

A theory for the formation of large agrarian empires

Peter Turchin
Department of Ecology and Evolutionary Biology
University of Connecticut
Storrs, CT 06235, USA
and
Santa Fe Institute
Santa Fe, NM 87501, USA

e-mail: turchin@santafe.edu

Abstract

Before the Industrial Revolution the greater part of the inhabitable world was occupied by small-scale societies and large territorial states were, comparatively speaking, a rarity. Nevertheless, between 3000 BCE and 1800 CE there were at least 60 agrarian “megaempires” that controlled at the peak an area equal to or greater than one million of squared kilometers. What were the social forces that kept together such huge agrarian states? A clue is provided by the empirical observation that over 90 percent of megaempires originated at steppe frontiers—zones of interaction between nomadic pastoralists and settled agriculturalists. I propose a model for one route to megaempire. The model is motivated by the imperial dynamics in East Asia (more specifically, the interface between the settled farmers of East Asia and the nomads of Central Asia). It attempts to synthesize recent developments from theories of cultural evolution with insights from previous work by anthropologists on nomad/farmer interactions.

Bio-data: Peter Turchin is Professor of Ecology and Evolutionary Biology at the University of Connecticut. His books include *War and Peace and War* (Pi Press) and *Secular Cycles* (Princeton).

INTRODUCTION

Understanding the rise and fall of empires (large territorial states) is one of the most important research directions in world history. Beginning with Gibbon (1932), most theoretical efforts have been directed to the second part—the causes of imperial disintegration and fragmentation (Tainter 1988). The first part of the question, however, is theoretically more challenging. Large territorial empires are a comparative rarity in the historical record. Before the Industrial Revolution the greater part of the inhabitable world was occupied by small-scale societies, and even regions where empires repeatedly rose and fell were, as often as not, in the fragmented state. Thus, the really difficult question is not why large agrarian states kept disintegrating, but how they arose in the first place.

What were the social forces that kept together huge states controlling populations of millions or tens of millions spread over millions of squared kilometers? What were the preconditions for the rise of such “megaempires,” defined here as territorial states that controlled at the peak an area equal to or greater than one megameter squared (one million of squared kilometers)? Despite some promising approaches, reviewed, for example, by Michael Mann (1986), these questions remain essentially unanswered.

In this article I propose a model for one route to megaempire. The model is motivated by the imperial dynamics in a particular world region, East Asia (more specifically, the interface between the settled farmers of East Asia and the nomads of Central Asia). It attempts to synthesize recent developments from theories of cultural evolution with insights from previous work by anthropologists on nomad/farmer interactions. The main focus of the paper is not on the origin of the state, but rather on the processes explaining how small states scale up to megaempires.

“IMPERIOGENESIS” IN THE EAST ASIAN REGION

As will be addressed in greater detail in a later section, the spatial distribution of megaempire occurrence is highly clumped. Many world regions never served as loci for such states, while in other regions empires rose sporadically and in a few locations repeatedly. One world region, unique in that it has seen a continuous sequence of rise and fall of empires since the Bronze Age, is China (Table 1). Beginning with the Qin unification in the third century BCE, the periods of fragmentation between successive unifications generally did not exceed a century (although some unifications were partial, see Table 1 for details).

Table 1 also quantifies another striking pattern that has been noted by world historians (e.g., McNeill 1963). Only one out of the fifteen unifications listed in Table 1—the establishment of the Ming dynasty in the fourteenth century—did not originate in the North. All others started in the North: three from the northeast (Liao peninsula, Manchuria), three from the north central (Huang He region), and eight from the northwest (most often from the Wei River Valley). The importance of the North for Chinese empires is also suggested by the location of capitals (Table 1). Capitals tend to be located in areas where rulers draw most of their support. What is particularly interesting is that in the beginning of the second millennium C.E. the economic center of

China has shifted south to the Yangtze river, but the political center remained in the North.

The political centers of Chinese empires, thus, were located not in geopolitically safe locations, but on, or near China's "perilous frontier" (borrowing from the title of the insightful book by Thomas J. Barfield). Before the intrusion of the imperialist powers in the nineteenth century, China was never seriously threatened from any other direction than the North (Barfield 2001). By contrast, there was almost continuous military pressure in the North, from Turco-Mongolic steppe dwellers in the Northwest and the Tungusic people of Manchuria in the Northeast, periodically punctuated by successful invasions that occupied northern China and, on two occasions, the whole of China. Is it possible that this geopolitical pressure may help us understand why northern China was such a "hotspot of imperiogenesis"?

The importance of the steppe frontier in Chinese history has not escaped the attention of scholars (e.g., Lattimore 1967, Barfield 1989). Of particular interest are the ideas of anthropologists Thomas Barfield and Nikolai Kradin on the interactions between the Chinese and the steppe nomads (Barfield 1989, 1990, 1994, Kradin 2000, 2002, 2005). These authors showed that the scale of political organization among the nomads was directly proportional to the size of the neighboring agrarian empires. Thus, the greatest imperial confederations of nomads in world history (the Xiongnu, the Turks, and the Mongols) arose on the steppe side of the northern China frontier. In other words, the exceptionalism of the East Asian imperiogenesis hotspot was mirrored in the exceptionalism of repeated rise of gigantic imperial confederations in the steppe. Furthermore, there was a substantial degree of temporal synchrony between the rise of the steppe imperial confederations and Chinese empires—Xiongnu and Qin/Han, Turks and Sui/Tang, Mongols and Sung (Barfield 1989).

Barfield and Kradin (Barfield 1989, 1990, 1994, Kradin 2000, 2001, 2002, 2002, 2005) argue convincingly that political organization of the pastoral nomads on a large scale is impossible without a nearby settled society. The reason is that the nomadic society does not produce the surplus in the form useable to support the state. Collecting taxes from the nomads is problematic because free nomads are at the same time warriors well able to resist any coercive pressure brought on them. Nomads also have another option of resisting taxation that farmers do not have—moving their herds away. Finally, the chief product of the steppe economy, livestock, cannot be easily stored (unlike grain produced by agrarian economies). Thus, political organizations of the nomads had to draw their resources from the agrarian societies—by robbing the farmers, by extorting tribute from the agrarian states, or by controlling trade routes. Kradin refers to this type of political organization as *xenocratic* (from *xenos* "stranger" and *kratos* "power"). This argument leads to a hypothesis about why there is a correlation between the sizes of agrarian states and nomadic confederations. As the scale of agrarian states in East Asia increased, the nomads needed to cooperate on an increasingly greater scale in order to overcome the state defenses for successful raiding, or present a credible threat to extort the tribute. Additionally, larger and richer sedentary states possessed greater wealth that nomads could extract, thus enabling larger nomadic polities. That is why Barfield calls

the nomadic confederations the “shadow empires”, because their size mirrored the size of agrarian states (Barfield 2001).

The direction of causation proposed by Barfield, then, is that the appearance of agrarian megaempires explains the rise of nomadic imperial confederations. This appears to be a valid inference. However, if the presence of a large agrarian state exerted an integrative influence on the nomads, why should not the presence of a large nomadic confederation have a similar effect on the farmer societies? In the next section I propose a verbal model exploring the possible consequences of such a nonlinear feedback loop.

A MODEL OF MIRROR STATE FORMATION ON THE FARMING/NOMADIC FRONTIER

Cooperation as the basis of society

Before I discuss the logic of the model, I need to deal with the general question raised in the introduction—what were the social forces that held together agrarian states? I explored this issue in greater detail elsewhere (Turchin 2006: Chapter 5), so here I only give a brief summary. The first important point is that attempts to build a theory of society by assuming that all people behave in entirely self-interested manner have been unsuccessful. This premise was clearly enunciated in the fifteenth century by Niccoló Machiavelli in *The Prince*, and in the twentieth century was formalized as the rational choice theory (for a systematic application of the rational choice theory in sociology, see Coleman 1990). Another influential current in social theorizing, which also assumes that the society is made up of rational agents, was the theory of social contract, going back at least to *The Leviathan* of Thomas Hobbs. Unfortunately, all attempts to develop a theory of society based on Hobbesian contractarianism founder on the free-rider problem (Kraus 1993). More generally, it is becoming clear that if people acted on a purely rational (self-interested) basis, they would never be able to cohere into society (Collins 1992:9). The fundamental problem is that for a society to exist and function, its members must produce public goods (for example, collective defense), in which the costs are born by each contributor, while the benefits are shared evenly across all members. The rational behavior in such situations is to benefit from the public-good production by others, while not contributing oneself. If all society members are such self-interested agents, then no public good would be produced, and there will be no society, but merely a collection of free agents. Thus, the basis of society can come only from precontractual solidarity (Durkheim 1915).

Fortunately, the rational choice theory does not describe well the behavior of real people (Bowles and Gintis 2002, Fehr and Gächter 2002). The recent flowering of research in experimental economics has shown that there is a great degree of variation among humans in both modern and traditional societies (Henrich et al. 2004). Some people indeed behave as rational agents (following David Hume, let's call them “knave”).¹ In most societies, however, the majority is motivated not only by the prospect of gain and avoidance of punishment, but also by *social norms* that promote social

¹ “Political writers established it as a maxim, that, in contriving any system of government ... every man ought to be supposed to be a *knave* and to have no other end, in all his actions, than his private interest” from David Hume, *Essays, moral, political, and economic*.

cooperation (and no societies consisting entirely of rational agents have been found so far). For example, experimental economists showed that about half of experimental subjects coming from industrial societies were willing to pay money to punish free riders in the public goods game (Fischbacher et al. 2001) —a distinctly non-rational strategy. But it was the action of such “moralistic punishers” (or “moralists”, in short) that stabilized cooperation and enabled the group to produce public goods. In the absence of punishment, cooperation unraveled because the free riders took advantage of cooperators, following which the cooperators themselves stopped contributing to the common pot, because they did not want to be taken advantage of.

How do these insights from experimental economics help us to understand the dynamics of state formation on the steppe frontier? One of the most important types of collective good for small-scale societies is mutual defense, as well as its logical counterpart—collective predation on other groups. Collective defense is a public good, because its costs are born privately (the probability of being killed or seriously injured), while its benefits (group survival, or successful defense of the group’s territory, crops and herds) are shared automatically among all members. Thus, the rational strategy is to free ride—let others bear the dangers of fighting the invader, while benefiting from their ability to repel the attack. However, groups consisting entirely, or largely of knaves (free riders) will be unable to cooperate in collective defense, and thus will not persist. On the other hand, groups of moralists, motivated by the social norm of sharing equally in collective defense, would present a unified front to the invaders, and will have a good chance of repelling them. To be successful, a group does not need to consist entirely of moralists. As long as there are enough moralists in the group to bring it to the cooperative equilibrium (by forcing knaves to contribute to the common good), the group will be capable of effective collective action. Still, clearly the smaller the proportion of knaves in the group, the more cohesive and effective it is in producing collective goods.

It is not my intention to imply that group composition (in terms of knaves, moralists, etc) is the only factor explaining the capacity of groups for collective action. In addition to group composition, other important factors that contribute to group success in warfare and other socially meaningful activities include religion, military organization, and many more. I lack space to expand on this issue here and thus refer the reader to my previous publications on this issue (Turchin 2003: Chapter 3, 2006: Chapter 5). What is important for the argument in this paper is that different groups are characterized by different degrees of group cohesion and solidarity and, therefore, different abilities for concerted action—in warfare as well as in other endeavors. As Charles Darwin argued in *The Descent of Man*, more cohesive groups will have a better chance of surviving and displacing less cohesive groups and, as a result, the cooperative behaviors will spread in the population.

Perhaps the greatest theorist of group cohesion was the fourteenth century Arab sociologist Abd-ar-Rahman ibn Khaldun (Ibn Khaldun 1958). The key concept of Ibn Khaldun was *asabiya*, which can be loosely translated as “group feeling.” According to Ibn Khaldun, *asabiya* produces “the ability to defend oneself, to offer opposition, to

protect oneself, and to press one's claims" (Ibn Khaldun 1958:I:289). Preponderance of *asabiya* allows a group to dominate the rivals (Ibn Khaldun 1958:I:lxix).

Ibn Khaldun was a native of the Maghreb (northwestern Africa), and was intimately familiar with the social interactions on the nomadic/agricultural frontier. His insight into the social forces underlying the creation of nomadic confederations of the Berbers is extremely useful for the task of understanding imperial dynamics at the opposite end of Afroeurasia, on the East Asian steppe frontier. However, the outcomes of the nomad/farmer interactions in the Maghreb were different—there were no mighty agrarian empires in northwestern Africa, similar to the Chinese ones, and the scale of Maghrebin states was much smaller than those put together by the Turco-Mongolian peoples. The reasons for such divergent outcomes will be discussed below; for now, I accept the basic insight of Ibn Khaldun and explore how it can be employed in modeling the situation in East Asia.

The model

The starting point of the model is provided by two fundamental postulates. First, there is a steep environmental gradient in average rainfall within the modeled area. The well-watered side of the ecological frontier is inhabited by settled agriculturalists, while the arid zone provides sustenance to pastoral nomads. Second, the relations between farmers and nomads are characterized by a strong anisotropy. On one hand, "pure nomads" (that is, nomads who do not have opportunity to practice even limited forms of agriculture) cannot produce certain necessities (such as grain), which they must somehow obtain from the agriculturalists (Khazanov 1984). On the other hand, the nomads enjoy a preponderance of military power over the farmers, which became especially pronounced with the invention of mounted archery in the beginning of the first millennium B.C.E. In other words, the pastoralist nomads have both an incentive and ability to take agricultural products away from farmers by force.

Let us now consider the consequences of these premises for the interactions between farmers and nomads. Suppose that initially the modeled area is occupied by small-scale pastoralist and agricultural communities (band societies, autonomous villages). Because the nomads need the products of agriculture and enjoy the preponderance of military power, they routinely rob their farmer neighbors. For example, a nomadic community could send its warriors on a raid at, or soon after the harvesting time, when agricultural products are most abundant. A successful raid can be devastating to a farming community. Not only it loses a large part of resources it needs to survive until the next harvest, but the nomads may also kill men of fighting age and abduct women and children as slaves. Thus, raiding pressure from the steppe imposes a severe selective regime on farming communities. Internally cohesive communities, which developed effective ways of military organization, have a better chance of surviving a nomad raid. Less cohesive and militarily able polities are likely to disintegrate in the aftermath of a raid (which does not mean the physical death of all community members, many of whom could be accepted into surrounding, more fortunate groups).

The raiding pressure, thus, over a long period of time should have selected for more cohesive polities better capable of repelling the nomads. Because of the military superiority of the steppe warriors, however, there are limits to how effective a single farming community could become in defending itself. The only effective way of resisting the nomad pressure is for several local communities to unite into a “meta-community” with a larger defensive force that would offset the military advantage of the nomads.

There are at least three ways in which this scaling up in polity size could be achieved. The simplest and most direct way is for the leaders of several neighboring communities to form an alliance that would coordinate mutual defense. The concept of joint defense was broadly known within Eurasia, as attested by the common motif of the Quarreling Sons and the Bundle of Twigs, which is found in the Greek, Talmudic-Midrashic, and Indic literatures, as well as in many folkloric traditions (Moses 1987).

One potential problem to such voluntary defensive alliances, however, is that they are vulnerable to free riding. An element of coercion may be necessary to stabilize the cooperative equilibrium, analogously to the dynamics in the public goods game, discussed earlier. Therefore, the second route to joint defense, and probably a more realistic one, is for one community to annex several neighbors by force, resulting in a simple chiefdom (tribal society). A chiefdom should be better at coordinating its defensive actions than an alliance of independent communities, because the chief has the capacity to punish any subordinate community that does not contribute to the common defense.

The third route to scaling up is a variant of the “roving bandits–stationary bandits” argument of Mancur Olson (2000). In this scenario, an enterprising group of nomads may decide to make a permanent move to the agricultural side of the frontier, where they subjugate several farming communities, thus converting themselves from roving to stationary bandits. With time, these nomads assimilate to the farmers’ language and culture (or vice versa), so the end point is similar to the one in the second route, except that the chiefly elites are of nonnative (steppe) origin.

Whatever the particular mechanism of scaling up, the end result is the same. By virtue of its increased size, the scaled-up farming polity has shifted the balance of military power in its favor. Now the selective pressure for scaling up begins to work on the steppe side of the frontier. The nomad communities need to band together in order to nullify the numerical advantage of the agrarian polities. We expect that ability to conduct successful raids would again require a mixture of cooperation and coercion, but with a lesser degree of coercion than on the farming side. First, there is less of an opportunity to free ride (those who do not join the raiding party do not share in the booty). Second, as we discussed above, it is more difficult to coerce a nomadic group. It is interesting that the Quarreling Sons motif appears in the *Secret History of the Mongols* no fewer than four times (Moses 1987). (Naturally, in the steppe version, it is not twigs that are bundled together to illustrate the value of cooperation, but arrows). Such a frequent occurrence possibly reflects the difficulty of uniting the nomads by force, and a greater emphasis on ideological means to achieve the same end.

The scaled-up steppe polities regain their military advantage, which puts pressure on the agrarian polity to scale up again, by evolving from simple to complex chiefdoms. The initial anisotropy in military power on the farming-steppe frontier, thus, sets up an autocatalytic process by which the pressure for scaling up is brought on both the farming and nomadic polities, resulting in a runaway evolution of polity sizes on both sides of the frontier. The autocatalytic process is stopped either by running out of space or due to the logistics limitation on empire sizes. Furthermore, once centralized agrarian states arose, the steppe nomadic confederations learned that they could extort tribute simply by threatening raids (Barfield 1989). However, the pressure for maintaining large size of the confederation did not slacken because of the need to pose a credible threat to the agrarian empire.

Note that although the starting point of the model is local communities, this is not a critical assumption. The main argument is concerned with the social scaling-up process—from polities to “metapolities” (a *polity* is an independent political organization; depending on the number of control levels it can be a local community, simple chiefdom, complex chiefdom, state, or empire). Thus, simple chiefdoms evolve into complex ones. The latter, in turn, evolve into archaic states, and so on all the way to agrarian megaempires and imperial confederations of the nomads.

Application of the model to the North China region

The model described above is quite abstract, and it is worth discussing in what ways it oversimplifies the complex situation on the Chinese Inner Asian frontier. The first postulate of the model (a steep environmental gradient) appears to describe well the area encompassing the imperial core in northern China and the deserts and steppes to its northwest. One needs to travel only 300 km from Chang'an (modern Xi'an) to reach the Ordos desert (the homeland, or one of the homelands of the Xiongnu), and even less from Beijing to arrive at the outskirts of the Gobi. Thus, the farming communities of northern China were within easy raiding distance for the steppe nomads. Experiments with well-fed Mongolian horses indicate that nomads could traverse this distance (300 km) in a week or less, while a trans-Gobi raid (1800 km) would require 25 days of traveling time (Hoang 2001).

The second postulate, the military superiority of the nomads, also appears to describe well the history of Chinese relations with the steppe, especially after the nomads mastered mounted archery. Only a centralized Chinese state was able to mount successful defense against the nomad raids. The periods of internal disunity in China were typically associated with increased raiding pressure from the steppe. Even with China unified, many Chinese rulers found it necessary to buy the nomads off with tribute, thinly disguised as “gifts” (Barfield 1989).

One feature of the model that may strike the historians of China as unrealistic is that at the starting point the modeled area is populated by smallest scale polities, local communities. Indeed, my focus is on the advantage that the invention of mounted archery gave the pastoral nomads, and by the time this military technology spread to East Asia,

the social evolution in this region had already led to the rise of complex chiefdoms and archaic states. For example, the earliest state in China probably arose during the Erlitou period (1900–1500 BCE) (Liu et al. 2002). I have two responses to this potential critique. First, the nomads probably enjoyed some degree of military superiority over the farmers even before mounted archery. The invention of the chariot and its spread from the Great Steppe during the second millennium BCE provides one example. Second, the starting point is not the main feature of the model. The model explains why there is a repeated transition from polities to metapolities (the runaway social scaling-up process) on both sides of the steppe frontier. The main question of interest is not the starting point, but the process leading to the end point—the rise of very large-scale polities, the megaempires.

What were the limits that eventually stopped the runaway growth of frontier empires? For the nomadic imperial confederations, the answer is, probably, that they simply run out of the steppe. The greatest nomadic empire, that of Chinggis Khan and his successors, encompassed the whole of Eurasian steppe (with the exception of the western extension of the Steppe, the Hungarian Plain, which was conquered, but then abandoned). By contrast, the size of the Chinese empires was probably limited by the logistic constraints. The Chinese empires ventured into such agricultural regions as Southeast Asia and the Korean peninsula, but did not annex them in a systematic fashion.

This consideration brings us to the question of why agrarian megaempires and imperial nomadic confederations did not form in the Maghreb, the area studied by Ibn Khaldun. The answer suggested by the model is that the agrarian polities of Northwest Africa simply lacked the deep hinterland of China into which they could expand. Within the Maghreb (North Africa west of Egypt), agriculture is possible only within a thin band stretching along the Mediterranean coast. As a result, these agrarian states lost the scaling-up race to the nomadic confederations and were conquered by them, setting up the dynastic cycling dynamics so admirably described by Ibn Khaldun. The great majority of desert-originating dynasties in the Maghreb (Zirids, Marinids, Wattasids, etc), thus, reached only medium size.

Only three states founded by the Berbers broke through the 1 Mm^2 threshold: the Fatimids, the Almoravids, and the Almohads (Table 1), and these achievements were possible only because these states started, or moved their center of gravity outside the Maghreb. The Fatimids originated in Tunisia in 909, when the Kutama Berbers were converted to the Shia Islam, but very soon they conquered Egypt and moved the capital to Cairo (in 969). It was the agricultural wealth of Egypt that financed their expansion into the Levant and the Arabian Peninsula. The Almohads moved in the other direction. After defeating the Almoravids in 1144, they first conquered Andalusia before turning East to Algeria and Tunis. Their capital was moved to Seville in 1170. Finally, the Almoravids began their imperial career outside the Maghreb, on the steppe frontier with the agrarian Ghana empire (which they ended up conquering; later they also added to their empire the Moorish Spain). Thus, all three of these Berber megaempires interacted with, and eventually moved their centers of mass to large and productive agricultural areas outside Northwest Africa.

MEGAEMPIRES: THE SPATIAL AND TEMPORAL PATTERNS

Although the development of the model was motivated by historical dynamics in East Asia, the interactions assumed by the model are generic to Afroeurasia, and thus it would be interesting to determine how well its predictions hold beyond East Asia. The empirical basis for the test is provided by Rein Taagepera's (1978a, 1978b, 1979, 1997) compilation of territorial dynamics of historical states, which has been systematized and posted on the web (<http://www.irows.ucr.edu/>) by Chase-Dunn and coworkers (Chase-Dunn et al. 2007), and further extended by Turchin, Adams, and Hall (2006). Our focus is on land-based empires (both agrarian and nomadic) whose peak territory exceeded 1 Mm², and not on thalassocratic (sea-based) powers, such as Athens, Venice, or modern maritime empires of the European Great Powers. Furthermore, because we are interested in pre-industrial states our list of empires goes only until 1800.

There are 61 historical megaempires in our list (Table 2), all but one (Inca) in Afroeurasia. The great majority (56 out of 61) of these empires are situated in, or next to the arid belt that runs through Afroeurasia, from Sahara in the West to Gobi in the East (Figure 1). The exceptions include three European states (the Roman and Carolingian Frankish empires; Lithuania-Poland) and one empire in Southeast Asia (Khmer). In fact, Lithuania-Poland is a "semi-exception"—although it started in the swamps and forests of northeastern Europe, its expansion during the fourteenth century into former Kievan lands exposed it to very painful contact with first the Golden Horde and later the Khanate of Crimea. In summary, the relationship between location on a steppe frontier and the rise of megaempires is not a rigid one, but there is strong statistical regularity.

Figure 2 presents the temporal evolution of the largest empire size in the Taagepera–Chase-Dunn database (Chase-Dunn et al. 2007). During the third and second millennia BCE the maximum empire size fluctuated between 0.3 and 1 Mm², albeit with a gradual trend up (these were several Egyptian empires, the Akkad, and the Shang). Between 800 and 200 BCE, however, there was a rapid increase of maximum territory, from 0.4 Mm² in 900 BCE to 1.4 in 670 BCE (the neo-Assyrian empire), then to 5.5 in 500 BCE (the Achaemenid Persia) and finally to 9.0 Mm² in 180 BCE (the Xiongnu). It is remarkable that this dramatic upsweep in the maximum area coincided almost precisely with the Axial Age (Jaspers 1953), usually dated to 800–200 BCE. After the upsweep of the Axial Age, the maximum imperial size continued to increase, but at a much slower rate (Figure 2). Karl Jaspers himself speculated that the great religious and philosophical breakthroughs of the Axial Age were a response to the political and social instability brought on by the increased pressure from the nomadic steppe dwellers, which, in turn, was due to the military breakthrough of mounted archery.

We know that the first imperial steppe confederation, that of the Scyths, was instrumental in bringing the end to the neo-Assyrian empire, and that the Achaemenid struggle against the Scyths/Sakai was in many ways similar to the struggle of the Han Empire against the Xiongnu. Was the simultaneous development during the Axial Age of military superiority of the nomads, the rise of great world religions, the radical increase in the world urbanization rate (Korotayev 2006), and the upsweep in the imperial size just a coincidence? One possible explanation of this remarkable pattern is offered by the model

developed in this paper. The military superiority of the nomads forced agrarian states to scale up to very large sizes in order to resist the pressure from the steppe. One cultural mechanism for holding together ethnically diverse people in new megaempires was the presence of a unifying, metaethnic (that is, capable of integrating multiethnic populations) ideologies, such as Zoroastrianism in the Achaemenid empire, Buddhism in the Mauryan empire, and Confucianism in the Han empire.

DISCUSSION

The main argument in this paper is that steppe frontiers are very special places for politogenesis and, especially, imperiogenesis; places where very large territorial states are much more likely to arise than elsewhere. To account for this broad macrohistorical generalization, I propose a model of the social scaling-up process that leads to the coevolution of agrarian megaempires and nomadic imperial confederations facing each other across the steppe frontier.

The model advanced in this paper was strongly influenced by the warfare theories of the origin of the state (Carneiro 1970, Webster 1975, Wright 1977). However, this influential current in anthropological theory, in my opinion, suffers from its failure to integrate the insights of Ibn Khaldun's *Muqaddimah*, as well as much more recent developments from the evolutionary theory. State formation involves much coercion and profit-seeking, but an additional key ingredient is cooperation, as I argued above.

The basic idea of evolution of cooperation by group selection was clearly formulated by Charles Darwin (1871): "Selfish and contentious people will not cohere, and without coherence, nothing can be effected. A tribe possessing ... a greater number of courageous, sympathetic, and faithful members, who were always ready to warn each other of danger, to aid and defend each other ... would spread and be victorious over other tribes." During the twentieth century group selection first went through a period of uncritical acceptance, resulting in a lot of bad theorizing, followed by a backlash, when the concept was completely repudiated by evolutionary biologists. The influential figures in sociobiology, such as E. O. Wilson (1975, 1979), preferred to emphasize kin selection and reciprocity as chief evolutionary forces explaining cooperation. The rejection of group selection by evolutionary biologists was reflected in sociological literature (Sanderson 1999).

In recent years, however, group selection, or rather *multilevel selection* (which emphasizes that natural selection can operate at all levels—genes, cells, individuals, and groups) has staged a comeback (Wilson and Wilson 2007). The role of multilevel selection in the evolution of human sociality has been supported with both mathematical models and empirical analyses (Sober and Wilson 1991, Richerson and Boyd 1998, Bowles 2006). As a result of these studies, we now have good understanding of how cooperation evolved in small-scale societies. The great remaining scientific puzzle is the evolution of our ability to combine into huge cooperating groups consisting of millions of unrelated individuals. This is not to say that all inhabitants of, say, the Han empire were selfless altruists. However, the diametrically opposing view, that the Han empire was

held together by force and greed alone, is equally untenable. There had to be a group of “moralistic” individuals broad enough to provide cohesion to this huge conglomerate of humanity that at the peak numbered more than 60 million people spread over 6 Mm² of territory.

The basic premise of my model is that the evolution of such huge societies occurred in a series of steps of adding an extra level of political organization resulting in increasing social and political complexity. This is not a new concept. The level of political integration is reflected in such typologies of societies as band–tribe–chiefdom–state of Elman Service (1975) and is coded in the Human Relations Area Files (Ember and Ember 2001). Furthermore, we know that a general empirical pattern in historical dynamics is the cyclic rise and decline of political complexity (Marcus 1998, Turchin and Nefedov 2008). What has been an unresolved puzzle, however, is why in most places, for example, in the prehistoric southeastern United States, the cycle is between simple and complex chiefdoms (Anderson 1996), while in a few special places, such as East Asia, the social complexity repeatedly went through many more levels of political organization—all the way to megaempire—followed by disintegration down to the level of states and complex chiefdoms. The answer to this puzzle, as I argued in this paper, must be sought in the interactions of farmers and nomads on the steppe frontier.

Steppe frontiers are a subset of what I termed in my previous work (Turchin 2003, 2006) *metaethnic frontiers*—areas where two metaethnic communities come in contact and conflict. *Metaethnic communities* (from the Greek *meta*—beyond and *ethnos*—ethnic group, nation) are the broadest, supranational groupings of people that include not only “civilizations” (Toynbee 1956, Huntington 1996), such as the Western, Islamic, or Sinic, but also other broad cultural groupings, e.g. the Iron Age Celts or Turco-Mongolian steppe nomads. Typically, cultural difference is greatest between people belonging to different metaethnic communities; sometimes this gap is so extreme that people deny the very humanity of those who are on the other side of the metaethnic fault line. Metaethnic frontiers are zones where groups come under enormous pressure, and where ethnocide, culturicide, and genocide, but also ethnogenesis, commonly occur. Intense intergroup competition creates conditions for the emergence of groups with high capacity for collective action that eventually transform themselves into expansionist states (Turchin 2006).

Steppe frontiers often anchor very deep cultural divides. Pastoralist nomads and settled farmers have very different ways of life, and other divisive cultural markers tend to follow suit. For example, the farming people in eastern Europe tended to adopt Christianity, while the nomads and seminomads (Volga Bulgars, the Golden Horde, and the Crimean Tatars) converted to Islam, thus adding a religious dimension to the farmer–nomad antagonism. As a result, steppe frontiers tend to be among the most intense metaethnic frontiers, which is why there is a strong statistical association between these frontiers and megaempires. Less intense metaethnic frontiers should be associated with expansionist states of lesser scale. This expectation is borne out by the empirical test, which I conducted for Europe during the first two millennia CE (Turchin 2003: Chapter 5). The history of Europe is known much better than the rest of Afroeurasia, and thus it

was possible to achieve a much more detailed quantification of metaethnic frontiers. I also considered a much broader spectrum of polities, compared to the present study, by including in the database all states that had peak territory greater than 0.1 Mm². My conclusion was that there was a strong statistical correlation between the locations of frontiers and regions where expansionist states originated. Thus, the metaethnic frontier theory has been now tested in one region, Europe, where a detailed quantification of frontiers and medium-scale polities is possible (Turchin 2003: Chapter 5), and more broadly within Afroeurasia by focusing on only one kind of frontier and the very largest states (this paper). Because both studies yielded results supportive of the metaethnic frontier theory, our confidence that we have identified a valid macrohistorian generalization is correspondingly enhanced.

Acknowledgments

The author wishes to express gratitude to Eugene Anderson for help with compiling the information in Table 1, and to Nikolay Kradin for information on Central Asian nomads and for comments on the manuscript. Many thanks to Sergey Gavrilets and Andrey Korotayev for their comments on the manuscript. This research has been supported by the research grant from NSF 05-520 Human and Social Dynamics. The manuscript was written while I was on sabbatical at the Santa Fe Institute. I am grateful for the support and the intellectually stimulating atmosphere of the Institute.

Bibliography

- Anderson, D. G. 1996. Fluctuations between simple and complex chiefdoms: cycling in the late prehistoric southwest. Pages 231-252 *in* J. F. Scarry, editor. Political structures and change in the prehistoric southeastern United States. University Press of Florida, Gainesville, FL.
- Barfield, T. J. 1989. The perilous frontier: nomadic empires and China. Blackwell, Oxford, UK.
- Barfield, T. J. 1990. Tribe and state relations: the Inner Asian perspective. Pages 153-182 *in* P. S. Khoury and J. Kostiner, editors. Tribes and state formation in the Middle East. University of California Press, Berkeley, CA.
- Barfield, T. J. 1994. The devil's horsemen. Pages 157-182 *in* S. P. Reyna and R. E. Downs, editors. Studying war: anthropological perspectives. Gordon and Breach, Langhorn, PA.
- Barfield, T. J. 2001. The Shadow Empires: Imperial state formation along the Chinese-nomad frontier. Pages 11-41 *in* S. E. Alcock, T. N. D'Altroy, K. D. Morrison, and C. M. Sinopoli, editors. Empires. Cambridge University Press, Cambridge.
- Barfield, T. J. 2002. Turk, Persian, and Arab: Changing Relationships between Tribes and State in Iran and along its Frontiers. *in* N. R. Keddie and R. Matthee, editors. Iran and the Surrounding World: Interactions in Culture and Cultural Politics. University of Washington Press, Seattle.
- Bowles, S. 2006. Group competition, reproductive leveling and the evolution of human altruism. *Science* **314**:1569-1572.
- Bowles, S., and H. Gintis. 2002. Homo reciprocans. *Nature* **415**:125-128.
- Carneiro, R. L. 1970. A theory of the origin of the state. *Science* **169**:733-738.
- Chase-Dunn, C., T. D. Hall, and P. Turchin. 2007. World-systems in the biogeosphere: urbanization, state formation and climate change since the Iron Age. *in* A. Hornborg, editor. World System History and Global Environmental Change. Columbia University Press, New York.
- Coleman, J. S. 1990. Foundations of Social Theory. Belknap Press, Cambridge, MA.
- Collins, R. 1992. Sociological Insight. Oxford University Press, New York.
- Darwin, C. 1871. The Descent of Man, and Selection in Relation to Sex. John Murray, London.
- Durkheim, E. 1915. The elementary forms of the religious life, a study in religious sociology. Macmillan, New York.
- Ember, C. R., and M. Ember. 2001. Cross-Cultural Research Methods. AltaMira Press, Walnut Creek, CA.
- Fehr, E., and S. Gächter. 2002. Altruistic punishment in humans. *Nature* **415**:137-140.
- Fischbacher, U., S. Gächter, and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* **71**:397-404.
- Gibbon, E. 1932. The decline and fall of the Roman Empire. Modern Library, New York.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, and H. Gintis. 2004. Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies. Oxford University Press, New York.
- Hoang, M. 2001. Genghis Khan. Saqi Books.
- Huntington, S. P. 1996. The Clash of Civilizations and the Remaking of World Order. Simon and Schuster, New York.

- Ibn Khaldun. 1958. *The Muqaddimah: An Introduction to History*. Translated from the Arabic by Franz Rosenthal. Pantheon Books, New York.
- Jaspers, K. 1953. *The Origin and Goal of History*. Routledge & Kegan Paul, New York.
- Khazanov, A. M. 1984. *Nomads and the outside world*. Cambridge University Press, Cambridge.
- Korotayev, A. 2006. Urbanization and Political Development of the World System: A comparative quantitative analysis. Pages 115-153 *in* P. Turchin, L. Grinin, V. C. de Munck, and A. Korotayev, editors. *History and Mathematics: Historical Dynamics and Development of Complex Societies*. URSS, Moscow.
- Kradin, N. N. 2000. Nomads, world-empires, and social evolution. Pages 314-336 *in* N. N. Kradin, A. V. Korotayev, D. M. Bondarenko, and V. A. Lynshi, editors. *Alternative routes to civilization (in Russian)*. Logos, Moscow.
- Kradin, N. N. 2002. *The Hunnu Empire (in Russian)*. Logos, Moscow.
- Kradin, N. N. 2005. From tribal confederation to empire: the evolution of the Rouran society. *Acta Orientalia Academiae Scientiarum Hungaricae* **58**:149-169.
- Kraus, J. S. 1993. *The Limits of Hobbesian Contractarianism*. Cambridge University Press, Cambridge, UK.
- Lattimore, O. 1967. *Studies in Frontier History: collected Papers 1928-1958*. Mouton, Paris.
- Liu, L., X. Chen, Y. K. Lee, H. Wright, and A. Rosen. 2002. Settlement patterns and development of social complexity in the Yiluo Region, North China. *Journal of Field Archaeology* **29**:75-100.
- Mair, V. 2005. The North(west)ern Peoples and the Recurrent Origins of the "Chinese" State. Pages 46-84 *in* J. A. Fogel, editor. *The Teleology of the Modern Nation-State: Japan and China*. University of Pennsylvania Press, Philadelphia.
- Mann, M. 1986. *The sources of social power. I. A history of power from the beginning to A.D. 1760*. Cambridge University Press, Cambridge, UK.
- Marcus, J. 1998. The peaks and valleys of ancient states: an extension of the dynamic model. Pages 59-94 *in* G. M. Feinman and J. Marcus, editors. *Archaic states*. School of American Research Press, Santa Fe, NM.
- McNeill, W. H. 1963. *The Rise of the West*. New American Library, New York.
- Moses, L. 1987. The quarelling sons in the Secret History of the Mongols. *Journal of American Folklore* **100**:63-68.
- Olson, M. 2000. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. Basic Books, NY.
- Richerson, P. J., and R. Boyd. 1998. The evolution of human ultrasociality. *in* I. Eibl-Eibesfeldt and F. K. Salter, editors. *Ethnic Conflict and Indoctrination*. Berghahn Books, Oxford.
- Sanderson, S. K. 1999. *Social Transformations: a General Theory for Historical Development*. Rowman and Littlefield, Lanham, MD.
- Service, E. R. 1975. *Origins of the State and Civilization: The Process of Cultural Evolution*. Norton, New York.
- Sober, E., and D. S. Wilson. 1991. *Unto Others: the evolution and psychology of unselfish behavior*. Harvard University Press, Cambridge, MA.
- Stearns, P. N. 2001. *The encyclopedia of world history, 6th edition*. Houghton Mifflin, Boston.

- Taagepera, R. 1978a. Size and duration of empires: systematics of size. *Social Science Research* **7**:108-127.
- Taagepera, R. 1978b. Size and duration of empires; growth-decline curves, 3000 to 600 BC. *Social Science Research* **7**:180-196.
- Taagepera, R. 1979. Size and duration of empires: growth-decline curves, 600 B.C. to 600 A.D. *Social Science History* **3**:115-138.
- Taagepera, R. 1997. Expansion and contraction patterns of large polities: context for Russia. *International Studies Quarterly* **41**:475-504.
- Tainter, J. A. 1988. *The Collapse of Complex Societies*. Cambridge University Press, Cambridge, UK.
- Toynbee, A. J. 1956. *A study of history*. Oxford University Press, London.
- Turchin, P. 2003. *Historical dynamics: why states rise and fall*. Princeton University Press, Princeton, NJ.
- Turchin, P. 2006. *War and Peace and War: The Life Cycles of Imperial Nations*. Pi Press, NY.
- Turchin, P., J. M. Adams, and T. D. Hall. 2006. East-West Orientation of Historical Empires and Modern States. *Journal of World-Systems Research* **12**:219-229.
- Turchin, P., and S. Nefedov. 2008. *Secular cycles*. Princeton University Press, Princeton, NJ.
- Webster, D. 1975. Warfare and the Evolution of the State: A Reconsideration. *American Antiquity* **40**:464-470.
- Wilson, D. S., and E. O. Wilson. 2007. Rethinking the theoretical foundations of sociobiology. *Quarterly Review of Biology* **82**:327-348.
- Wilson, E. O. 1975. *Sociobiology: The New Synthesis*. Harvard University Press, Cambridge, MA.
- Wilson, E. O. 1979. *On Human Nature*. Harvard University Press, Cambridge, MA.
- Wright, H. T. 1977. Recent Research on the Origin of the State. *Annual Review of Anthropology* **6**:379-397.

Table 1. Imperial unifications in China from the Shang era to the present. Compiled from various sources: Eugene Anderson, University of California (pers. comm.), Mair (2005), and Stearns (2001). Not all of these states controlled the territory exceeding 1 Mm², and thus only some of them are reflected in Table 2.

<i>Unification</i>	<i>Period</i>	<i>Ethnicity</i>	<i>From</i>	<i>Capital</i>
Shang	1766–1122 BCE	??	North Central (Huang He)	Anyang (Huang-He)
W. Zhou	1122–771 BCE	Frontier Han ("Western Barbarians")	Northwest (Wei River Valley)	Loyang (Huang He)
Qin	221–206 BCE	Frontier Han	Northwest (Wei River Valley)	Xianyang (Wei)
Han	202 BCE–220	Han	Northwest (confluence of Wei and Huang)	Chang'an (Wei)
W. Jin	280–316	Han	North Central (Huang He)	Loyang (Huang He)
N. Wei (partial, N)	386–534	To-ba (Mongolian)	Northwest	Loyang (Huang He)
Sui	581–618	Han	Northwest (Wei River Valley)	Chang'an (Wei)
Tang	618–907	Han (ruling family of Turkic descent)	Northwest (Wei River Valley)	Chang'an (Wei)
Liao (partial, N)	907–1125	Kitan (Altaic?)	Northeast S. Manchuria	Beijing
N. Sung (partial, w/o N)	960–1127	Han	North Central (lower Huang He, around Kaifeng)	Kaifeng (Huang He)
Jin (partial, N)	1115–1234	Jurchen (Tungus)	Northeast Manchuria	Beijing
Yuan	1206–1368	Mongol	Northwest (Mongolia)	Beijing
Ming	1368–1644	Han	South Central (from Nanjing area)	Beijing
Qing	1644–1911	Manchu (Tungus)	Northeast (Manchuria)	Beijing
Communist	1949–	Han	Northwest (Long March to Wei River Valley; unification from there)	Beijing

Table 2. Megaempires in the historical record

Date (peak)	Empire name	World Region	Area (Mm²)
-1300	Egypt (New Kingdom)	Africa	1.00
350	Axum	Africa	1.25
969	Fatimid	Africa	4.10
1120	Almoravid	Africa	1.00
1200	Almohad	Africa	2.00
1380	Mali	Africa	1.10
1400	Mameluk	Africa	2.10
1527	Inca	America	2.00
-176	Hsiung-Nu	Central Asia	9.00
405	Juan-Juan	Central Asia	2.80
557	Turks	Central Asia	6.00
800	Uigur	Central Asia	3.10
800	Tufan (Tibet)	Central Asia	4.60
850	Khazar	Central Asia	3.00
1100	Hsi-Hsia	Central Asia	1.00
1210	Khorezm	Central Asia	2.30
1210	Kara-Khitai	Central Asia	1.50
1270	Mongol	Central Asia	24.00
1310	Golden Horde	Central Asia	6.00
1350	Chagatai	Central Asia	3.50
1405	Timurid	Central Asia	4.40
-1122	Shang	East Asia	1.25
-50	China-Early Han	East Asia	6.00
579	Liang	East Asia	1.30
715	China-Tang	East Asia	5.40
947	Liao (Kitan)	East Asia	2.60
980	China-Sung	East Asia	3.10
1126	Chin (Jurchen)	East Asia	2.30
1450	China-Ming	East Asia	6.50
1790	China-Manchu	East Asia	14.70
117	Rome	Europe	5.00
441	Huns (Atilla's)	Europe	4.00
555	East Roman	Europe	2.70
814	Frankish	Europe	1.20
1000	Kiev	Europe	2.10
1025	Byzantine	Europe	1.35
1480	Lithuania-Poland	Europe	1.10
1683	Ottoman	Europe	5.20
1800	Russia	Europe	15.50
1290	Khmer	Southeast Asia	1.00
-250	Mauryan	South Asia	5.00

200	Kushan	South Asia	2.00
400	Gupta	South Asia	3.50
648	Harsha (Kanyakubia)	South Asia	1.00
1312	Delhi	South Asia	3.20
1690	Mughal	South Asia	4.00
1760	Maratha	South Asia	2.50
-670	Assyria	Southwest Asia	1.40
-585	Media	Southwest Asia	2.80
-500	Achaemenid Persia	Southwest Asia	5.50
-323	Hellenistic (Alexander's)	Southwest Asia	5.20
-301	Seleucid	Southwest Asia	3.90
0	Parthia	Southwest Asia	2.80
550	Sassanian Persia	Southwest Asia	3.50
750	Caliphate	Southwest Asia	11.10
928	Samanid	Southwest Asia	2.85
980	Buyid (Buwahid)	Southwest Asia	1.60
1029	Ghaznavid	Southwest Asia	3.40
1080	Seljuk	Southwest Asia	3.90
1190	Ayyubids	Southwest Asia	2.00
1310	Il-Khanid	Southwest Asia	3.75

Figure 1 (on next page). Spatial distribution of megaempires within Afroeurasia. The approximate locations of imperial centers are indicated with codes related to empire names in Table 2. These locations are especially approximate in areas which were “hotspots of imperiogenesis” (e.g., East Asia), where there simply was not space to fit all names appropriately. The colors indicate the distribution of major types of ecologic communities. Of particular interest are the yellow (temperate grassland/desert) and light brown (subtropical desert).

KEY			
 Tropical rain forest	 Woodland /shrubland	 Subtropical desert	 Tundra
 Tropical seasonal forest/ savannah	 Temperate grassland/ desert	 Temperate rain forest	 Alpine
	 Boreal forest	 Temperate seasonal forest	 Polar ice cap



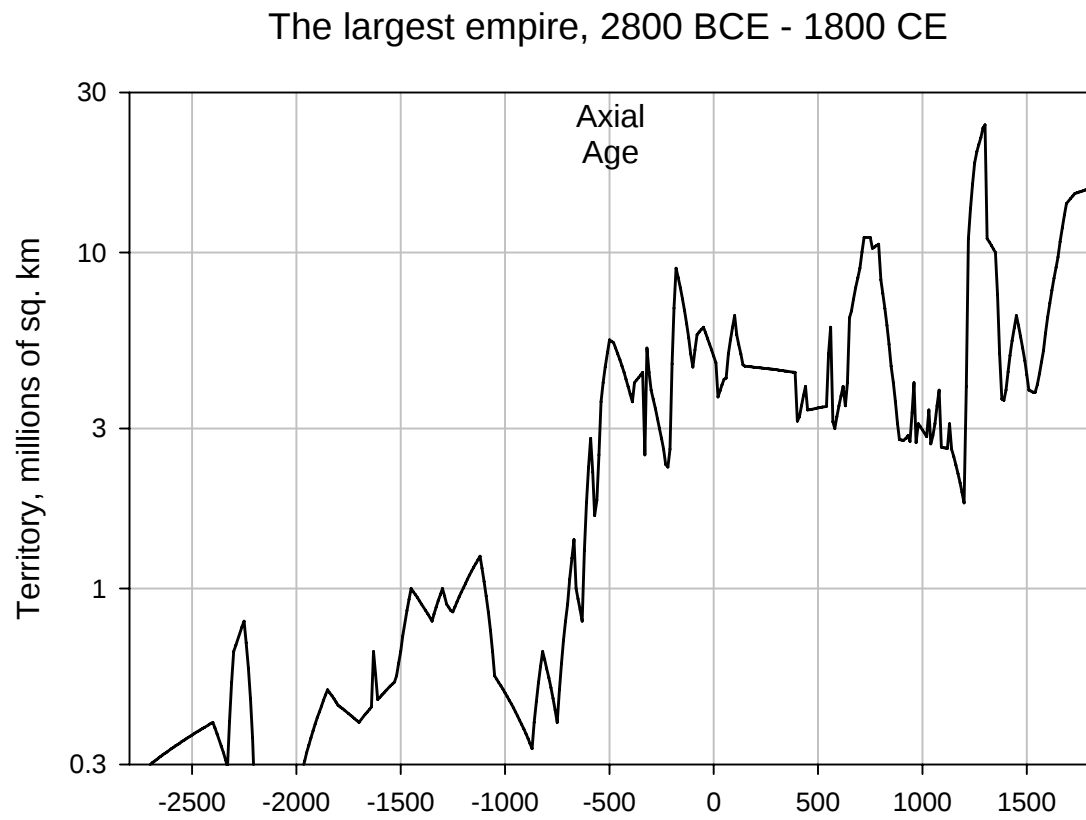


Figure 2. The area of the largest empire between 2800 BCE and 1800 CE. Note the log-scale for territorial size.