The Origins of Governments: From Anarchy to Hierarchy*

Matthew Baker,[1] Erwin Bulte,[2] and Jacob Weisdorf[3]

[1]: Department of Economics, Hunter College and The Graduate Center – City University of New York, 695 Park Avenue, New York, NY 10021 USA, Fax 212 772 5398, Email: matthew.baker@hunter.cuny.edu

[2]: Development Economics Group, Wageningen University, and Department of Economics, Tilburg University, Netherlands, Fax: +31 13 466 3042, E-mail: Erwin.bulte@wur.nl

[3]: Department of Economics, University of Copenhagen, 1455 Copenhagen, Denmark, Fax: +45 32 30 30 00, E-mail: jacob.weisdorf@econ.ku.dk

VERSION MAY 19, 2008

Abstract: We analyze development trajectories of early civilizations where population size and technology are endogenous, and derive conditions under which such societies optimally "switch" from anarchy to hierarchy – when it is optimal to elect and support a ruler. The ruler provides an efficient level of law and order, but creams off part of society's surplus for his own consumption. Switching to hierarchy occurs if the state of technology exceeds a threshold value, but societies may also be "trapped" at lower levels of technology – perpetuating conditions of anarchy. We present empirical evidence based on the Standard Cross Cultural Sample that support the model's main predictions.

Keywords: Origins of institutions, common defense, raiding, hunter-gatherers, SCCS.

* The authors gratefully acknowledge the feedback from seminar participants at University College London, University of Cambridge, the ETH (Zurich, Switzerland) and the First Conference on Early Economic Developments in Copenhagen 2006. Erwin Bulte's research is funded by the Dutch National Science Foundation (N.W.O.).

The Origins of Governments:

From Anarchy to Hierarchy

1. Introduction

A nice political economy literature has evolved that deals with the size distribution of nations, focusing on the gains and losses of breaking up or integrating regions (e.g. Bolton et al. 1996, Bolton and Roland 1997, Alesina and Spolaore 1997). Among the key issues that play a role are potential efficiency gains in terms of public good provision from creating larger units, versus associated heterogeneity costs (in terms of policy preferences) and distributional issues. One relevant application concerns scale economies in the production of military might, so it is not surprising that the impact of international conflict on endogenous border formation has been studied in some detail (e.g. Alesina and Spolaore 2005, 2006). One interpretation of this political economy literature is that it addresses the question why the whole world is not integrated in a single nation (Bolton et al. 1996). In this paper we consider the polar opposite question — why do we find clustering of a set of homogenous individuals with similar abilities and preferences into "hierarchical groups" (or nations) in the first place? This amounts to an enquiry into the origins and evolution of hierarchic structure in societies. Following Alesina and Spolaore we focus on conflict as a guiding theme.

According to the definitions of Hirshleifer (1995), the world was characterized by phases of 'amorphy' and 'anarchy' prior to the emergence of nation states ('hierarchy'). *Amorphy* refers to societies without storage, where resources are consumed on the move. This is clearly the relevant state of affairs for most of mankind's history – we have been mobile hunter-gatherers living from hand to mouth for millennia, and some societies still are. Such societies did not need (organized) defense to prevent others from stealing belongings. Other than the territory occupied for foraging, there was simply not much to steal. This situation changed after the agricultural transition, some 10,000 years ago for early farming societies (see Weisdorf 2005). Along with several other major changes to the human lifestyle, this involved production of surpluses and storage of commodities in a systematic way (e.g., Fernandez-Armesto 2001). Growing crops is a seasonal activity, so that reliance on storage is necessary to survive from one harvest to another. Indeed, crops maturing in the field are also, to some degree, stored assets. While possibly efficient in terms of aggregate production, storage also opens the door to theft – enabling a new economic sector (one of thieves and raiders) to emerge. *Anarchy*, then, refers to a system of spontaneous order in which agents can seize and defend resources without regulation from above. From a situation of anarchy societies may evolve towards *hierarchy*, where defense decisions are made by a central authority to incorporate the positive external effects of defense. This likely happens when the gains from such a transition for the people exceed the costs.

How important is the presence of storage in provoking a transition to hierarchy and greater social organization for the purposes of defense? Anthropologists have long recognized that capacity for production and storing resources has a profound impact on the social structure of society, a transformation that is evident even when comparing hunter-gatherers. Woodburn (1982), for example, draws a distinction between simple and complex hunter-gatherers. Complex hunter-gatherers featured some degree of hierarchy, admitted a degree of inequality among citizens and chiefs, and possessed economies that relied heavily on stored resources. They were also more likely to engage in organized warfare, own specific tracts of land, and practice slavery.[1] Kelly (1995, p. 311), in his synopsis of the differences between simple and complex hunter-gatherers puts it succinctly: "… *storage carries with it the seeds of conflict*." Anthropologists have also invoked the link between storage and hierarchy at higher levels. In his classic text, Harris (1997, p. 295) points out that taxation is only possible if there is some degree of storage capacity present in the society, and also argues that there are greater possibilities for the development and maintenance of hierarchy in the presence of resources that are storable.[2]

The transition from amorphy to anarchy and on to hierarchy, triggered by the accumulation of wealth ('lootable assets'), implies an evolutionary trajectory characterized by (i) the absence of thieving and raiding, followed by (ii) the rise of thieving and raiding and decentralized defense levels, which in turns might give way to (iii) the emergence of hierarchic structure to provide efficient levels of law and order. The corollary is that in times of economic downturns and loss of assets, hierarchic societies may regress to anarchy. Indeed, there is evidence of this, too. For example, Bell (1971) links the rise and fall of Egyptian kingdoms to variations in affluence. During years of prosperity, Egyptians were represented by a strong and wealthy kingship that upheld a stable society. But periods of economic misery (e.g. prolonged periods of drought) would cause kingships to break down.[3] Similar evidence exists for the fall of the Maya civilization and Roman Empire, and the plight of the society on Easter Island (see Diamond 2005). Indeed, even up to the time of the Industrial Revolution, shifts between economic prosperity/misery and hierarchy/anarchy existed. Between 1560 and 1770, for example, England experienced "numerous

[1] The classic example are the hunter-gatherer peoples of the Pacific Northwest of North America. Kelly (1995, p. 302) lists a few other societies that fit this mold, including some peoples of California, the Ainu in Japan, and the Calusa of Florida.

[2] As an example, Harris (1997, p. 296) describes work in Earle (1989), Hommon (1986) and Kirch (1984) on the origins and development of Hawaiian society, arguing that the lack of capacity to produce a storable grain inhibited the development of a larger, centralized state on the island chain.

[3] For example, around 2180 BCE the Egyptian Kingdom, known as Dynasty VI, collapsed. During the following so-called *Egyptian Dark Age*, which lasted 20-25 years, "hardly any form of civil disorder was absent, ranging from strife between districts, to looting and killing … to individual crime run riot, to revolution and social anarchy" (*ibid*. p. 7).

periods of political turmoil, internal warfare, and important changes in political regimes" (Clark 1996, p. 568).

This study provides a model that links technological progress to the existence of hierarchic structures upholding the rule of law. Kings or ruling elites (in what follows we use these terms interchangeably) can provide such a service and arguably represent the predecessor of nation states. Kings taxed their farmers to finance the provision of law and order, as well as their own consumption. While the political economy literature on the size distribution of nations mentioned above typically assumes a government maximizing the benefits of the median voter, we introduce a selfish ruler maximizing its own surplus. The tradeoff for the people, therefore, is not to balance efficiency gains (public goods) versus heterogeneity costs, as in most of the existing literature on jurisdictional size. Instead, they choose between being raided by thieves versus being taxed by a king. We analyze the economic incentives for the transition from anarchy to hierarchy against the backdrop of population growth and endogenous technical change in production.

There are two important prior contributions related to our work.[4] First, Usher (1989) provides a model of so-called dynastic cycles – alternating periods of peace and prosperity on the one hand, and chaos and decline on the other. A ruling elite provides socially optimal levels of defense, taking the public good nature of deterrence effort into account, bringing stability and setting the stage for populations to grow. However, diminishing returns to labor will eventually lower incomes (and taxable surplus), eventually undermining the incentive for a ruler to provide public goods. The result is a fall back into chaos and a decline in human population density, which in turn raises incomes. Eventually, sufficient taxable surplus warrants the re-entry of a new king. An important difference between Usher (1989) and our work is that the transition from hierarchy

[4] In addition to these two contributions, our paper is also somewhat related to earlier work on the origins of firms. For example, Alchian and Demsetz (1972) argue that putting in place a "central common party to a set of bilateral contracts" facilitates efficient organization of the joint inputs, which is remeniscent of our story.

to anarchy is the result of a choice by the ruler in Usher, whereas we assume the people make such choices.

Second, and closest to our work, Grossman (2002) considers the question whether producers are better off with centralized defenses organized by a ruler or with decentralized defense in a situation of anarchy. We extend Grossman's approach along three dimensions. First, we employ a conventional conflict, or contest, function to describe the interaction between producers and raiders, and hence the benefits from a switching towards hierarchy. Second, we develop a dynamic model where population size and the level of technology evolve endogenously so that we can analyze the evolution of anarchy to hierarchy. Unlike Grossman, our model suggests the existence of multiple stable equilibria, underdevelopment traps in anarchy, and so on. Third, we explore some of the model's implications using an existing dataset on lifestyle and material culture of indigenous peoples at various stages of development (the Standard Cross-Cultural Sample, or SCCS dataset). Among other things, this data includes information on the technological sophistication and governance structure of a varied group of peoples, both past and present.

The paper proceeds as follows. In section 2, as a motivating exercise, we demonstrate several empirical relationships between population density, the state of technology, and the state of hierarchy. In Section 3 we present the static model. We start off by introducing the anarchic case of a society without a ruler, and then analyze the consequences of introducing a selfish king. We compare payoffs for producers, and determine the conditions under which a king would be voted into office. In Section 4 we introduce population growth and technological change and consider the dynamics of institutional development in more detail. We interpret our specification of technical change as the gradual transition from foraging to early agriculture, and the accompanying emergence of stored (and hence 'lootable') commodities. In section 5 we present and discuss the main results. Finally, section 6 concludes.

2. A Cross-Cultural Portrait of the Evolution of Hierarchy

In this section, we describe how the emergence of hierarchy relates to some of the most basic characteristics of the economy. The evidence we present is intended to motivate our theoretical presentation and provide the reader some insights into how hierarchy apparently emerges from anarchy. Since we seek to discuss the evolution of hierarchy from a state in which there is no higher authority or government, it is useful to employ data that allows comparison of societies in which there is no form of higher government with societies in which some hierarchy exists. Our approach is to study the incidence of hierarchy in a cross-section of different cultures.

We employ the Standard Cross Cultural Sample (henceforth SCCS), an extensive and well-documented cross-cultural data set originally developed in the work of Murdock and White (1969).[5] The SCCS contains information on the technology, environment, and culture of 186 different cultures, drawn from ethnographic reports assembled at different places and times from a spectrum of indigenous cultures. As separate data points, it includes sub-Saharan African hunter gatherers, Native American hunter-gatherers, European peoples, large-scale agricultural nation-state cultures of Meso-America (such as the Aztecs), and historical nation-state peoples (such as the ancient Hebrews and Egyptians), among others. The geographical distribution of the societies in the SCCS is displayed on Map 1, which we shall discuss in more detail momentarily. The majority of the cultures in the SCCS were sampled at a time coinciding with or just after contact with western cultures (the mean date of contact among the cultures is 1850, but some observations

---

[5] For more detailed descriptions of the data and its use in economics, see Baker (2005), Baker and Miceli (2004), and Pryor (1985). The dataset has expanded to include approximately 2000 variables, and is currently maintained by William Divale, who distributes updated versions of the SCCS in the journal World Cultures. We have also indicated the original source of the data where possible.

– for example, the ancient Egyptians – date from considerably earlier). These cultures can be taken as a reasonably representative of the cultural and technological diversity of human history.[6]

< Map 1 about here >

The SCCS contains some information that broadly captures the notion of hierarchy that we shall develop in our theory. While there are several candidate variables in the SCCS that might be used to measure the degree of hierarchy, available variables vary greatly in their quality and completeness. We settle on a variable referred to as "Jurisdictional Hierarchy beyond the Local Community," which we henceforth refer to as *the degree of hierarchy*.[7] This variable runs on a scale from 1 to 5. A society earns a score of 1 if higher political authority is absent, a score of 2 if petty chiefs are present, and a 3 on the scale if it is a larger chiefdom. The scores 4 and 5 on the scale are reserved for state forms of government, where a society earns a 5 if the governance structure is that of a state with multiple jurisdictions.[8] Since our theoretical below model is basically a dichotomous model of the emergence of hierarchy, we also present some evidence using a dichotomous hierarchy variable. This hierarchy dummy variable is constructed by awarding a society a score of 0 if there is no higher political authority (i.e., a score of 1 on the degree of hierarchy scale), and a 1 for all other values.

We seek to offer evidence in support of the position that technological sophistication and, in particular, the presence of storage, are important determinants of hierarchy, and also that an important function of a hierarchy is policing. Thus, we employ a set of variables in the SCCS to measure technological sophistication, including the degree to which writing and record-keeping

---

[6] The mean contact year for societies in the SCCS is 1853. However, it must be said that there is great variance in the amount of contact the societies in the SCCS have had with the modern western world, and further that the SCCS does not include any information about the degree to which societies have had contact with other centers of development, such as the Far East.

[7] This variable was originally coded by George Murdock (1967) in the Ethnographic Atlas, which is a larger but less complete cross cultural data set.

[8] There are other variables that might proxy for the degree of hierarchy in the SCCS, but choices among these variables do not make a big difference in the analysis. For example, Tuden and Marshall's (1972) four-point scale

are present in a society, a measure of task specialization, a measure of whether or not storage and retention of surplus output is present in a particular society, and the contribution of agriculture to the food supply.[9] These variables, along with some other variables (which serve as instruments in some of our regressions), are described in Table 1. The SCCS also contains some limited information on the function of the state, and we include some information on whether or not police are present or absent.

Table 2 presents summary statistics, sorted by the presence or absence of hierarchy. A first point of interest is that the presence of hierarchy correlates strongly with the presence of police in any given society. This suggests that one of the most important functions of a hierarchy is enforcement of law and maintenance of property rights. The correlation coefficient between the police presence dummy and the presence of hierarchy is 0.68 (significant at 1%).

< Tables 1 and 2 about here >

To get a better feel for the relationship between hierarchy and technology, and in light of the fact that the technological variables plainly increase together, we have also constructed a univariate index of technology using principal components. The index is assembled from the first principal component of the three variables measuring technological specialization, agricultural importance and sophistication, and the presence of writing and record keeping. The first principal component captures a strongly positive association between the three variables and explains about 64 percent of the joint variation in the three variables.[10] We take this to mean that in one common dimension the overall level of technology in the society is increasing. We refer to the index so

---

rating the levels of sovereignty present in each society is also included in the SCCS. Our measure is heavily correlated with this variable (correlation coefficient of .77, significant at 1% level).

[9] The specialization scale variable, the reliance on agriculture scale variable, and the writing and record keeping variable were all coded originally by Murdock and White (1969).

[10] To be specific, the technology index is computed using the first principal component as follows: *technology*=.6217 *specialization* + .5678 *agriculture* + .5396 *writing*. This component explains 64% of the joint variation in these variables. As one would expect, all the coefficients are positive, meaning that the four indices all vary strongly in one direction. Somewhat surprisingly, the three scale variables enter into the first principal component with very similar

constructed from this first principal component as "the level of technology," or more succinctly, "technology." The resulting technology index has a mean of zero (by design), a minimum value of -2.34, and a maximum value of 2.40. A society for which the technology value is close to the lower limit can be thought of as one with a low reliance on agriculture, little or no record keeping and writing, and little task specialization. We maintain our fourth technological variable, measuring the presence or absence of storage, separately as we hope to offer some evidence that this variable is an important determinant of hierarchy independent of the behavior of the other measures of technology.

Table 3 breaks down each technological measure, the aggregate measure of technology, the scale variable measuring the importance of storage, population density, and whether or not police are present in the society by the degree of hierarchy. The results suggest that all of these things increase more or less together.

But what roles do technology, population density, and storage play in the emergence of hierarchy? It seems reasonable to posit that hierarchy, population density, and technology are all jointly determined. As we shall argue in the theoretical section of the paper, the adoption of hierarchy has consequences for the types of technology that may be used and the population density that may be sustained by a given society. Thus, a more rigorous theoretical analysis must rely on some sort of instrumentation strategy. Along these lines, consider once again Map 1, where societies that are 'hierarchical' according to our dichotomous variable are marked with squares or rectangles, while non-hierarchical societies are marked with circles. It is fairly clear from the geographical distribution that those societies that are hierarchical appear to be grouped closer to historical centers of civilization, and also appear to be centered in areas in which the environment might be considered richer.

weights, suggesting that a simple sum of these variables might also function quite well as an index of technological specialization.

Measures of environmental quality like those described in Table 1 should have little direct impact on the emergence and development of hierarchy. Environmental factors should work through other variables by allowing higher population density, which in turn impacts the adoption and creation of more sophisticated technologies. Thus, it seems reasonable to use as instruments for population density and technology underlying environmental characteristics. These environmental variables are described on table 1. One cannot be sure, however, from Map 1 that the clustering of hierarchical societies is due solely to similar population density and environment. To further control for this closeness and to allow for the possibility that hierarchical forms of government are learned and adopted from neighbors, we include the distance measures described in table 1 as further exogenous variables in regressions.[11]

Table 4 presents a series of regressions in which hierarchy is the dependent variable, and various measures of population density and technological sophistication are used as independent variables. Since technology and population density may be endogenous, and because the analysis may suffer from omitted variables and measurement error, we report both OLS results and IV results, including probit specifications in which the dependent variable is absence or presence of hierarchy. The first two columns on table 4 are the results of OLS estimation. The second column reports OLS results when the three separate technology measures are replaced by the first principal component of the three. Not surprisingly, the results are similar and suggest that population density, technology, and hierarchy are significantly positively related. The third column of table 4 presents instrumental variables results where we include distance controls in the main regression and use environmental characteristics as instruments. Hausman endogeneity tests for the IV regression reveal that one may reject the Null that population density, technology, and storage are

[11] To be completely clear, we shall use environmental variables as instruments, while we shall use distance measures as additional controls in the regressions. The argument is that hierarchical behavior may be learned or may spill over from neighboring societies, while environmental characteristics do not determine the degree of hierarchy directly, but only through their impact on technological sophisticaton and population density.

jointly exogenous. Column 3 also includes partial F-statistics for the first-stage regressions, and these suggest that our instruments perform reasonably well. Moreover, the J-statistic for the regression suggests that the instruments are orthogonal to the model error term.

Since our dependent variable is close to being dichotomous, we also include some probit model estimates on table 4, columns 4 and 5. Column 4 presents a probit model mirroring the OLS specification in column 1; the results are similar to previous results. Column 5 estimates the probit model, including the distance controls and treating the technology and population density variables as endogenous. The results are again similar to the IV estimation in column 3.

This evidence is merely meant to be suggestive of a few fundamental features of some of the root determinants of the degree of hierarchy in a society, and is not intended to be a rigorous test. Nevertheless, it is worth summarizing what we consider to be the crucial points that one may take from this data. First, there appears to be a relationship between the presence of hierarchy and policing. Second, hierarchy appears in part to be caused by greater degrees of technological sophistication and the presence of storage surpluses. Third, when addressing the endogeneity through instrumentation, it does not appear that population density is a crucial factor in driving the evolution of hireranchies. In the subsequent sections, we lay out a model that describes how this process might work, and that is consistent with these 'stylized facts'.

3. The Static Model

3.1 The case of anarchy

Consider a society or group of n members, some of which earn a living as a producer (or early farmer) and others are supporting themselves by raiding producers, taking part of their output. People are risk neutral and amoral, preferring to become thieves whenever that profession is more profitable than producing. The total population is described as follows:

(1) $n = n_f + n_r$,

where n_f denotes the number of producers, and n_r denotes the number of raiders in society. Each raider devotes his entire endowment of time (one unit) to raiding. In contrast, producers split their time between production and defense effort to protect their output from raiders. The time constraint for producers is given by $l + d = 1$ where l is labor devoted to production and d is defense effort by the producer.

There are two inputs in production, land and labor. The production function is given by $f(q,l) = Aq^{\beta} l^{\mu}$, where β, μ ≤ 1. Land is available in fixed supply, and we normalize the total land base to unity. After career choices, each producer is allotted an equal share of available land upon which to produce, so each producer receives ($1/n_f$) units of land. To simplify the analysis and exposition, and without affecting the qualitative results, we set μ=1 in what follows so that the production function for any producer, net of any theft of output, is given by:

(2) $$\pi_f = A n_f^{-\beta} l s ,$$

In equation (2), A is a parameter measuring the state of technology (which will be made endogenous below) and s is the share of output retained by the producer. We specify the contest or share function as follows:[12]

(3) $$s = \frac{d}{d + \theta \frac{n_r}{n_f}} ,$$

where the parameter θ is a proxy for (lack of) security and measures the ease with which output can be stolen. This parameter must take on values less than one-half, as will become evident below. This share function is consistent with the idea that all raiders devote all of their efforts to

[12] Note there is a subtle difference in Grossman's method and the one used here. Grossman assumes the share of resources retained depends upon the fraction of effort devoted to defense, so s = 1 when d = 1, regardless of the proportion of raiders to defenders. Our specification is arguably more natural, and is certainly more commonly used in

theft, and each farmer bears an equal fraction of total raiding effort. Using the time constraint and the share function, producers' payoffs are defined as:

$$\pi_f = \frac{A(1-d)}{n_f^{\beta}} \frac{d}{d+\theta\psi}, \tag{4}$$

where $\Psi = n_r/n_f$, or the ratio of raiders to producers in the economy. Under autarky, each producer chooses defense efforts to maximize (4) while taking the number of raiders in the population as given. The optimal choice of defense is simply:

$$d_A = \sqrt{(\theta\psi)^2 + \theta\psi} - \theta\psi\ . \tag{5}$$

Plugging (5) into (4) gives the following expression for the optimal return to producing and defending for any combination of raiding and producing, n_r and n_f, in the economy:

$$\pi_f = A\frac{\left(1+\theta\psi - \sqrt{(\theta\psi)^2+\theta\psi}\right)\left(\sqrt{(\theta\psi)^2+\theta\psi} - \theta\psi\right)}{\sqrt{(\theta\psi)^2+\theta\psi}}. \tag{6}$$

Next, consider the payoffs for raiders. Assuming each raider gets an equal share of the total 'take', these payoffs consist of the total resources stolen from producers divided by the number of raiders:

$$\pi_r = \frac{A(1-d)}{n_f^{\beta}}\left(1-\frac{d}{d+\theta\psi}\right)\frac{1}{\psi}. \tag{7}$$

Using the equilibrium value of defense in (5), we have:

$$\pi_r = A\theta\frac{1+\theta\psi - \sqrt{(\theta\psi)^2+\theta\psi}}{n_f^{\beta}\sqrt{(\theta\psi)^2+\theta\psi}}. \tag{8}$$

In equilibrium, people are indifferent between producing and raiding so that the returns from each career choice should be the same. Equating (6) and (8) results in the following expression for the ratio of raiders to producers in the economy (or the raid ratio):

---

the literature. A more general specification would raise contest effort of raiders and farmers to a certain power R. Baye et al. (1994) demonstrate that a symmetric Nash equilibrium then exists, even for R>2.

(9) $$\frac{n_r}{n_f} \equiv \psi = \frac{\theta}{1-2\theta}.$$

Hence, the fraction of raiders in the population only depends on security parameter θ. Amorphy, as defined by Hirschleifer, thus occurs in the special case where θ=0. As θ increases, so does the relative number of raiders. When θ=½, all agents prefer to be raiders and per capita income in the economy has completely evaporated. In equilibrium, we have the following number of producers:

(10) $$n_f = n\frac{(1-2\theta)}{1-\theta}.$$

The returns to productive labor are given by:

(11) $$\pi_f = \pi_r = An_f^{-\beta}(1-2\theta).$$

Upon substituting (10) in (11) we get an expression for equilibrium income of raiders and farmers:

(12) $$\pi_f^A = \pi_r^A = An^{-\beta}\left(\frac{1-2\theta}{1-\theta}\right)^{-\beta}(1-2\theta) = An^{-\beta}(1-2\theta)^{1-\beta}(1-\theta)^{\beta}.$$

Note that the returns in (12) depend upon the degree of security present in society, and in particular go to zero as θ goes to ½. We now investigate how these outcomes are affected when a potential king enters the scene.

3.2. The case of hierarchy

Following Grossman (2002), under hierarchy, the ruler organizes defense on behalf of everybody in society. While individual people still use their own effort to protect their output, the ruler decides on the allocation of time between production and defense.[13] The key difference with anarchy is that the king takes into account that the number of raiders is endogenous with respect to the level of defense. In other words, the ruler internalizes the external effect that raising defense induces some raiders to become producers instead, as opposed to the decision of individual agents

[13] In what follows we ignore problems of free riding and cheating, and simply assume that the ruler is able to force "his people" to the optimal allocation of time. For an early discussion of this simplification, refer to Williamson (1973).

that just balances foregone production versus the share of output that can be retained. However, the king also reserves the right to collect taxes, and the cost of hierarchy is that the hierarch is concerned about his own welfare when setting taxes. The payoffs to being a producer are now defined as:

(13) $$\pi_f = \frac{(1-\tau)A(1-d_H)}{n_f^{\beta}} \frac{d_H}{d_H + \theta\psi},$$

where d_H denotes the centralized level of defense and τ is the tax rate chosen by the hierarch. Assuming only legitimate activities are subject to taxation, the returns to being a raider are now given by:

(14) $$\pi_r = \frac{A(1-d_H)}{n_f^{\beta}} \left(1 - \frac{d_H}{d_H + \theta\psi}\right)\psi^{-1}.$$

The hierarch must set the level of defense and tax rate so that the following condition is satisfied:

(15) $$\frac{(1-\tau)A(1-d_H)}{n_f^{\beta}} \frac{d_H}{d_H + \theta\psi} \geq \frac{A(1-d_H)}{n_f^{\beta}} \left(1 - \frac{d_H}{d_H + \theta\psi}\right)\psi^{-1},$$

otherwise, he collects no tax revenues. Equation (15) reduces to the following condition:

(16) $$(1-\tau)d_H \geq \theta.$$

The total tax revenues collected by the hierarch are then defined by $y_H = n(1-d_H)A\tau / n_f^{\beta}$. The hierarch chooses defense to maximize tax revenues subject to a "participation constraint" (16), or the condition that people prefer to produce rather than to raid. Note this choice involves an elementary tradeoff. While the government wishes to raise the tax rate, he is restricted in his ability to do so by the fact that raising taxes implies he should simultaneously mandate higher levels of defense (which comes at the expense of foregone production). The reason is that raising the tax rate makes a "switch" to raiding more attractive, *ceteris paribus*, which must be offset by an increase in defenses. The solution to the hierarch's problem is:

(17) $$d_H = \sqrt{\theta}\,.$$

Not surprisingly, centralized defense levels are higher than the anarchic level of defenses, reflecting that the king considers the impact of defense decisions on career choice while individual producers do not.[14] Plugging in the optimal value of τ and d_H into (11) gives the following expression for producers' income under a ruler:

(18) $$\pi_f^H = (1-\sqrt{\theta})\sqrt{\theta}An^{-\beta}\,.$$

We can now solve for the critical security level where people are indifferent between being raided and taxed by equating (12) and (18). The critical parameter is defined by the following condition:

(19) $$(1-\sqrt{\theta})\sqrt{\theta} > (1-2\theta)^{1-\beta}(1-\theta)^{\beta}\,.$$

Since the right-hand-side, RHS, of (21) monotonically decreases in the magnitude of θ for any value less than ½ and the left-hand-side increases and then decreases, there exists a unique intersection point. Denote this threshold value as θ*. For sufficiently large values of θ we find that $\pi_f^H > \pi_f^A$.

While this model is suited for discussing the tradeoff between autarky and an appropriative government, it says nothing about how the degree of government is influenced by population and technological progress, or about the different possible phases in the development process of societies over time. A dynamic model is necessary to consider these issues.

4. A Dynamic Model

We now introduce equations of motion to capture the intertemporal development of population size and the state of technology. Following Galor (2005) and others, we assume that technical

[14] Note that (15) defines a new constraint on parameter θ: since the time constraint implies $d \le 1$, interior solutions can only occur for $\theta < 1$ as well. Since this "hierarchy constraint" is less stringent than the anarchy constraint, $\theta < ½$, it will be satisfied automatically.

progress is one of the main drivers of the model, and aim to capture the process of innovation and depreciation (erosion) of the state of technology as follows:

(20) $$\dot{A} = \upsilon n_f^{\alpha} A^{\gamma} - \delta A,$$

where $\dot{A} = dA/dt$. The first term on the RHS captures innovations and the second term captures erosion of technology (δ denotes a depreciation parameter). The specification in (20) assumes the state of technology is increasing in the current state of technology, A, and in the population of producers which may be due to some 'innovation by doing' argument, or because higher population density allows for a faster exchange of ideas, etc. (e.g. Shekhar et al., 2006). Although this is not important for the model that follows, we envisage the gradual increase in A over time as the gradual transition from foraging to early farming.

One implication of $\dot{A} > 0$ is that output becomes easier to steal – increases in output (surpluses) not only imply that guarding output becomes harder, theft may also be facilitated due to changes in the nature of production. For example, if producers rely more on harvesting produce that they have planted or sown themselves, then raiders may have relatively easy access to the stock when it is on the field. This means it is realistic to assume that the security parameter θ is not invariant with respect to technology levels. We capture this idea by assuming a linear relation between the state of technology and the ease with which output can be stolen (the results are robust with respect to alternative monotonic relationships):

(21) $$\theta \equiv zA,$$

where z is just a scaling parameter. In line with the earlier discussions on threshold values for the security parameter θ, specification (21) defines an upper limit on productivity for the anarchy and hierarchy models to be sensible – respectively $A < \overline{A} = 1/2z$ and $A < \widetilde{A} = 1/z$. Failing to satisfy these conditions implies all become raiders (for the anarchy model) or producers allocate all their time to defenses and there is no production (for the hierarchy model).

Recall people are indifferent between anarchy and hierarchy when $\theta = \theta^*$ (or $A = A^*$). For $\theta > \theta^*$ (or $A > A^*$) the dynamics of the state of technology are determined in the context of a society governed by a ruler. Instead, for $\theta < \theta^*$ (or $A < A^*$), anarchy prevails and some people prefer to raid rather than produce. This means we must derive two different segments of the $\dot{A} = 0$ isocline associated with (20), or two distinct segments where technology is constant. One segment is relevant for the anarchy context (where some people are raiding, $\Psi>0$, and not contributing to the accumulation of knowledge) and another one is relevant for the hierarchy context ($\Psi=0$). Solving for $\dot{A} = 0$ under anarchy and hierarchy results in, respectively:

(22a) $$n = \frac{1 - zA}{1 - 2zA}\left(\frac{\delta}{\upsilon}A^{1-\gamma}\right)^{1/\alpha}, \text{ and}$$

(22b) $$n = \left(\frac{\delta}{\upsilon}A^{1-\gamma}\right)^{1/\alpha}.$$

Equations (22a) and (22b) imply that, at $A=A^*$, there is a discontinuity in the $\dot{A} = 0$ isocline, which jumps downwards then (note $0 < (1\text{-}zA)/(1\text{-}2zA) < 1$).

Specification of a population growth process completes the system's dynamics. Following a large literature on early economic civilizations (e.g. Brander and Taylor 1998, Horan et al. 2005) we describe population dynamics by a simple Malthusian process:

(23) $$\frac{dn}{dt} = \varphi n\left(-1 + \frac{\pi^i}{S}\right),$$

where φ is a population growth parameter, i denotes anarchy or hierarchy, and S is a food intake threshold (minimum caloric intake for self-maintenance). For $\pi < S$, the population shrinks and for $\pi > S$ it grows. We ignore the king's income as his fertility decisions are relatively unimportant relative to those of the population at large, and use the people's income as provided in (12) and (18) (but qualitatively similar outcomes obtain if we would base the analysis on total income in the economy). Condition (23) is readily solved for the $\dot{n} = 0$ isocline by setting $\pi = S$ and, again, there are 2 distinct isocline segments depending on whether A is greater or smaller than A^*. The isocline segments for anarchy and hierarchy, respectively, are given by:

(24a) $$n = \left(\frac{A(1-2zA)^{1-\beta}(1-zA)^{\beta}}{S} \right)^{1/\beta}, \text{ and}$$

(24b) $$n = \left(\frac{A(1-\sqrt{zA})\sqrt{zA}}{S} \right)^{1/\beta}.$$

With these building blocks in place, we can now turn to an analysis of the dynamics of population growth and technical change, and the consequences for institutional change.

5. Model results

Before we analyze the dynamics and steady states in some detail, it is instructive to consider the impact of technical change on some of the model's key variables in equilibrium. In Figure 1 we demonstrate that the anarchic raid ratio (ψ), or the number of raiders divided by their victims, increases as the state of technology increases. While an increase in productivity increases the returns to both producing and raiding (so that the net effect is neutral)[15], there is an additional effect due to the assumption that the production and storage of surpluses enhances the ease with which output can be stolen – tipping the balance in favor of becoming a raider. The net effect of technical change on income of producers and raiders is ambiguous, depending on the level of technical change (see Figure 3 below). Technical change raises the value marginal product of labor allocated to production, but since fewer people actually produce the net effect on income is unclear. Specifically, if we consider the case where society is in anarchy before and after the

[15] Note that the argument A does not appear in (9) – the raid ratio is only a function of security parameter θ.

increase in A (so that equation (12) describes income throughout), we can use (21) to show that technical change is detrimental for income if:

$$(25) \quad \frac{d\pi}{dA} = 1 - 4zA < 0,$$

which is the case whenever $A > 1/4z$, or when the raid ratio curve is sufficiently steep (see Figures 1 and 3).

< Figure 1 about here >

Next, consider the case of hierarchy. In Figure 2 we show that an increase in the state of technology lowers the equilibrium tax rate. Since technical progress makes raiding more attractive, the ruler aims to neutralize the incentive to switch from producing to raiding, and lowering the tax rate so that producers can retain a larger share of their output is one approach to doing this. An alternative approach to restoring the balance between producing and raiding would be to raise the levels of defense, but this comes at the cost of foregone production. To find the optimal tax rate, the ruler balances the foregone profit share from low taxes versus the foregone output from raising defense levels. Hence, while technical advances ambiguously impact on producers' income under conditions of anarchy, when ruled by a king it makes producers unambiguously better off – it allows them to retain a greater share of a larger pie. One may use (18) to verify that $d\pi^H / dA > 0$.

< Figure 2 about here >

In Figure 3 we provide equilibrium income for anarchic and hierarchical societies. As explained above, the relation between the state of technology and income under anarchy is non-monotonous. The critical level of technology A^*, and the associated critical security parameter θ^*, is defined by the unique interior intersection of the two income curves.

< Figure 3 about here >

5.1 Analysis of the dynamic system

The analysis until now implies that simple models with exogenous technical change (i.e., $\dot{A} = A_0 e^{\sigma t}$ where σ is a parameter) will necessarily undergo the transition from anarchy to hierarchy: eventually the level of technology will be sufficiently high to elect a king in power. Is this also true for our dynamic model with endogenous technical change? The answer is no. For different parameter combinations we obtain qualitatively different results. In what follows we illustrate this important point by presenting two interesting cases.

In Figure 4 we provide the phase plane for one possible configuration of parameters, providing the isoclines as derived in (22) and (24).[16] On the horizontal axis the state of technology is depicted, and the vertical axis provides population numbers (or density). As before, A^* denotes the critical technology level, so that for $A < A^*$ the anarchy segments (22a, 24a) are relevant, and for $A > A^*$ the hierarchy segments (22b, 24b). The present configuration has two potentially stable interior steady states: in both the anarchy and hierarchy sectors of the phase plane the $\dot{A} = 0$ isocline cuts the $\dot{n} = 0$ isocline from below, which implies that each steady state is a focus or node. The stability of the steady states depends on the relative slopes of the isoclines, and may be analyzed further by examining the eigenvalues of the linearized system in equilibrium.

< Figure 4 about here >

The existence of a potentially stable steady state in the anarchy sector implies that societies may "get stuck" to the left of A^* and never elect a king. These low-tech societies prefer to go untaxed and provide defense on a decentralized basis. We interpret these societies with low productivity levels and little or no storage of commodities as stable foragers' societies. Other societies, however, will cross the security (or productivity) threshold – possibly because the trajectories cycle outward. Depending on the trajectory, such societies will eventually support a king providing efficient levels of security, or cycle back towards the anarchic segment. Indeed, as in Usher (1989) it may be feasible to have 'dynastic cycles' where societies grow and collapse and undergo periodic phases of anarchy.[17] The model predicts farmers' societies where storable and 'lootable' surpluses are produced will settle down in the hierarchic steady state. This is a prediction that is amenable to empirical testing in section 5.[18]

The phase plane suggests that societies are more likely to become hierarchical if the initial level of technology is 'high' (i.e. close to the threshold) or if the population is 'large.'[19] In contrast, societies with few people are likely to evolve towards anarchy, even if their initial level of technology places them to the right of the threshold. The reason is that such societies will gradually lose their level of technical sophistication as new innovations fail to keep pace with the erosion of knowledge (similar to the findings in Shekhar et al., 2006). As an interesting aside, while population size is important during the transition phase, the model predicts that *equilibrium* population size of the anarchy and hierarchy steady state cannot be unambiguously ranked. Indeed, for the combination of parameters underlying Figure 4 we find that the hierarchical society is smaller than the anarchic one. Again, this is a model implication that may be amenable to empirical testing.

In Figure 5 we present an alternative phase plane based on a different set of parameters.[20] This is the case where the $\dot{A}=0$ isocline cuts the $\dot{n}=0$ isocline from *below* in the anarchy sector (as before), and from *above* in the hierarchy sector of the phase plane. The anarchy steady state is again a focus or node, but the hierarchic steady state is now a saddle point. The separatrix

---

[16] To be precise, Figure 4 is based on z=1, α=0.3, β=0.3, δ=0.9, γ=0.2, ν=1.0 and S=0.205.

[17] But unlike the model of Usher, the dynamics are not governed by diminishing returns in production and a king that voluntarily steps down – it is driven by cycles of population abundance that affect productivity levels and the ease with which output can be stolen (affecting the incentives of the people to support a king, or not).

[18] An additional effect, not captured by the model, is that agriculture and storage facilitate financing a hierarch as it makes it much easier to collect taxes: agricultural output is readily observable, and storage enables accumulatation of resources to pay for taxes.

[19] The intuition for the latter is evident: if societies start sufficiently "high" in the phase plane, they will necessarily cross the $A = A^*$ threshold. The reason is that technical change occurs at a rapid rate in societies with many people. Even if the population "shrinks", the level is high enough to induce a transition from foraging to early agriculture.

[20] Figure 5 is based on z=1, α=0.135, β=0.3, δ=0.9, γ=0.9, ν=1.0 and S=0.15.

associated with the saddle point divides the hierarchy sector in two parts – below the separatrix the system is pushed (back) towards anarchy, above if the system is pushed on a path of sustained growth where population increases and technical mutually increase each other. In other words, an anarchic society that faces a sufficiently large (temporary) productivity shock – say, a climatic shock or significant technical innovation (possibly through technology import) – will face a fundamentally different future. This will happen if due to the shock the system "jumps" above the separatrix. Not only will it elect a king in power, it will also experience sustained growth where population growth spurs technical change, and where the income increase following technical change permits Malthusian style population growth.

< Figure 5 about here >

This is not intended as an exhaustive analysis of all potential dynamic outcomes. For example, some parameter values imply the anarchic steady state disappears (i.e., the isoclines do not intersect for values of A below A^*) and others imply it changes into a saddle point. The latter possibility suggests that for some initial conditions (specifically: for starting points below the anarchy separatrix) society will become trapped in a vicious circle of population decline and technical regression, and eventually go extinct. For other initial conditions, of course, the transition to hierarchy is inevitable. Detailed knowledge about underlying parameters of societies is necessary to predict the qualitative nature of its long-term fate.

6. Discussion and Conclusions

Why do some societies evolve from anarchy to hierarchy, and why do other societies fail to make that transition? We propose a novel explanation by focusing on internal conflict (raiding) and the tradeoffs between decentralized and centralized levels of defense. While centralized levels of defense are efficient – taking external effects into account – and increase total output relative to anarchic equilibrium levels, there is an associated cost for producers as the ruling elite is capable to tax away part of the surplus for its own consumption. The main result of the paper is that producers prefer to be taxed, rather than be raided, if production technology is sufficiently advanced. Using a dynamic model with endogenous population and technology we demonstrate that the evolution from anarchy to hierarchy is possible, but not inevitable – some (foraging) societies will prefer to stay egalitarian and anarchic.

Our empirical analysis, based on the Standard Cross Cultural Sample, supports the model's main predictions. Using OLS and IV estimation techniques we find that technology levels are an important driver of the degree of hierarchy in indigenous cultures, and that 'storage' plays a particularly important role in this respect. The latter result makes perfect sense in the context of our model because, in line with several anthropologists, we argue that the ability to store output is one of the main determinants of raiding and conflict (enhancing the appeal of efficient levels of defense). Consistent with the theory model, our empirical work also downplays the role of population size as a determining factor in the emergence of hierarchic structures. The potential efficiency gains from centralized public good provision are therefore not necessarily an overriding concern for community members when deciding about whether to elect a king, or not.

The theoretical model is very stylized and ignores many salient features of reality. For example, we don't consider that some communities may be able to coordinate on optimal levels of defenses in the absence of a ruler, or that security parameter θ may be subject to technical change (related to technical advances in production, or otherwise). We also ignore the issue of external conflict – including fights over territories of different communities – which may clearly provide an important impetus for centralized organization of defense (see Baker 2003). Finally, by assuming all community members have identical skills we assume away that there may be internal conflict over whether to switch from anarchy to hierarchy, or not. Members that are more productive than others, or have access to more productive land, will prefer to make the transition at some lower critical value of technology. This could result in societies "breaking up" or in some faction imposing its will on another. Analyzing these issues, and others, is left for future work.

Literature cited

Alchian, A.E. and H. Demsetz, (1972), "Production, Information Costs, and Economic Organization," *American Economic Review* 62, pp. 777-795.

Alesina, A. and E. Spolaore, (1997). "On the Number and Size of Nations," *Quarterly Journal of Economics* 112, pp. 1027-1056

Alesina, A. and E. Spolaore, (2005). "War, Peace, and the Size of Nations," *Journal of Public Economics* 89, pp. 1333-1354.

Alesina, A. and E. Spolaore, (2006). "Conflict, Defense Spending and the Number of Nations," *European Economic Review*, In Press.

Baker, M. J., (2003), "An equilibrium model of land tenure in hunter-gatherer societies," *Journal of Political Economy* 111, pp 124-173.

Baker, M. J., and Thomas Miceli (2005), "Land Inheritance Rules: Theory and Cross-Cultural Analysis," *Journal of Economic Behavior and Organization* 56(1): pp. 77-102.

Baker, M. J., (2005), "A Structural Model of the Neolithic Revolution," United States Naval Academy Department of Economics Working Paper 2005-9.

Baye, M.R., D. Kovenock and C.G. de Vries, (1994), "The Solution to the Tullock Rent-Seeking Game when R > 2: Mixed-Strategy Equilibria and Mean Dissipation Rates," *Public Choice* 81, pp. 363-380.

Bell, B. (1971), "The Dark Ages in Ancient History," *American Journal of Archaeology* 75(1): pp. 1-26.

Bolton, P. and G. Roland, (1997). "The Break-up of Nations: A Political Economy Analysis," *Quarterly Journal of Economics* 112, pp. 1057-1090.

Bolton, P., G. Roland and E. Spolaore (1996), "Economic Theories of the Breakup and Integration of Nations," *European Economic Review* 40, pp. 697-706.

Brander, A. and M. Taylor (1998), “The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use,” *American Economic Review* 88, pp. 119-38.

Cashdan, Elizabeth (2001), “Ethnic Diversity and Its Environmental Determinants: Effects of Climate, Pathogens, and Habitat Diversity,” *American Anthropologist* 103: pp. 968-991.

Diamond, J. (1997), *Guns, Germs, and Steel: The Fates of Human Societies*, New York: WW Norton.

Fernandez-Armesto, F. (2000), *Civilizations*, London: Macmillan.

Galor O. (2005), “From Stagnation to Growth: Unified Growth Theory,” *Handbook of Economic Growth,* 2005, 171-293.

Gray, J. Patrick, ed. (2006), *World Cultures: Journal of Comparative and Cross Cultural Research. Journal* Jamaica, NY: Social Sciences Department, York College – City University of New York.

Grossman, H.I. (2002), ”’Make us a king’: anarchy, predation, and the state,” *European Journal of Political Economics* 18, pp. 31-46.

Grossman, H.I. and S.J. Noh (1994), “Proprietary public finance and economic welfare,” *Journal of Public Economics* 53, pp. 187-204.

Hirschleifer, J., (1995), “Anarchy and its breakdown,” *Journal of Political Economy* 103, pp. 26-52

Horan, R.D, E.H. Bulte and J.F. Shogren (2005), “How trade saved humanity from biological exclusion : an economic theory of Neanderthal extinction,” *Journal of Economic Behavior and Organization* 59, pp. 297-323

Murdock, George P., and Caterina Provost (1973), ”Measurement of Cultural Complexity,” *Ethnology* 12: pp. 379-392.

Murdock, George P., and Douglas White (1969), ”The Standard Cross-Cultural Sample,” *Ethnology* 8: pp. 329-369.

Price, T.D. (1995), "Social Inequality at the Origins of Agriculture," in G.M. Feinman and T.D. Price (eds) *Foundations of Social Inequality*, London: Plenum Press, pp. 129–151.

Frederic L. Pryor, 1986. The Adoption of Agriculture: Some Theoretical and Empirical Evidence. *American Anthropologist* 88:pp. 894-897.

Shekhar A, C.-J. Dalgaard and O. Moav (2006), "Technological Progress and Regress in Pre-industrial Times," *CEPR Discussion Paper No* 5454

Usher, D., (1987), "Theft as a paradigm for departures from efficiency," *Oxford Economic Papers* 39, pp. 235-252

Usher, D., (1989), "The dynastic cycle and the stationary state," *American Economic Review* 79, pp. 1031-1044

Weisdorf, J.L. (2005), "From Foraging to Farming: Explaining the Neolithic Revolution," *Journal of Economic Surveys* 19, pp. 561-586

Williamson, O.E. (1973), "Markets and Hierarchies: Some Elementary Considerations," *American Economic Review* 63, pp. 316-325

FIGURE 1
RAID RATIO AND TECHNOLOGY

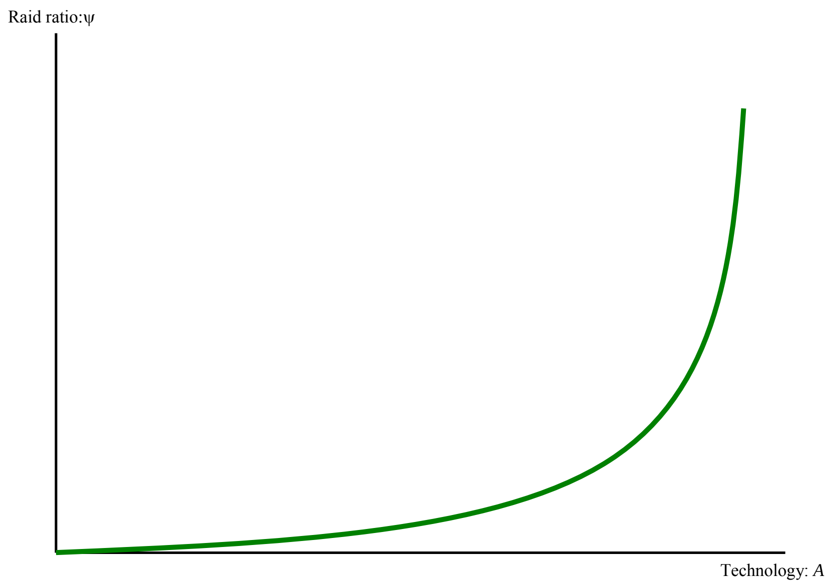


FIGURE 2
OPTIMAL TAX RATE AND TECHNOLOGY

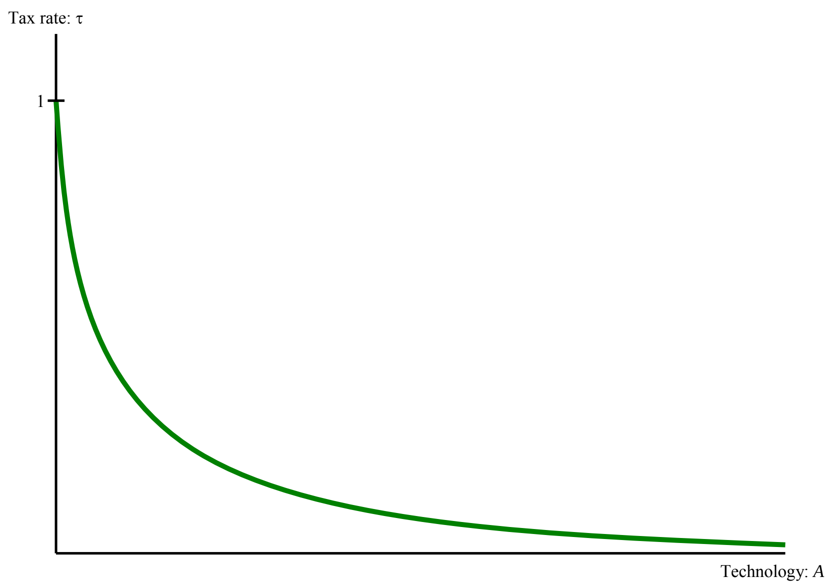


FIGURE 3
PAY-OFFS AND TECHNOLOGY

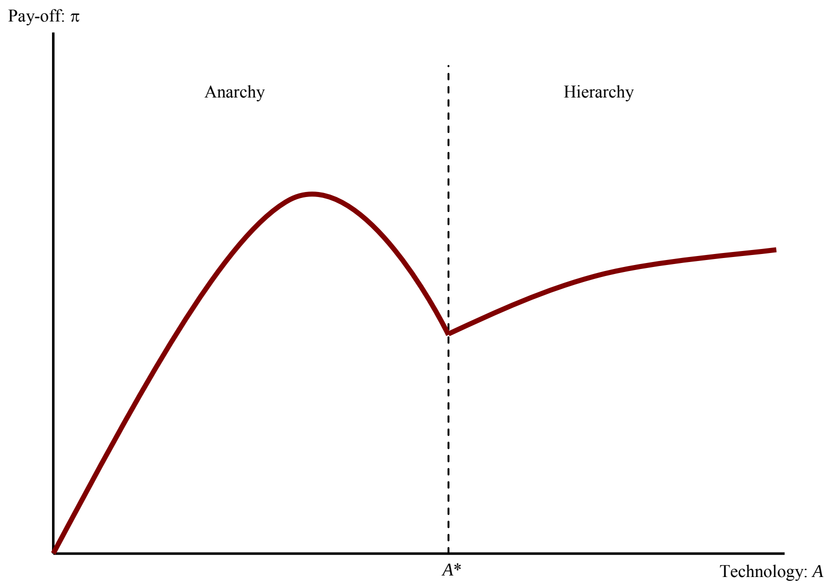


FIGURE 4
POPULATION AND TECHNOLOGY DYNAMICS
(TWO POTENTIALLY STABLE STEADY STATES)

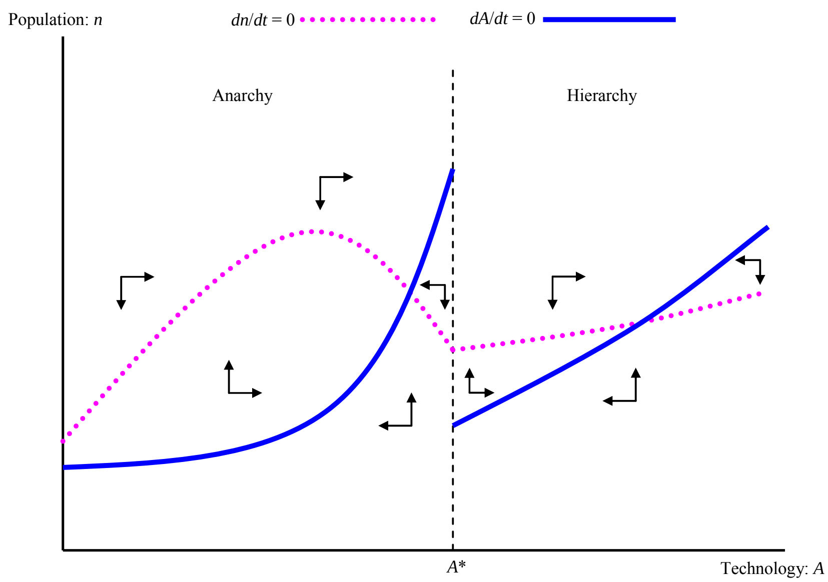
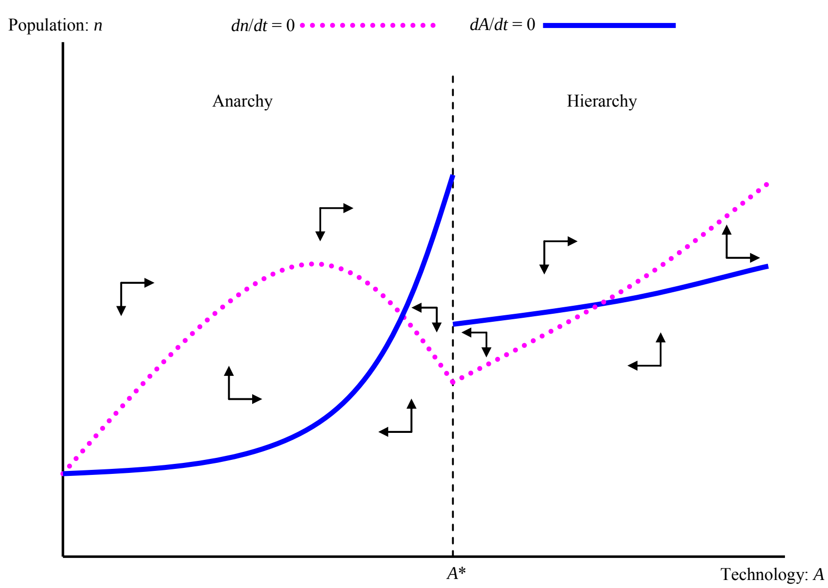
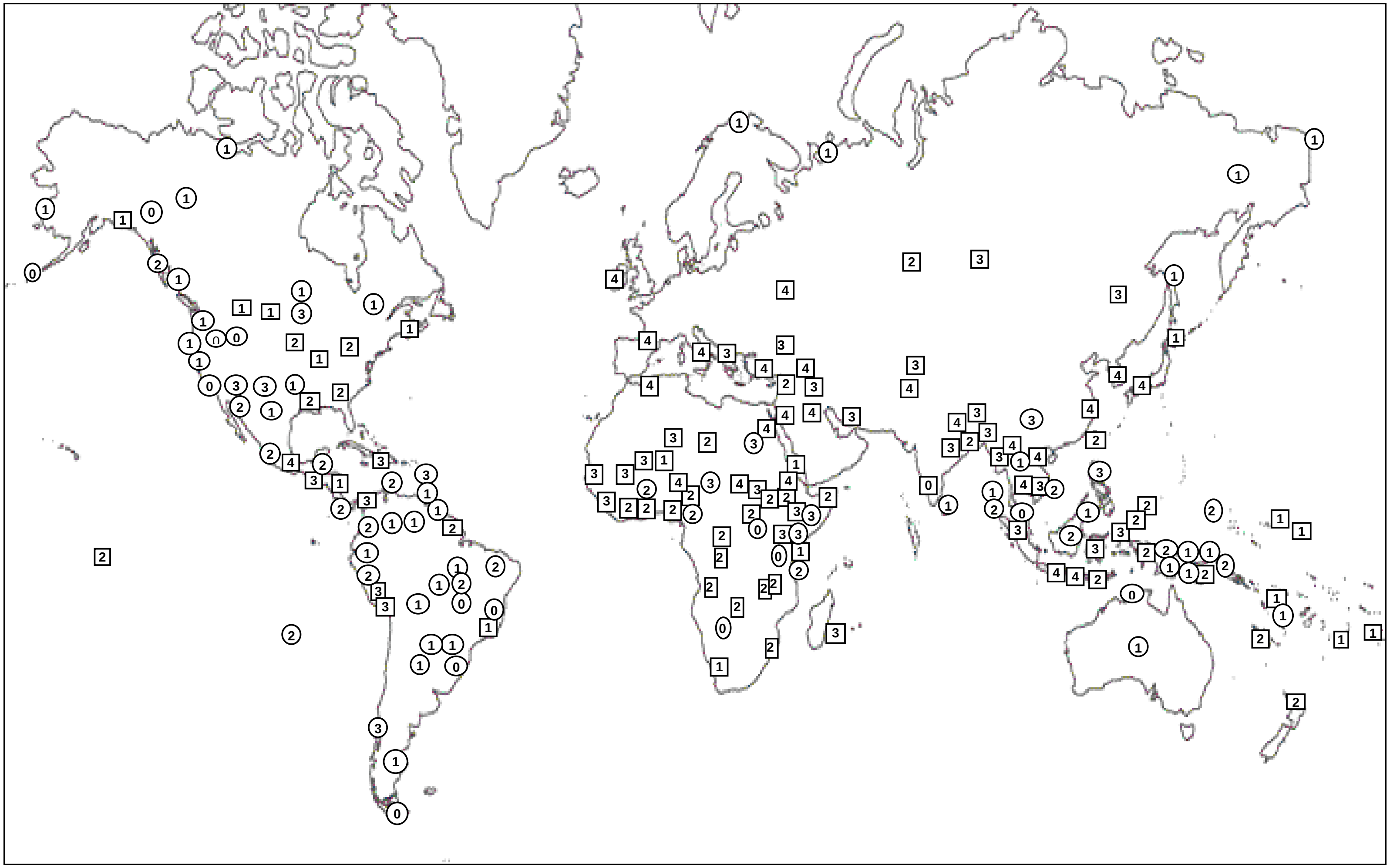


FIGURE 5
POPULATION AND TECHNOLOGY DYNAMICS
(ONE POTENTIALLY STABLE STEADY STATE, ONE SADDLE-PATH STABLE)





Map 1: The geographic distribution of technology and hierarchy in the SCCS

TABLE 1: DESCRIPTION OF VARIABLES AND SOURCES USED IN THE ANALYSIS

	Description	Source
Technology and population		
Jurisdictional Hierarchy	=1 if none, =2 if petty chiefdoms, =3 if larger chiefdoms, =4 if states, =5 if there are multi-layered states	SCCS
Policing present	=1 if police are present in the society, =0 otherwise	SCCS
Agriculture Scale - Importance of Agriculture and its Nature	=1 if none, =2 if <10% of food supply, =3 if >10% and of secondary importance in food supply, =4 if primary but not intensive, =5 if primary and intensive	SCCS
Storage Surplus	=1 none or barely adequate, =2 simple storage or adequate, =3 if complex and more than adequate	SCCS
Technological specialization	=0 if no specialization present, =1 if pottery only, =2 if loom weaving only, =3 if metalwork only, =4 if smiths, weavers, and potters	SCCS
Writing and Record-Keeping	=0 none, =1 Mnemonic devices, =2 Non-written records, =3 True writing, no records, =4 True writing, records	SCCS
Population Density	=1 if < 1 persons per square mile, =2 1-5 persons per square mile, =3 if 5-25 persons per square mile, =4 if 25-100 persons per square mile, =5 100 persons per square mile	SCCS
Technology index	1st Principal component of Specialization, Writing, and Agricultural Contribution	
	Exogenous Variables	
Environmental Characteristics		
Mean Rainfall	Mean yearly rainfall (cm)	SCCS, from Cashdan (2003)
Climate suitability for agriculture	Scale ranging from 0 (impossible) to 4 (very good) developed by Pryor using FAO/UNESCO reports	SCCS, from Pryor (1986)
Soil suitability for agriculture	Scale ranging from 0 (impossible) to 4 (very good) developed by Pryor using FAO/UNESCO	SCCS, from Pryor (1986)
Land slope	Scale ranging from 2 to 4, 2=steep, 4=relatively flat	SCCS, from Pryor (1986)
No. habitats w/in 200 miles	Based on counting the number of vegetation types, ocean and lake presence within 200 mile diameter	SCCS, from Cashdan (2003)
Geography/Time		
Distance from fertile crescent	Calculated using society coordinates in SCCS, with the fertile crescent at 45E, 35N (.786, .611 in radians)	Baker (2005)
Closer to another hearth?	=1 if closest to another original hearth of agriculture (Northeastern U. S., Central America, South China)	Baker (2005)
Distance from closest hearth	Calculated using society coordinates in SCCS, with the Northeastern U. S., Central America, and South China as other hearths.	Baker (2005)
"Vertical" distance from fertile crescent	Calculated to be miles north or south from fertile crescent.	Baker (2005)
"Vertical" distance from closest hearth	Miles north or south from nearest hearth of agriculture/civilization.	Baker (2005)

TABLE 2: SUMMARY STATISTICS OF ALL VARIABLES, BY PRESENCE OR ABSENCE OF HIERARCHY

Variable	All Societies N=186	Hierarchical Societies N=102	Non-Hierarchical Ssocieties N=82
Policing Dummy	.33 (0.47)	0.53*** (0.50)	.10*** (0.30)
Population Density Scale	2.86 (1.56)	3.48*** (1.36)	2.04*** (1.39)
Technology Index – First Principal Component of Agriculture, Spec., and Writing Scales.	0 (1.41)	0.64*** (1.31)	-0.83*** (1.00)
Contribution of Agriculture	3.45 (1.51)	3.97*** (1.21)	2.77*** (1.58)
Technological Specialization	3.09 (1.41)	3.65*** (1.32)	2.38*** (1.21)
Writing and Record Keeping	2.35 (1.47)	2.89*** (1.62)	1.65*** (0.85)
Storage Surplus	1.81 (0.72)	1.92** (0.70)	1.67** (0.72)
Exogenous Variables			
Mean rainfall (cm/year)	140.74 (106.00)	136.72 (99.88)	147.84 (113.92)
Climate suitability for agriculture	3.13 (1.16)	3.29* (0.92)	2.96* (1.40)
Soil suitability for agriculture	2.07 (0.77)	2.18** (0.68)	1.91** (0.86)
Land slope	3.29 (0.74)	3.28 (0.73)	3.33 (0.73)
Number habitats w/in 200 miles (N=172)	3.93 (1.35)	3.87 (1.32)	3.95 (1.36)
Distance from Fertile Crescent (miles)	4996.70 (2456.02)	4137.96*** (2495.33)	6036.95*** (1961.32)
Vertical Distance from Fertile Crescent (miles)	1859.50 (1248.37)	1623.39*** (1160.26)	2144.03*** (1298.53)
Closer to another hearth?	0.66 (0.47)	0.5*** (0.50)	0.85*** (0.36)
Distance to closest hearth (miles)	2395.27 (1256.75)	2237.04* (1312.33)	2594.23* (1170.73)
Vertical distance to closest hearth (miles)	1647.46 (1077.28)	1452.05*** (1109.03)	1884.60*** (1001.32)

** Difference in means significant at 5%
*** Difference in means significant at 1% (assuming unequal variances)

TABLE 3: AVERAGE POPULATION DENSITY AND TECHNOLOGY BY THE DEGREE OF HIERARCHY.

Degree of Hierarchy	*Obs.*	*Population Density*	*% Police Presence*	Technology variables				
				Writing scale	Agriculture scale	Specialization scale	Storage scale	Technology scale
1	82	2.04	0.09	1.65	2.77	2.37	1.67	-0.83
2	48	2.94	0.27	2.06	3.44	2.96	1.87	-0.17
3	23	3.74	0.61	2.73	4.13	3.69	1.70	1.16
4	19	3.89	0.79	4.21	4.47	4.47	2.11	1.67
5	12	4.50	1.00	4.41	5.00	5.00	2.25	2.18

TABLE 4: SOME EMPIRICAL MODELS OF HIERARCHY AND TECHNOLOGY

Dependent variable	Degree of Hierarchy	Degree of Hierarchy	Degree of Hierarchy	Hierarchy Dummy	Hierarchy Dummy
Method	OLS	OLS	IV, GMM	Probit	Probit, IV
Population Density	0.165*** (0.054)	0.108** (0.053)	-0.386 (0.279)	0.273*** (0.090)	-0.402 (0.510)
Technology Index	-	0.557*** (0.059)	1.347*** (0.339)	-	1.687** (.760)
Agricultural scale variable	0.043 (0.057)		-	0.032 (0.094)	-
Technological specialization scale	0.240*** (0.057)		-	0.235** (0.095)	
Storage and Surplus	0.137 (0.090)	0.123 (0.093)	.736* (.386)	0.118 (0.153)	1.531* (0.917)
Writing and Record-Keeping	0.339*** (0.049)		-	0.331*** (0.097)	-
Constant	-0.314 (0.215)	1.563*** (0.214)	1.220 (0.807)	-2.337*** (0.422)	-2.079*** (2.084)
Distance Controls?	No	No	Yes	No	Yes
R^2	0.560	0.522	0.533	0.286 (pseudo)	-
Adjusted R^2	0.548	0.514	0.528	-	-
Hansen J-Statistic: (p-val)			.702 (.704)		
Partial F-statistics:					
Population Density			6.46		
Technology Index			5.32		
Storage and Surplus			3.04		
Obs.	184	184	170	184	170

Notes: * significant at 10%, **=significant at 5%, ***= significant at 1%. In the case in which an instrumental variables regression is reported, the instruments used are the environmental and geographical variables (described in table 1). Standard IV estimates are obtained using GMM to allow for heteroskedasticity of unknown form. Probit IV estimates are obtained using a two-step estimator.