	(2)	(3)	(4)	(5)
CALORIC DIFF	0.139***	0.268***	0.195***	0.198***	0.271***
(CER - TUB)	(0.0345)	(0.0334)	(0.0307)	(0.0315)	(0.0358)
MAX CALORIES	0.0791**	-0.103**	0.00835	0.0138	-0.0981**
(ALL CROPS)	(0.0374)	(0.0412)	(0.0336)	(0.0353)	(0.0457)
Precipitation	-0.0995***				
	(0.0238)				
Temperature Abs Latitude		0.0781***			
		(0.0183)			
Elevation			0.120***		
			(0.0154)		
Ruggedness				0.0302**	
				(0.0153)	
Abs Latitude					-0.0670***
					(0.0205)
r2	0.161	0.146	0.160	0.136	0.141
N	982	982	982	982	982

The table reports cross-sectional OLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is a dummy that identifies societies that cultivate cereal grains as main crop. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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 C.2: Potential Crop Yields and Choice of Crops - Robustness Checks 2

	Dep. Variable: Major crop is cereal grains (dummy)					
	(1)	(2)	(3)	(4)	(5)	(6)
CALORIC DIFF	0.211***	0.209***	0.256***	0.198***	0.207***	0.276***
(CER - TUB)	(0.0308)	(0.0310)	(0.0307)	(0.0313)	(0.0313)	(0.0630)
MAX CALORIES	-0.00949	-0.00947	-0.0804**	-0.0143	-0.00862	-0.235***
(ALL CROPS)	(0.0336)	(0.0338)	(0.0366)	(0.0341)	(0.0338)	(0.0758)
Major River	-0.0359**					
	(0.0144)					
Distance Coast		0.0355**				
		(0.0154)				
Pct. Malaria			0.0711***			
			(0.0152)			
Pop Dens. 1995				0.0668***		
				(0.0154)		
Hist Pop Dens					0.0324	
					(0.0323)	
Pop Dens						0.235***
						(0.0332)
r2	0.138	0.137	0.149	0.148	0.137	0.313
N	982	982	982	966	982	144

The table reports cross-sectional OLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is a dummy that identifies societies that cultivate cereal grains as main crop. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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 C.3: Cereals and Hierarchy - 2SLS. Controlling for geography.

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community				
	(1)	(2)	(3)	(4)	(5)
	2SLS	2SLS	2SLS	2SLS	2SLS
MAIN CROP: CEREALS	0.478	0.599**	0.900**	0.887**	0.590**
	(0.570)	(0.298)	(0.394)	(0.396)	(0.300)
MAX CALORIES	0.178	0.172***	0.0731	0.0725	0.167**
(ALL CROPS)	(0.120)	(0.0653)	(0.0771)	(0.0846)	(0.0693)
Precipitation	-0.112				
	(0.0744)				
Temperature		-0.0734*			
		(0.0394)			
Elevation			-0.0631		
			(0.0635)		
Ruggedness				-0.0126	
				(0.0377)	
Abs Latitude					0.0622
					(0.0402)
N	952	952	952	952	952
F excl instrum.	15.39	59.50	37.45	36.76	55.55
A-R Test (p-val)	0.403	0.0458	0.0185	0.0205	0.0502

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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 C.4: Cereals and Hierarchy - 2SLS. Controlling for isolation and population density.

	Dependent variable: Jurisdictional Hierarchy Beyond Local Community				
	(1)	(2)	(3)	(4)	(5)
	2SLS	2SLS	2SLS	2SLS	2SLS
MAIN CROP: CEREALS	0.840**	0.870**	0.777**	1.317*	0.730**
	(0.356)	(0.366)	(0.329)	(0.685)	(0.328)
MAX CALORIES (ALL CROPS)	0.0899	0.0835	0.0631	0.0250	0.0317
	(0.0695)	(0.0706)	(0.0659)	(0.103)	(0.0636)
Major River	0.102***				
	(0.0356)				
Distance to Coast		-0.0323			
		(0.0364)			
Pop Density (HYDE)			0.257**		
			(0.125)		
Pop Density (SCSS)				0.415**	
				(0.183)	
Pop Density 1995					0.334***
					(0.0481)
N	952	952	952	142	936
F excl instrum.	43.86	41.93	40.91	17.63	37.13
A-R Test (p-val)	0.0160	0.0149	0.0161	0.0243	0.0223

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas. The dependent variable is Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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 C.5: Cereals and Hierarchy - 2SLS. 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
MAIN CROP: CEREALS	0.707***	1.109***	0.845**	1.040***	0.304***	0.841***	1.080***	0.994***
	(0.0630)	(0.188)	(0.333)	(0.245)	(0.0762)	(0.236)	(0.302)	(0.257)
MAX CALORIES (ALL CROPS)			0.0692				-0.0542	
			(0.0646)				(0.0546)	
DEPENDENCE ON AGRICULTURE				0.334				-0.574
				(0.298)				(0.583)
CONTINENT FE	NO	NO	NO	NO	YES	YES	YES	YES
N	952	942	942	952	952	942	942	942
F excl instrum.		162.7	52.46	63.39		118.7	74.18	28.21
A-R Test (p-val)		0.000	0.00859	0.000		0.000	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. The dependent variable is Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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 C.6: Cereals and Hierarchy - 2SLS. 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
MAIN CROP: CEREALS	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)
MAX CALORIES (ALL CROPS)			0.0667				-0.0489	
			(0.0520)				(0.0418)	
DEPENDENCE ON AGRICULTURE				0.327				-0.513
				(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. The dependent variable is Murdock's (1967) index of jurisdictional hierarchy beyond the local community and it takes the following values: 1 (no political authority beyond community), 2 (petty chiefdoms), 3 (larger chiefdoms), 4 (states), 5 (large states). The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. 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 C.7: Cereals and Surplus - 2SLS. Controlling for geography.

	Dependent variable: Existence of a farming surplus				
	(1)	(2)	(3)	(4)	(5)
	2SLS	2SLS	2SLS	2SLS	2SLS
MAIN CROP: CEREALS	0.774**	0.764***	0.921***	0.930***	0.681**
	(0.375)	(0.261)	(0.301)	(0.315)	(0.267)
MAX CALORIES (ALL CROPS)	0.0334	0.0387	0.00222	-0.0215	0.0534
	(0.0793)	(0.0686)	(0.0677)	(0.0811)	(0.0637)
Precipitation	-0.0344				
	(0.0785)				
Temperature		-0.0281			
		(0.0475)			
Elevation			-0.155***		
			(0.0543)		
Ruggedness				-0.109	
				(0.0714)	
Abs Latitude					0.0511
					(0.0468)
N	139	139	139	139	139
F excl instrum.	10.41	19.42	15.50	14.83	15.68
A-R Test (p-val)	0.0162	0.00198	0.000	0.000875	0.00822

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas.The dependent variable is a dummy that identifies societies that produce a farming surplus. The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. "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 C.8: Cereals and Surplus - 2SLS. Controlling for isolation and population density.

	Dependent variable: Existence of a farming surplus				
	(1)	(2)	(3)	(4)	(5)
	2SLS	2SLS	2SLS	2SLS	2SLS
MAIN CROP: CEREALS	0.823***	0.851***	0.820***	0.848***	0.916***
	(0.277)	(0.275)	(0.300)	(0.288)	(0.314)
MAX CALORIES	0.0215	0.0191	0.0132	0.0208	0.0117
(ALL CROPS)	(0.0625)	(0.0626)	(0.0589)	(0.0530)	(0.0616)
Major River	0.0363				
	(0.0409)				
Distance to Coast		-0.0150			
		(0.0448)			
Pop Density (HYDE)			0.0291		
			(0.0379)		
Pop Density (SCSS)				-0.00815	
				(0.0847)	
Pop Density 1995					0.00146
					(0.0358)
N	139	139	139	139	137
F excl instrum.	15.86	17.09	13.35	17.91	12.99
A-R Test (p-val)	0.00127	0.000635	0.00353	0.000	0.00111

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas.The dependent variable is a dummy that identifies societies that produce a farming surplus. The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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 C.9: Cereals and Surplus: 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
MAIN CROP: CEREALS	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)
MAX CALORIES (ALL CROPS)			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
A-R Test (p-val)		0.0000310	0.000326	0.0000119		0.0000802	0.00548	0.0000920

The table reports cross-sectional 2SLS estimates and the unit of observation is the society in Murdock's Ethnoatlas.The dependent variable is a dummy that identifies societies that produce a farming surplus. The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. Societies that live on lands that are suitable for neither cereals nor roots and tubers are excluded from 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 C.10: Cereals and Surplus: OLS and 2SLS. 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
MAIN CROP: CEREALS	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)
MAX CALORIES			-0.0368				-0.0215	
(ALL CROPS)			(0.0501)				(0.0473)	
DEPENDENCE ON				-0.362				-0.244
AGRICULTURE				(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.The dependent variable is a dummy that identifies societies that produce a farming surplus. The main regressor is a dummy that identifies society in which the major crop is a cereal grain. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. DEPENDENCE ON AGRICULTURE is the percentage calorie dependence on agriculture for subsistence. 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 C.11: Cereals and Hierarchy - Panel Regressions

	Dep. Variable: Government above tribal level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CALORIC DIFF	0.188***	0.270***	0.280***	0.235***	0.252***	0.259***	0.192**
(CER - TUB)	(0.0683)	(0.0835)	(0.0758)	(0.0855)	(0.0890)	(0.0840)	(0.0791)
MAX CALORIES		-0.159	-0.189	-0.150	-0.110	-0.145	-0.161
(ALL CROPS)		(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 countries characterized by a supra-tribal government. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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 C.12: Cereals and Hierarchy - Panel Regressions. Robustness Checks: Excluding years 1500-1750

	Dep. Variable: Hierarchy Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CALORIC DIFF	0.198***	0.272***	0.282***	0.235***	0.249***	0.260***	0.190**
(CER - TUB)	(0.0720)	(0.0889)	(0.0811)	(0.0912)	(0.0946)	(0.0892)	(0.0846)
MAX CALORIES		-0.145	-0.176	-0.140	-0.0889	-0.130	-0.148
(ALL CROPS)		(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 dependent variable is an hierarchy index: it equals 0 if there is not a government above tribal level, 0.75 if the political organization can be at best described as a paramount chiefdom and 1 otherwise. CALORIC DIFF (CER-TUB) is the standardized difference between the maximum potential calorie yield per hectare that can be obtained from cereals versus the one that can be obtained from either roots or tubers. MAX CALORIES (ALL CROPS) is the standardized maximum potential calorie yield per hectare that can be obtained from cultivating the most productive crop among cereal grains, roots and tubers. 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 C.1: Potential yields (calories per hectare) from cereal grains.

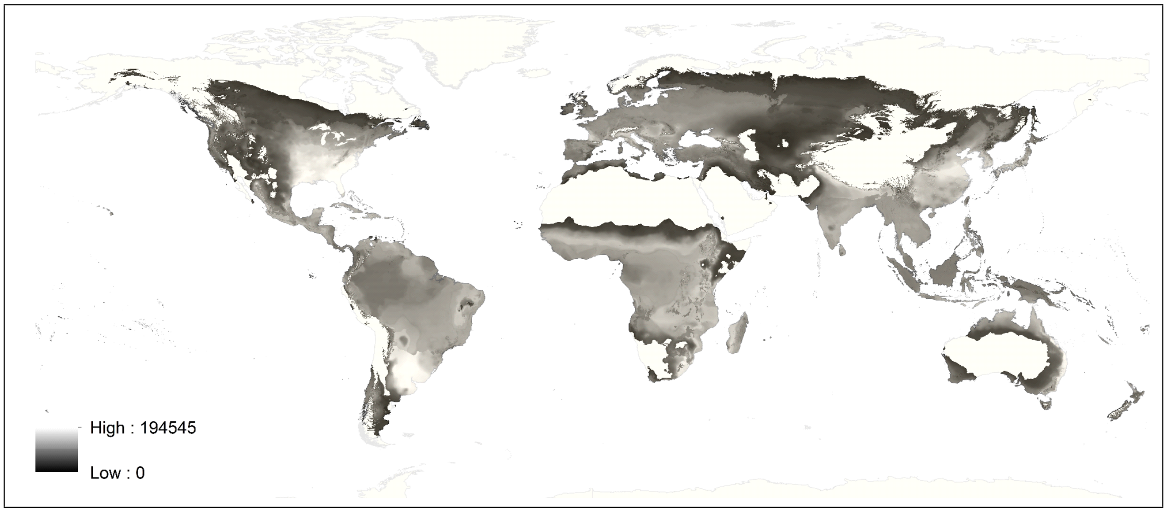


Figure C.2: Potential yields (calories per hectare) from roots and tubers

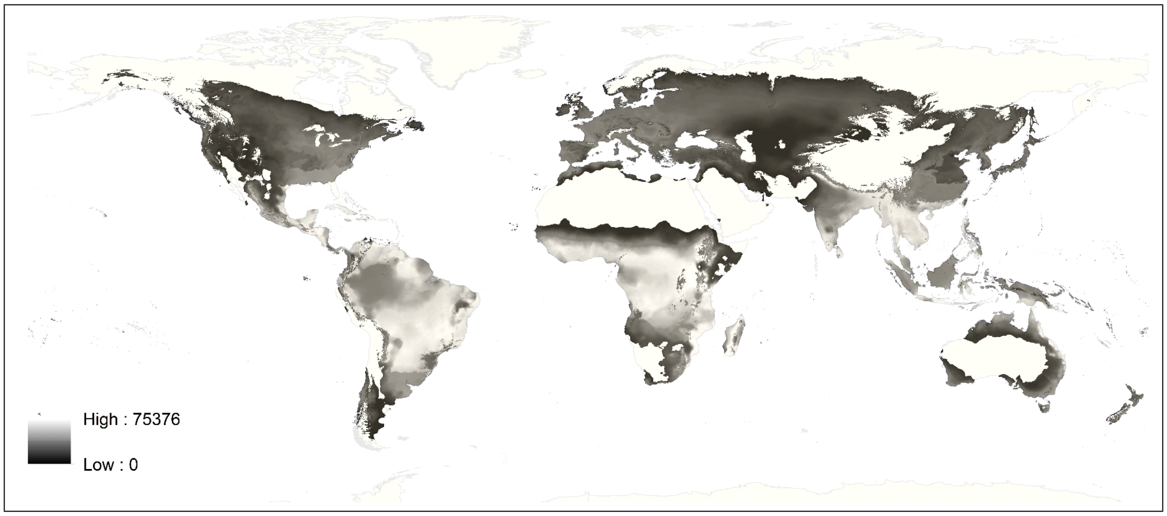


Figure C.3: Optimal crop in terms of caloric yields among cereals, roots and tubers