Fractured-Land and Political Fragmentation

Jesús Fernández-Villaverde, Mark Koyama, Youhong Lin, Tuan-Hwee Sng*

July 3, 2020

Abstract

Patterns of political unification and fragmentation have crucial implications for comparative economic development. Diamond (1997) famously argued that "fractured-land" was responsible for China's tendency towards political unification and Europe's protracted political fragmentation. We build a model with granular geographical information to quantitatively gauge the effects of "fractured-land" on state formation in Eurasia. We find that a broad version of the "fractured-land" hypothesis that takes into account both topographical features and the location of productive agricultural land is necessary and sufficient, within our model, to account for China's recurring political unification and Europe's persistent political fragmentation. In particular, the existence of a core region of high land productivity in Northern China plays a central role in our simulations. We discuss how our results map into observed historical outcomes and assess how robust our findings are.

Keywords: China; Europe; Great Divergence; State Capacity; Political Fragmentation; Political Centralization

JEL Codes: H56; N40; P48

*Jesús Fernández-Villaverde, Department of Economics, University of Pennsylvania. Email: jesusfv@econ.upenn.edu. Mark Koyama, Department of Economics, George Mason University. Email: mkoyama2@gmu.edu. Youhong Lin, Center for Cliometrics Studies of China, Guangdong University of Foreign Studies. Email: lin.youhong@foxmail.com. Tuan-Hwee Sng (corresponding author), Department of Economics, National University of Singapore. Email: tsng@nus.edu.sg. We are grateful for comments from seminar audiences at Brown University, Toulouse School of Economics, UC Irvine, Vancouver School of Economics, and from discussions with Siwan Anderson, Emmanuelle Auriol, Lisa Blaydes, Jean-Paul Carvalho, Latika Chaudhary, Dan Bogart, Oded Galor, Saum Jha, Nippe Lagerlöf, Gary Richardson, Mohamed Saleh, Paul Seabright, Stergios Skaperdas, and Felipe Valencia.

Here begins our tale. The empire, long divided, must unite; long united, must divide. Thus it has ever been.

Romance of the Three Kingdoms, Chapter 1.

1 Introduction

The economic rise of western Europe is often attributed to its political fragmentation. See, among many others, Jones (2003), Mokyr (2016), and Scheidel (2019). In this reading of the historical record, a persistently polycentric and competitive state system created incentives to innovate and invest in state capacity and, thus, laid the foundations of Europe's institutional development. Correspondingly, many explanations of China's comparative failure to achieve sustained economic growth focus on its long history as a centralized empire and the barriers to riches that such a centralization induced. But what factors account for the prevalence of political fragmentation in Europe and the prominence of political centralization in China? Therefore, the answer to this question is fundamental for our understanding of the onset of modern economic growth.

Researchers have proposed numerous mechanisms for the divergence in political fragmentation across the two extremes of the Euroasia landmass. A particularly popular mechanism, offered by Diamond (1997, 1998), argues that "fractured-land" such as mountain barriers, dense forests, and rugged terrain impeded the development of large empires in Europe in comparison to other parts of Eurasia.

However, the "fractured-land" hypothesis is not without its critics. For instance, Hoffman (2015b) points out that China is, in fact, more mountainous than Europe. Peter Turchin and Tanner Greer have advanced similar arguments in a blog form.[1] Turchin goes as far as defending the claim that it is not Europe's fragmentation that needs explanation, but China's precocious and persistent unification.

The "fractured-land" hypothesis has also been challenged for being static and overly deterministic. Hui (2005, 1) contests the idea that China was "destined to have authoritarian rule under a unified empire," while contending that Europe's political fragmentation was a

[1] See, for details, http://peterturchin.com/cliodynamica/why-europe-is-not-china/ and http://scholars-stage.blogspot.com/2013/06/geography-and-chinese-history-fractured.html.

highly contingent outcome. After all, China has not always been unified. As the opening lines from *The Romance of Three Kingdoms* above remind us, China has experienced long periods of fragmentation throughout its history. Besides, the degree of political fragmentation in Europe has varied considerably over time.[2]

This paper provides a quantitative investigation of the "fractured-land" hypothesis and the criticisms that it has received.[3] To do so, we explicitly model the dynamic process of state-building over time to explore if, and how, fractured-land shaped and interacted with inter-state competition. Making use of rich data on topography, climate, and land productivity, we simulate this model at a fine grid-cell geographical level and look at the resulting probability distribution of political structures. Obtaining probability distributions over historical outcomes is particularly relevant for our investigation because they gauge the role of structure versus contingency in shaping observed degrees of political unification.

Our main finding is that "fractured-land" indeed provides a robust explanation for the political divergence observed at the two ends of Eurasia: a unified China and a fragmented Europe. In addition, our model allows us to distinguish between two versions of the "fractured-land" hypothesis. First, in a narrow sense, scholars have equated the fractured-land with the presence of mountainous and rugged topography. Second, a broader definition of fractured-land takes into account the location of productive agricultural land.

We document that topography alone is sufficient, but not necessary, to explain political fragmentation in Europe and unification in China. The location of Europe's mountain ranges ensured that there were several distinct geographical cores of equal size that could provide the nuclei for future European states, whereas China was dominated by a single vast plain between the Yangtze and the Yellow River. But the presence of a dominant core region of high land productivity in China –in the form of the North China Plain– and the lack thereof in Europe can also explain political unification in China and division in Europe.

In our simulation exercise, it is only when we neutralize the effects of fractured-land in the broad sense that Europe and China cease to move at different paces toward political

[2]There is, as well, a subtle question about how we measure political fragmentation before the rise of the modern nation-state. Can we consider the Holy Roman Empire as a unified polity? Under Otto I (r. 962–973), perhaps yes. Under Francis I (r. 1745–1765), most likely no.

[3]Other tests of various elements of Diamond's hypothesis include Turchin et al. (2006) and Laitin et al. (2012).

unification. Our model, therefore, suggests that broad geographical features that went beyond ruggedness were indeed crucial to understand why China unified and Europe remained fragmented. Specifically, our analysis highlights the importance of having core geographical regions of high land productivity unbroken by major mountain, desert, or sea barriers.

A battery of robustness tests assesses how our quantitative results depend on the assumptions and calibration of the model. A summary of those tests is that they confirm the key role of fractured-land in the broad sense.

Our analysis contributes to several literatures in political economy and economic history. First, we complement a long-standing literature that attributes the rise of western Europe to its multi-state system by investigating the *causes* of Europe's political fragmentation. Without being exhaustive, the literature includes Montesquieu (1989), Pirenne (1925), Hicks (1969), Jones (2003), Hall (1985), Rosenberg and Birdzell (1986), Baechler (1975), Cowen (1990), Tilly (1990), Chaudhry and Garner (2006), Mokyr (2007), Karayalcin (2008), Chu (2010), Olsson and Hansson (2011), and Lagerlöf (2014).

Second, we add to the literature on state formation in Europe and China. One element within this literature emphasizes the importance of the invasion threat from the steppe in Chinese state development (Lattimore, 1940; Grousset, 1970; Huang, 1988; Barfield, 1989; Gat, 2006; Turchin, 2009; Chen, 2015; Ko et al., 2018). Another strand emphasizes the importance of war and military competition in the formation of European states (Parker, 1988; Tilly, 1990; Downing, 1992; Voigtländer and Voth, 2013; Gennaioli and Voth, 2015). We develop an argument that emphasizes the interaction between geographical fractionalization, agricultural productivity, and military competition.

However, in our model, we ignore the role of improvement in military technology through inter-state competition highlighted by Hoffman (2015b). According to Hoffman, this tournament eradicated polities that were unable to compete and led to an acceleration in military and political technologies, which evolved comparatively rapidly as a result of intensive learning-by-doing. Our results show that such an improvement is not necessary to account for the comparative political structures of Europe and China. However, in a richer model than the one we handle here, the political-military tournament could be a complement to the "fractured-land" hypothesis.

Our investigation also relates to two papers that examine the causal link between geography and state fragmentation. Turchin et al. (2013) study the formation of empires in Eurasia. These

authors argue that the intensification of warfare –a process heavily influenced by proximity to the Eurasian steppe and the subsequent antagonistic relations between the nomadic steppe and settled agriculturalists– favored the evolution of ultrasocial traits and the rise of large-scale states by putting pressure on premodern polities forcing them to strengthen and to invest in state capacity as a defensive response. Kitamura and Lagerlöf (2019) find that mountain ranges and rivers have an influence on the location of political boundaries in and around Europe. We build on these studies and focus on the roles of topography and land productivity as sufficient conditions to explain state formation at both ends of Eurasia.

One novel feature of our model is the role played by agricultural productivity in giving rise to conflict and thereby prompting larger states to coalesce. In this respect, our analysis shares some similarities with Acharya and Lee (2018), who develop a model in which economic development generates rents that lead to the formation of territorial states.

Finally, we contribute to the literature on the relationship between geography and economic and political outcomes. Geography can shape economic outcomes directly, for instance via access to trade routes or vulnerability to disease vectors (i.e., Sachs, 2001) or indirectly via its effect on political institutions (e.g., Acemoglu et al., 2001, 2002, 2005). We provide an example of the latter phenomenon: geography mattered in Chinese and European history because it gave rise to a centralized state in China and resulted in fragmentation in Europe.

Notice that while our findings speak to the literature on the origins of sustained economic growth (Galor and Weil, 2000; Galor, 2005, 2011), we do not investigate how centralization and fragmentation drove or retarded long-run growth. A reader can accept our conclusion that the "fractured-land" hypothesis is quantitatively sound without embracing the idea that a polycentric state system was behind the great divergence between Europe and China.

The remainder of the paper is organized as follows. Section 2 outlines the main arguments advanced in favor of, and against, geographical explanations for observed patterns of political unification and fragmentation in China and Europe. Section 3 considers European and Chinese geography and, motivated by it, builds a model of conflict and inter-state competition that explicitly integrates geographical characteristics. Section 4 calibrates the model and Section 5 presents the quantitative results. Section 6 discusses some extensions of the model. Section 7 discusses some aspects of European and Chinese history in light of our model. We conclude in Section 8.

2 Fractured-Land?

Anthropologists, geographers, historians, and sociologists have long suggested that early states could only form where there was a sufficiently large area of productive agricultural land. The land had to be productive enough (in per-unit terms times the size of the area) to generate a food surplus, which was needed to feed and clothe a political elite and its bureaucracy. This output also needed to be appropriable (Mayshar et al., 2016). Furthermore, the land tended to require geographical boundaries that made it possible for political authorities to coerce the population into transferring these surpluses to the political elite (e.g., Carneiro, 1970). Indeed, agrarian states struggled to project power into rugged, hilly, or mountainous lands where such coercion was too costly (Mayshar et al., 2017; Scott, 2017).

Based on these ideas, geographers use the concept of a geographical core to describe the nucleus of successful states (Whittlesey, 1944; Pounds and Ball, 1964; Hechter and Brustein, 1980). Geographers argue that the cores of most successful states were based around self-contained geographical regions characterized by areas of fertile agricultural land with good transport connections and defensible from external invasion.[4] Thus, regions of the world with many potential geographical cores were bound to see an earlier and faster unification into states.

A direct consequence of this argument is that, since Europe's geographical features were less favorable to the formation of these early states, the posterior history of the continent, plagued by fragmentation, was decisively determined by its topographical peculiarities. Many authors (e.g., Jones, 2003; Kennedy, 1987) have made this argument, but Jared Diamond's formulation of this "fractured-land" hypothesis (Diamond, 1997, 1998) is, perhaps, the most influential among scholars.[5]

[4] Thus: "[t]he geographical pattern of the states of Europe had, in general, taken shape before the age of modern nationalism ... However profoundly they may have been modified and their expansion influenced by the forces which make up modern nationalism, most European states grew in fact by a process of accretion from germinal areas which have come, after Derwent Whittlesey, to be called 'core-areas"' (Pounds and Ball, 1964, 24).

[5] Jones (2003, 226) notes that the "topographical structure of" Europe, "its mountain chains, coasts and major marshes, formed boundaries at which states expanding from the core-areas could meet and pause" and that "these natural barriers helped to hold the ring between the varied ethnic and linguistic groups making up the European peoples." Kennedy (1987, 17) similarly states that Europe's political diversity was "largely" due to its geography: "There were no enormous plains over which an empire of horsemen could impose its swift dominion; nor were there broad and fertile river zones, like those around the Ganges, Nile, Tigris and Euphrates, Yellow and Yangtze, providing the food for masses of toiling and easily conquerable peasants. Europe's landscape was much more fractured, with mountain ranges and large forests separating the scattered population centers in the valleys; and its climate altered considerably from north to south and west to east."

Diamond makes the following observations: (1) China was not threatened by the presence of large islands off the mainland of China (Taiwan was too small and Japan too far away); (2) the Chinese coastline was smooth compared to the European coastline; (3) most importantly, unlike Europe, China was not fractured by high mountains and dense forests.[6]

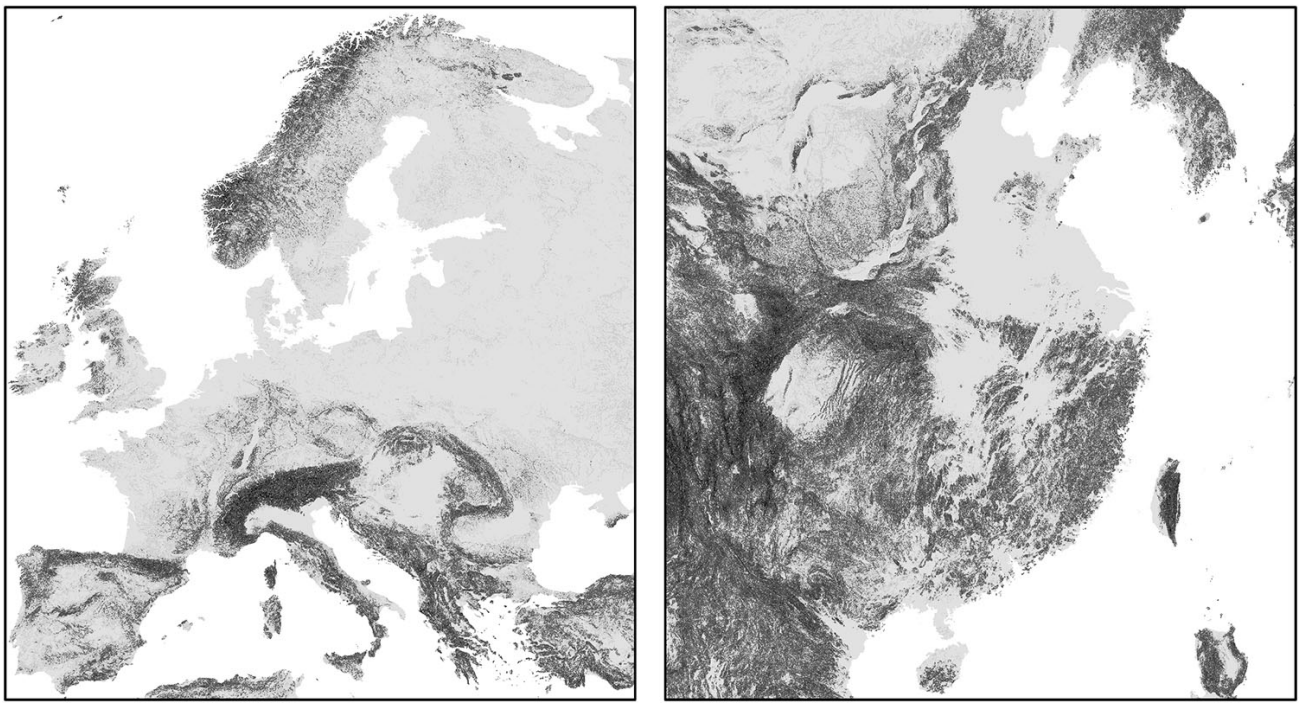


Figure 1: Ruggedness in Europe and China proper.

The claims of the fractured-land hypothesis have come under heavy criticism. Hoffman (2015b, 109–112) observes that, in fact, China is significantly more mountainous than Europe. He notes that over 37% of modern China is defined as mountainous in comparison to little more than 10% of Europe. Even if one restricts attention to the so-called China proper, i.e., the traditionally agrarian part of China south of the Great Wall and east of the Tibetan Plateau, more than 33% was elevated above 1,000m compared to only around 6% in Europe. A simple

[6] In his own words, Diamond (1998, 433), "[...] the ultimate reason for Europe's political fragmentation emerges from a glance at a map of Europe [...] Seas, a highly indented coastline, high mountains and dense forests divide Europe into many peninsulas, islands and geographical regions, each of which developed political, linguistic, ethnic and cultural autonomy. Each such region became one more natural experiment in the evolution of technology and scientific inquiry, competing against other regions. Conversely, China has a much less indented coastline, no islands large enough to achieve autonomy, and less formidable internal mountain barriers. (Even China's two largest islands, Hainan and Taiwan are small: each has less than half the area of Ireland; neither was a major independent power until Taiwan's emergence in recent decades; and, until recently, Japan's geographical isolation kept it much more remote politically from the Asian mainland than Britain has been from mainland Europe.) China was linked from east to west by two parallel, long and navigable rivers, and was eventually linked from north to south by canals between those rivers. So once a unified Chinese state was founded, geography prevented any other state from gaining lasting autonomy in any part of China."

inspection of Figure 1 reveals that, indeed, Hoffman is right: China is more mountainous than Europe.

However, this criticism forgets that the crucial factor might not be the ruggedness of the terrain in general, but the exact location of either continent's mountainous regions. Beyond the total amount of ruggedness, Figure 1 illustrates that mountain ranges at or near the center of Western Europe play an important role in separating Italy and Spain from France and making core regions of central Europe (Switzerland, Austria) difficult to conquer.[7]

Moreover, much of the Northern European plain was historically covered with dense forest, which impeded conquest and the rapid movement of large armies. The critical role of the northern Europe forest in deterring Roman expansion is attested in the historical sources including Julius Caesar, Pliny the Elder, and Tacitus.[8]

Consequently, Europe comprises several cores: the British Isles, Scandinavia, the Iberian peninsula, and the Italian peninsula all form distinctive and discontinuous "regions" that stand out from any simple visual inspection of a map of Europe. The modern countries of France, the Low Countries, Germany, and Poland span what is known as the Northern European plain. The easternmost region part of this plain borders the Russian forest in the north-east, the steppe in the east, and the Carpathian mountains in the south; it corresponds loosely to modern Poland and the territory controlled by the Polish-Lithuanian Commonwealth in the early modern period. The central part of the plain corresponds to the modern country of Germany, while the western part of the plain is in what we call today France.

Meanwhile, the most mountainous regions in China are in the south and west, and they do not intersect the central plain in North China that historically played a crucial role in China's early unification. The Central Plain, which centers on the Yellow River basin, is blocked from

[7]During World War II, Switzerland planned a retreat to a "réduit national' in the central part of Switzerland in case of a German invasion. Similarly, by the end of the war, Germany undertook preliminary preparations to retreat to an Alpine redoubt in Southern Bavaria, Western Austria, and Northern Italy, although such preparations might have played more of a role of misinformation to divert Allied forces toward those areas than of practical use. The 12 Battles of the Isonzo during World War I between the Italian and Austro-Hungarian armies suggests that taking over those redoubts could be extremely costly, even for a modern army.

[8]The Romans "were forced to stop in their expansion and Empire building at the boundaries of the dense, virgin German forests whose inhospitable and somber nature was pictured in dark colors by such ancient writers as Tacitus, Pomponius Mella, and Marcellinus, who spoke of the forests as of something horrid and inaccessible and unsuited for human habitation" (Zon, 1920, 141). See also Begle (1900) and Howorth (1909). Tacitus describes Germania as a land that "bristles with forests or reeks with swamps." He describes various German tribes, the "Reudignians, and Aviones, and Angles, and Varinians, and Eudoses, and Suardones and Nuithones" as "all defended by rivers or forests" (Tacitus, 1877, 90 and 116).

Korea in the northeast by the Changbai Mountains and the Taihang Mountains in the west. The plain itself is flat, except for the Taishan mountains in Shandong and Dabie Mountains of Anhui. Southern China is more mountainous than the central Chinese plain. The Yunnan-Guizhou Plateau has particularly high elevation. Mountains and then dense forest divided Lingnan and Yunnan from Vietnam and Burma, respectively. Diamond (1997, 414) himself emphasizes the existence of a large core region capable of dominating the other regions in China:

> "China's heartland is bound together from east to west by two long navigable river systems in rich alluvial valleys (the Yangtze and Yellow Rivers), and it is joined from north to south by relatively easy connections between these two river systems (eventually linked by canals). As a result, China very early became dominated by two huge geographic core areas of high productivity, themselves only weakly separated from each other and eventually fused into a single core."

The arguments in the previous pages are qualitative. And, as such, they cannot be assessed quantitatively (e.g., *how* rough must the terrain be to make a difference for political unification?) or used to measure the role of structure versus contingency in the observed outcomes (perhaps China's early leaders were just luckier or have better leadership skills than their European counterparts?).

Can we bring quantitative data and a simple model of state formation and competition to the table and formally evaluate the "fractured-land" hypothesis and the range of distributions of probability that it can span? The next section introduces such a model.

3 Model

We now discuss our model. First, we describe the geographical space of our study. Second, we explain how we divide the space into hexagonal cells. Third, we introduce the geographic, climatic, and resource availability characteristics that index those cells. Fourth, we discuss how the size of polities evolves through conflict and secession.

3.1 The Geographical Space

Our geographical space of study includes most of Europe, North Africa, the Middle East, Continental Asia, and Japan. This area, plotted in Figure 2, is called –with different degrees of precision– the "Old World," Eurasia, and the Afro-Eurasian ecumene.

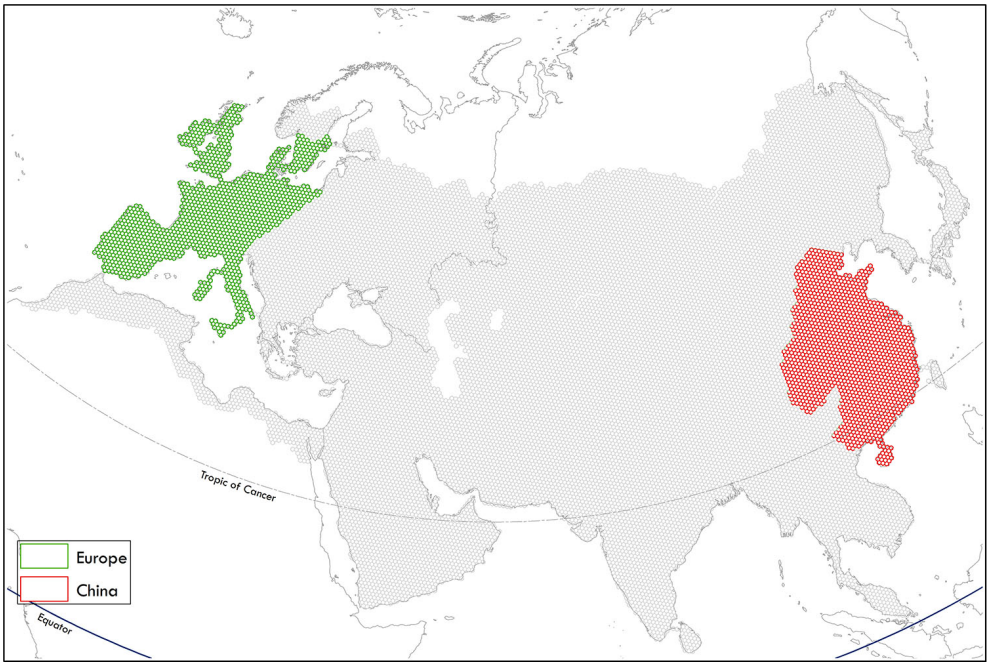


Figure 2: Study Area.

We consider this space our unit of study because, after the beginning of the Iron Age (c. 1200–1000 BCE), intense political, trade, and cultural contacts across the Afro-Eurasian ecumene started flourishing quickly.[9] As the renowned historian Marshall Hodgson (1922–1968) put it, our area of study corresponds to:

> "...the various lands of urbanized, literate, civilization in the Eastern Hemisphere, in a continuous zone from the Atlantic to the Pacific, [that] have been in commercial and commonly in intellectual contact with each other, mediately or immediately" (Hodgson, 1954, 716).

What phenomena did Hodgson have in mind? For instance, the Roman Empire and Han China traded indirectly and knew of each other's existence. A Roman delegation visited China

[9]The Iron Age onsets at slightly different times over Eurasia, with the earliest transitions in the Middle East and the latest in Northern Europe. For compactness of exposition, we will ignore such heterogeneity.

in 166 and the Chinese historian Yu Huan wrote a description of the Roman Empire –named *Daqin* by his contemporaries– sometime between 239 and 265. Roman commerce with the Indian subcontinent was lively, with the tariffs on it accounting perhaps for as much as one-third of the Empire's revenue (McLaughlin, 2010). Roman coins made their way to Japan and Buddhism had a presence in Rome.

In comparison, the Afro-Eurasian landmass regions that we ignore (northern parts of Scandinavia and Russia, Sub-Saharan Africa, etc.) were either excluded from the above-referenced networks of exchange due to geographical barriers or too thinly populated due to environmental constraints. Beyond minor interactions (the Mali Empire, Arab seamen in the East Coast of Africa, the Vikings in North America), Sub-Saharan Africa, the Americas, and Oceania developed independently from our area of study until the Age of Exploration starting in the second half of 15th century. Thus, after the 15th century, a different geographical space would be required, even if the consequences of the political structure at the end of the period of study are likely to have persisted to the present day.

Importantly, the space of study in Figure 2 has accumulated, for most of history, the majority of world human population and has been the origin of much the developments in technology and social and political forms (Kremer, 1993; Diamond, 1997). For instance, by the year 1500, around 85 percent of the world population lived in this space.[10] Understanding the dynamics of the political forms that evolved in this space is, therefore, critical for global economic history.

3.2 Dividing the Geographical Space

We divide the space of study in Figure 2 into 20,637 hexagonal cells of radius 28 kilometers, each capable of sustaining a polity and allowing armies to pass through it. This radius corresponds to the distance that a healthy adult travels by foot per day on flat terrain.[11] As such, a 28-kilometers hexagon roughly represents the surface that the simplest polities can monitor and

[10] We build this estimate from Table B-12 in Maddison (2001). The data does not exactly fit with our area of study. For example, Africa is not divided between North and Sub-Saharan Africa. However, Africa's total population in 1500 was 10.5 percent of the world population. Any reasonable breakdown of Africa's population between North and Sub-Saharan shares will give us roughly equivalent total shares for the area with which we are concerned with.

[11] This distance assumes a 7 hours march at a leisurely pace of 4 km/h. In Roman times, recruits were required to complete about 30 km in 6 hours in loaded marches. In the U.S. Army, the average march rate for foot soldiers is estimated to be between 20 to 30 km per day. See Headquarters, Department of the Army (2017, Figure 1-2).

defend with rudimentary Bronze Age technologies.

Of these cells, 1,434 cells cover "China," defined as the lands south of the Great Wall (cells with red borders in Figure 2). This region corresponds to the historical core of Imperial China until the Qing expansion to the West (Perdue, 2005). Another 1,307 cells, highlighted with green borders in Figure 2 are in (Western) "Europe," defined as the lands west of the Hajnal line running from Saint Petersburg to Trieste and delimiting the region of the so-called *European marriage pattern* (Hajnal, 1965). This marriage pattern is important because many historians have used it as a proxy for close cultural and social similarities of the loosely-called "western world."

3.3 Geographical, Climatic, and Resource Availability Characteristics

Our next step is to include variables measuring, in each cell, a vector of geography and climate characteristics, $\mathbf{x}$, and historical resource availability, y. Such characteristics will allow us to test Diamond's hypothesis that geographical features are central for the likelihood that we observe regional clustering of cells into empires.

In terms of geographical and climatic characteristics, we will consider, first, terrain ruggedness, x_{rugged}. We measure ruggedness by the average standard deviation of elevation, an index of topographic heterogeneity. Both plains and plateaus score low in this measure, while mountain ranges and valleys score high (Nunn and Puga, 2012). Figure 3 depicts x_{rugged}. There, we can see the high ruggedness of the Alps, the Balkans, Caucasus, and Himalayans and the low ruggedness of the Northern European plain, much of Russia, the Indian subcontinent, and North China.

Second, we assess whether the cell is part of a sea channel, x_{sea}. This variable aims at incorporating the "stopping power of water" for military conquest (Mearscheimer, 2001, 84). Seas constitute major barriers that impede the spread and expansion of states. An invasion across a sea channel is militarily risky and logistically challenging. In Figure 4, we represent the sea channels as dark blue cells ($x_{sea} = 1$). These cells include, for instance, the English Channel, the Sound, the Bosporus, and the Taiwan Strait.

Third, we will measure whether the cell is frigid, x_{frigid}. We classify a cell as frigid ($x_{frigid} = 1$) if it had an annual temperature below 0 degree celsius for six months or more during the Holocene epoch (8,000 BCE), which was relatively warm in historical context. Fourth, we check torridness,

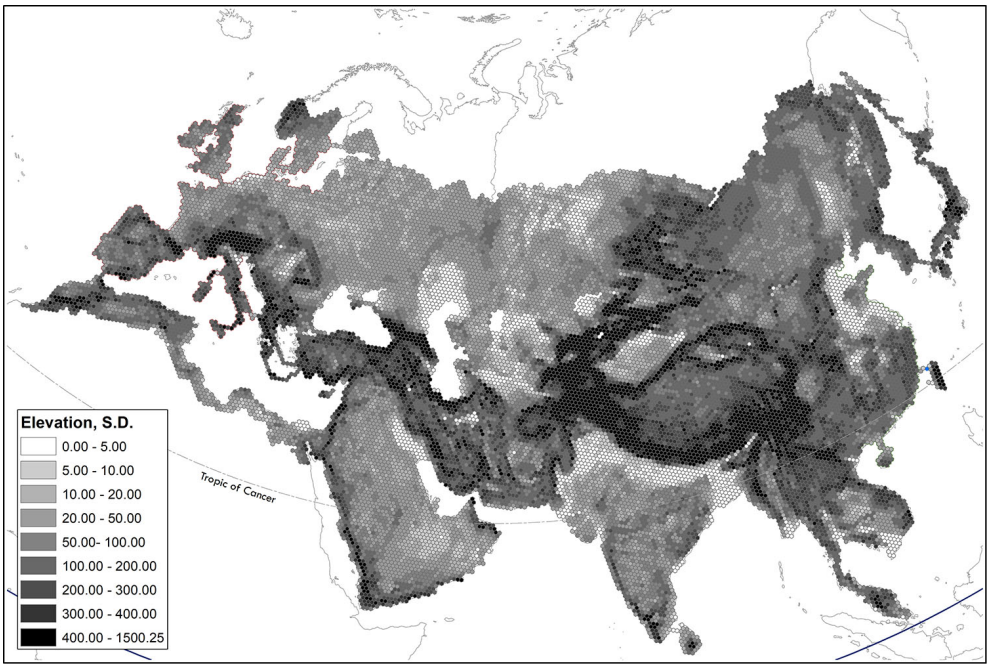


Figure 3: Ruggedness (Standard Deviation of Elevation).

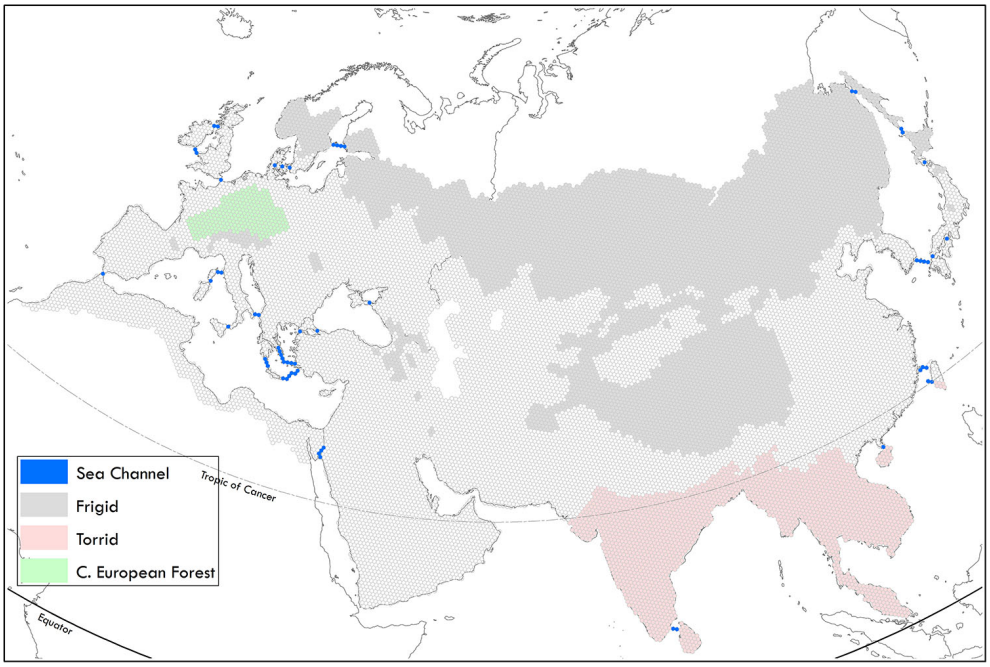


Figure 4: Auxiliary Barriers to Conquest.

with $x_{torrid} = 1$ if the cell is in the tropical or torrid zone based on the Köppen climate classification.

The terms "frigid" and "torrid" are borrowed from Aristotle. In Figure 4, the frigid climate

variable is depicted in light gray where $x_{frigid} = 1$. Most of these cells are in the northern frontier of our area of study or mountainous regions (the Himalayas, the Alps, the Caucasus). Similarly, the cells where $x_{torrid} = 1$ appear in pink. They are mostly clustered in the Indian Subcontinent and Southeast Asia.

Fifth, we gauge whether, according to Goldewijk et al. (2017), the cell was part of the ancient forests of central and northern Europe in 0 CE, with $x_{forest} = 1$ if it was. The role of forests requires a more detailed discussion. Historically, forests hindered the expansion of human activities (Zon, 1920). The ancient forests of Germany inhibited the expansion of the Roman empire into northern Europe, with the complete loss of three Roman Legions at the Battle of the Teutoburg Forest in 9 CE as a paradigmatic example of such difficulties.[12] Centuries later, the forests of Bohemia, Silesia, and Pomerania rendered similar protection to the German states against the Magyars and other invaders from the east, but also retarded state consolidation among the German states. Most recently, covering hills and mountains prevented the American colonists from reaching the crest of the Appalachians for some 200 years.

Central Europe, with its frost-free summers and plentiful rainfall throughout the year, is distinctly "forest-friendly" (Leuschner and Ellenberg, 2017). Furthermore, unlike forests in North America and East Asia, which retreated southwards during the glacial periods and returned north once the glaciers subsided, European forests were blocked from moving south by the Mediterranean and the east-west running mountains such as the Alps and the Pyrenees during glaciations. Consequently, many of the cold-sensitive tree species disappeared altogether. Compared with trees in North America and East Asia, those that made up the central European forests were limited in variety and were dominated by hardwood species such as oaks, beeches, and birches, which could not be cleared with primitive tools (Huntley and Birks, 1983). 90% of Germany was covered by forests in 750 (Kowarik and Körner, 2005). As late as 1700, about 40% of Germany remained forested (Wilson, 2012). By contrast, the loess region of northern and central China, where Chinese civilization first developed, was sparsely forested and largely covered by grass in the last 20,000 years (Jiang et al., 2013).

Thus, the vector $\mathbf{x}$:

$$\mathbf{x} = \{x_{rugged}, x_{sea}, x_{frigid}, x_{torrid}, x_{forest}\} \tag{1}$$

[12] From now on, we will omit "CE" when it is obvious from the context of the text.

includes one continuous geographical variable, x_{rugged}, and four binary ones. We will refer to these last four binary variables as auxiliary barriers to conquest. While we could consider other geographical and climatic characteristics, we found that these five variables were the most relevant to our analysis. We will return to this point later on.

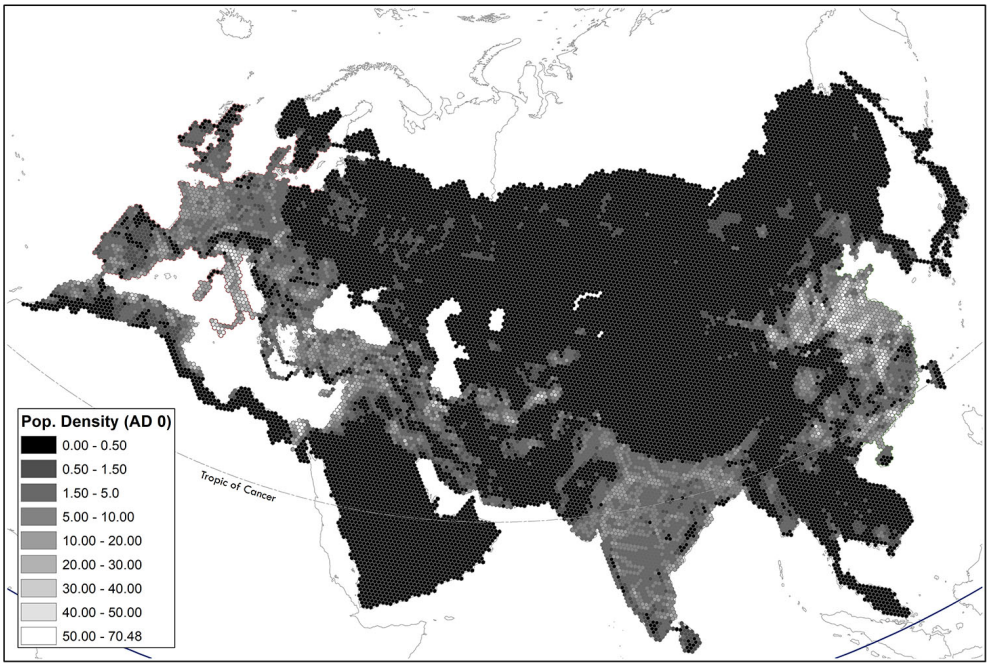


Figure 5: Population Density (0 CE).

We measure historical resource availability by population density in 0 CE as reported by Goldewijk et al. (2017). This measure is motivated by a simple Malthusian logic: before the Industrial Revolution, population density was directly linked to land productivity and its ability to support dense populations (Ashraf and Galor, 2011). This productivity could be either natural, e.g., a good climate, or human-made, e.g., the mastery of the cultivation of a cereal such as rice with a high-caloric yield. Productivity will determine the ability of the polity that controls it to mobilize resources for military purposes. Our measure is akin to the use of population density to proxy for prosperity before the Industrial Revolution by Acemoglu et al. (2002). Figure 5 depicts y, the population density in 0 CE. Here we can see the high densities in the Italian peninsula, the Indian subcontinent, and China as well as the low densities of Russia, the Arabian peninsula, and inner Asia.

Later, we will show that our results are robust to the use of alternative measures of resource availability, including agricultural productivity (Ramankutty et al., 2002), potential caloric yield

(Galor and Özak, 2016), and population densities in 1000 BCE and 500 CE.

3.4 The Evolution of Polities

Time t is discrete: $t = 0, 1, 2, \ldots$. At the initial period $t = 0$, each cell begins as an independent polity. Over time, some polities expand and come to rule a block of cells while other polities lose control of cells.

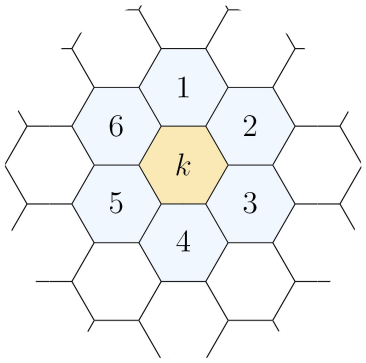


Figure 6: Cell k and adjacent cells.

As in our geographical space described above, each cell is a hexagon. Thus, the geographical space is filled by a hexagonal tiling, with each cell being bordered by adjacent cells 1–6 (Figure 6). We consider a regular tiling to impose ex ante homogeneity in the geographical shapes of polities. We prefer a hexagonal tiling to the other two regular tilings of the Euclidean plane because its vertex configuration is simpler than those of a triangular or square tiling. This simplicity reflects better, in our reading of historical data, the frontiers that most polities have had over time.

We describe now, in turn, conquest and secession of cells.

3.4.1 Conquest

In each period, a cell k finds itself in a border conflict with one of its adjacent cells with probability $\alpha \cdot y_k$, where $\alpha > 0$. For simplicity, we assume that when a cell experiences a border conflict, only one of its six borders is affected. Relaxing this assumption is straightforward, but it makes the model less transparent at the benefit of little additional insight. We make the probability of a border conflict depends on the productivity of the cell to capture the idea

that more productive cells are more tempting to neighbors to exploit (see Liberman, 1998, for evidence from industrial societies showing that, indeed, military conquest pays).

We assume that, conditional on cell k encountering a border conflict, the probability that its adversary is cell $\overline{k} \in \{1, 2, 3, 4, 5, 6\}$ is

$$\frac{y_{\overline{k}}}{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}, \tag{2}$$

where $y_1, ..., y_6$ are the respective productivities of the six adjacent cells. This assumption follows an intuition related to the paragraph above: two highly productive cells are more likely to be tempted into a conflict with each other than a low and a high productivity ones. A highly productive cell has a small incentive to enter into a conflict with a low productive one. Conversely, a low productive cell is afraid of engaging in a conflict with a high productive region.

A conflict between cells has two interpretations. If each cell is controlled by a different polity (as it occurs, for sure, in period 0), we think about this conflict as a war. The victor of this war, to be determined in the next paragraph, annexes the losing cell. If the cells are controlled by the same polity (as it will start occurring after a few periods), we think about this conflict as a political struggle for resources within the polity. The unified government will resolve the conflict by reallocating resources or through other policies in a manner that is inconsequential for our model.

Victory in a war between two polities is given by a contest function that depends on (a) the aggregate productivity of the polities in conflict, and (b) the geographical characteristics $\mathbf{x}$ of the cells in conflict. Specifically, if a war takes place between polities i and j, which controlled cells k and $\overline{k}$ respectively, polity i wins with probability:

$$\pi_i = \frac{Y_{i,t}}{(Y_{i,t} + Y_{j,t}) \times (1 + max\{\Theta \cdot \mathbf{x}_k, \Theta \cdot \mathbf{x}_{\overline{k}}\}}, \tag{3}$$

where $Y_{i,t}$ ($Y_{j,t}$) denotes the sum of productivities of all cells controlled by polity i (j) at period t, and $\mathbf{x}_k$ ($\mathbf{x}_{\overline{k}}$) denotes the geographical characteristics of cell k ($\overline{k}$) and Θ is a parameter vector that controls the weights of each geographical and climatic characteristic:

$$\Theta = \{\theta_{rugged}, \theta_{sea}, \theta_{frigid}, \theta_{torrid}, \theta_{forest}\}.$$

A random contest function such as (3) reflects two main ideas. First, more productive polities win more often, but the vagaries of war might bring victory, some times, to the weaker side. This may be due to random factors such as exceptional military leadership or strong state capability that we do not model here. Also, we specify that the relevant productivity depends on the sum of productivities of the cells of a polity, not the average productivity. Estonia, in 1939, was richer than the Soviet Union in per capita terms (see Norkus, 2019, for per capita output comparisons), but due to the difference in size, it could do next-to-nothing to resist annexation.

Second, the probability of victory is mediated by the geographical and climatic variables that make conquest harder or easier depending on the values of Θ. To see this notice that the probability of the war ending with no victor and, therefore, no annexation is:

$$1 - \pi_i - \pi_j = 1 - \frac{1}{1 + max\{\Theta \cdot \mathbf{x}_k, \Theta \cdot \mathbf{x}_{\overline{k}}\}},$$

which is strictly positive and is increasing in $max\{\Theta \cdot \mathbf{x}_k, \Theta \cdot \mathbf{x}_{\overline{k}}\}$. If $\theta_{rugged} >> 0$ (i.e., conquering very rough terrain is daunting, as scores of armies over millennia have discovered in Afghanistan), the probability of no annexation after a war that involves a cell with rough terrain is high.

Two secondary assumptions are worthy of further discussion. First, we assume that only the cell of the losing polity in the conflict is annexed, and not the whole polity. While complete conquest sometimes occur in history (think about the fall of the Sasanian Empire to the Arab invaders between 642 and 651), most conflicts end up with tradings of relatively small pieces of land (recall the dynastic struggles that plagued Europe during the early modern period and the subsequent small exchanges of territories).

Second, since a polity may have borders with multiple polities, it may face wars with several of them in the same period. We assume that a polity fighting more than one war will channel its resources proportionately according to the strengths of its adversaries. Otherwise, these wars are independent of each other. A good example of simultaneous struggles were the wars of Charles V, Holy Roman Emperor (r. 1519–1556), against many enemies across his domains. The Emperor always carefully weighted where to allocate his resources. His strategic choices were lamented by Francis I of France (r. 1515–1547) during his captivity in Madrid, but thoroughly enjoyed by the Elector John of Saxony (r. 1525–1532) while organizing the Schmalkaldic League.

We could generalize the previous two assumptions by allowing an annexation of larger parts

(or the totality) of a polity and the correlation of wars across frontiers. In our example above, Francis I and Suleiman the Magnificent (r. 1520–1566) signed an improbable alliance in 1536 against Charles V. While those generalizations are computationally straightforward, they require the introduction of many new free parameters. We prefer, instead, to keep our model tightly parameterized and enhance its interpretability even if at the cost of some realism.

3.4.2 Secession

To reflect the historical tendency for border regions in large states to seek secession, we allow border cells to secede from the polity they belong to with strictly positive probability in each period. A border cell is defined as one that shares an edge with one or more cells ruled by another polity.

We assume that border cell k's probability of secession is high if (a) the cell has a high $\Theta \cdot \mathbf{x}_k$ (i.e., geographical and climatic characteristics that make secession hard to suppress), (b) if the parent polity i controls a large number of cells (and is therefore heterogenous), or (c) if polity i has a long frontier relative to its interior (which increases the difficulty of monitoring and controlling the population). Specifically, the probability of border cell k seceding from polity i is:

$$\begin{aligned} \beta \times \Theta \cdot \mathbf{x}_k \quad &\times \quad \sum_{s}^{20,637} \mathbb{1}_i(s) \times \frac{\sum_{s}^{20,637} (\mathbb{1}_i(s) \cap 1_B(s))}{\sum_{s}^{20,637} \mathbb{1}_i(s)} \\ &= \quad \beta \times \Theta \cdot \mathbf{x}_k \times \sum_{s}^{20,637} (\mathbb{1}_i(s) \cap 1_B(s)), \end{aligned} \tag{4}$$

where $\mathbb{1}_i(s) = 1$ if cell s is ruled by polity i and $\mathbb{1}_i(s) = 0$ otherwise, and $\mathbb{1}_B(s) = 1$ if cell s is a border cell and $\mathbb{1}_B(s) = 0$ otherwise.

To simplify the analysis, we assume that if a polity is cut into disjoint parts due to war or succession, each part becomes a separate polity. Geographically-divided polities such as Pakistan between 1947 and 1971 seldom lived long. Note that our model applies to the "old world," and not to the Americas, Sub-Saharan Africa, and Oceania; where European sea empires expanded from the 15th century onwards separately from the metropolis. But even after the modern sea empires were created, the House of Hanover found much easier to prevail on Culloden Moor and Vinegar Hill than on Saratoga and Yorktown.

As before, for simplicity, we consider that each cell separates independently from other cells.

However, since a polity might have several cells sharing edges with other polities, it may suffer the separation of several cells in the same period.

3.4.3 Summary

As conflicts between polities and unrest within polities occur, state consolidation takes place over time as long as the probability of secession if not too high. Larger and more consolidated states have access to more resources, and hence are likely to consolidate further. However, some cells are more difficult to conquer than other cells due to their geographical and climatic characteristics. These features will lead to regular patterns of political concentration and fragmentation.

To summarize, the timing of events is as follows:

1. At $t = 0$, each cell is a separate polity (i.e., we have 20,637 polities).
2. At each time period, the probability of conflict breaking out in cell k is $\alpha \cdot y_k$, where $\alpha > 0$ and y_k is the productivity of cell k.
3. If cell k encounters a border conflict, only one of its six borders is affected. The conditional probability that its adversary is cell $\overline{k} \in \{1, 2, 3, 4, 5, 6\}$ is $\frac{y_{\overline{k}}}{y_1+y_2+y_3+y_4+y_5+y_6}$, where $y_1, ..., y_6$ are the productivities of the six cells bordering cell k.
4. If there is a conflict between adjacent cells controlled by different polities, a war takes place.
5. In a war between cells k and $\overline{k}$, controlled respectively by polity i and j, polity i wins and annexes cell $\overline{k}$ with probability given by the contest function (3).
6. A polity may fight no war, one war, or multiple wars at any period. If it fights multiple wars, it splits its resources proportionally according to the resources of its adversaries.
7. Cell k secedes from polity i with probability given by equation (4).

4 Calibration

To calibrate our model, we need to pick an initial and end point of the simulation, a time unit, and the values of seven parameters $\alpha, \beta, \theta_{rugged}, \theta_{sea}, \theta_{frigid}, \theta_{torrid}$, and θ_{forest}.

The initial and end points of the simulation are given by the geographical structure of our model and its associated historical dynamics. As we explained in Subsection 3.1, our model is designed to understand the evolution of Euroasia between, roughly, the beginning of the Iron Age (c. 1200–1000 BCE) and the dawn of the Age of Exploration in the second half of 15th century. These initial and end points gives us a total of around 2,500 years.

Let us explain our start and end choices in more detail. At the dawn of the Iron Age (c. 1200–1000 BCE), the "old world" was nearly fully fragmented. Even areas where larger polities existed previously, such as the Fertile Crescent, were recovering from the Late Bronze Age collapse: Egypt was transitioning through its third intermediate period, the palace economies of the Aegean had collapsed, and the Kassite dynasty of Babylonia and the Hittite Empire disappeared (see the classic account of this collapse in Drews, 1993, and, more recently, Cline, 2014). The Shang in China had achieved some progress in unification, but the documentary record of how effective their territorial control was is scant (Campbell, 2018, ch. 4).

The Age of Exploration integrated the whole world quickly. Juan Sebastián Elcano completed the first circumnavigation of the world in 1522, only 103 years after the Portuguese started the systematic exploration of the West African coast. And by 1565, the Manila galleon had opened a regular trade route between Europe, Asia, and the Americas (Giráldez, 2015).

Our choice of time unit must balance the need to have a detailed account of the evolution of political forms and the computational burden. Hence, we pick five years to get 500 simulation periods (2,500 years divided by 5). This time unit is also a reasonable approximation to the median length of many conflicts (which, in the data, have a huge variation.[13]

Fortunately, the values of all parameters in the model, except β, are time-independent, in the sense that they represent relative attractiveness or difficulties of conquest, which are inherently static properties (at least in the time horizon of human history). Therefore, our pick of an initial and end period and of a time unit of 5 years only matters in terms of how to map the lengths of outcomes in the model with the lengths of outcomes in the data.

We can move now to calibrate our seven parameters. Since $\alpha \cdot y_k$ determines the probability of conflict occurring in cell k, we set $\alpha = \frac{1}{y_{max}}$, where y_{max} is the productivity of the cell with

[13]Computing this variance become even more challenging once one realizes it is even hard to agree on what is a war. Think about the long conflict between the Spanish Empire and the Provinces of the Netherlands (1568-1648): Was it one long war or several consecutive ones?

the most resources in our dataset. In such way, $\alpha \cdot y = 1$ for the cell with the highest value of y and $0 \leq \alpha \cdot y \leq 1$ for all other cells.

We set $\beta = 0.000005$. Given the other parameter values in the model, this value of β implies that a cell that is at the 90th percentile of the ruggedness ranking and is controlled by a regime of 100 cells, approximately the size of Britain, will have a secession probability of 0.001 per period. We select a very low probability of secession to avoid biasing our results against Europe, which may be more likely to produce states that are uncompact in shape due to its long coastline.

We set $\theta_{sea} = \theta_{frigid} = \theta_{torrid} = 2$ so that the probability of a war ending with no annexation is twice the probability of it ending with annexation (i.e., $\frac{2}{3}$), if the war involves at least one cell that is a sea channel, or is too cold or too warm to be conducive to war. Since forests likely constituted a lesser obstacle to conquest, we set $\theta_{forest} = 1$.

Finally, we choose the value of θ_{rugged} so that $\theta_{rugged} \cdot x_{rugged} = 2$ for the cell at the 90th percentile of the ruggedness ranking. Below, we will conduct sensitivity tests to ensure that our results are not determined the precise values of these parameters.

Table 1 summarizes the calibration of the model's seven parameters.

Parameter	Value
α	$\frac{1}{y_{max}}$
β	0.000005
θ_{rugged}	$\frac{2}{x_{rugged}=90th \text{ percentile}}$
θ_{sea}	2
θ_{frigid}	2
θ_{torrid}	2
θ_{forest}	1

Table 1: Baseline calibration of the model.

5 Quantitative Results

Simulating our model in the computer is straightforward: we divide the geographical space into hexagonal cells, feed the geographical and climate characteristics and historical resource availability of each cell, and draw random paths of conflicts and secession. Since the evolution of the model is stochastic, replicating the idea that history is a mix of structure and contingency,

we simulate the model 19 times.[14] We checked that 19 simulations were enough to capture the ergodic distribution of events implied by the model.

Despite its extreme simplicity (and the omission of many plausible mechanisms of state formation), our model generates patterns of political consolidation and fragmentation that resemble those that we observe in history. Figure 7 depicts a map of our geographical space in a representative simulation in period 50. Following the calibration of time zero as the start of the Iron Age (c. 1200–1000 BCE) and a five-year time unit, this would correspond to c. 950–750 BCE. While nearly every cell is still an independent polity, we start seeing a consolidation of power in Northern China resembling the core areas of the Shang and the Zhou dynasties. In comparison, no large polities appear in Europe.

Figure 8 depicts the same simulation after 300 periods, which would correspond to c. 300–500 CE. We can see, in the east, a large polity in light green that closely matches historical China. In the West, one can see polities that roughly resemble Spain, Poland, or England. We can think about this moment as around the formation of the Germanic kingdoms that inherent the Western Roman Empire (including a Kingdom of the Suebi in the Northwest of the Iberian Peninsula, which existed in reality between 409–585). Interestingly, the Indian subcontinent is divided into quite a few polities, and the Arabian peninsula, given the low productivity of its land, is fragmented.

Figure 9 depicts the same simulation at the end of our 500 periods, which would correspond to 1300–1500. The large polity occupying China has expanded to the south toward Vietnam and Yunnan, and still clearly dominates East Asia. The polity controlling India has expanded toward the south, occupying an area similar to what the Mughal Empire reached at the death of Aurangzeb in 1707, but a couple of centuries earlier. In Europe, we see a unified Iberian Peninsula (as it happened between 1580 and 1640), polities corresponding rather closely to England, Scotland, Ireland, and Turkey, a polity comparable to European Russia, and a larger France.

Interestingly, at the end of the simulation, Southern Siberia is highly fragmented (Northern Siberia is excluded from our area of study), which also corresponds to the evidence in most of

[14] An odd number of simulations allows us to define the median simulation more easily. A short movie with a representative simulation can be seen here: https://www.dropbox.com/s/wm4jxqntuf9jz0j/Animation_Hexagon200514.mp4?dl=0.



Figure 7: Period 50

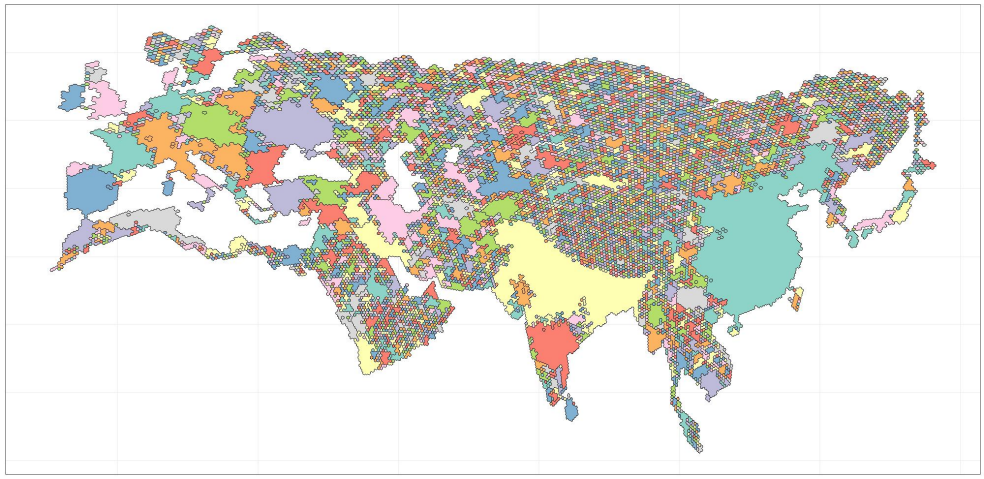


Figure 8: Period 300

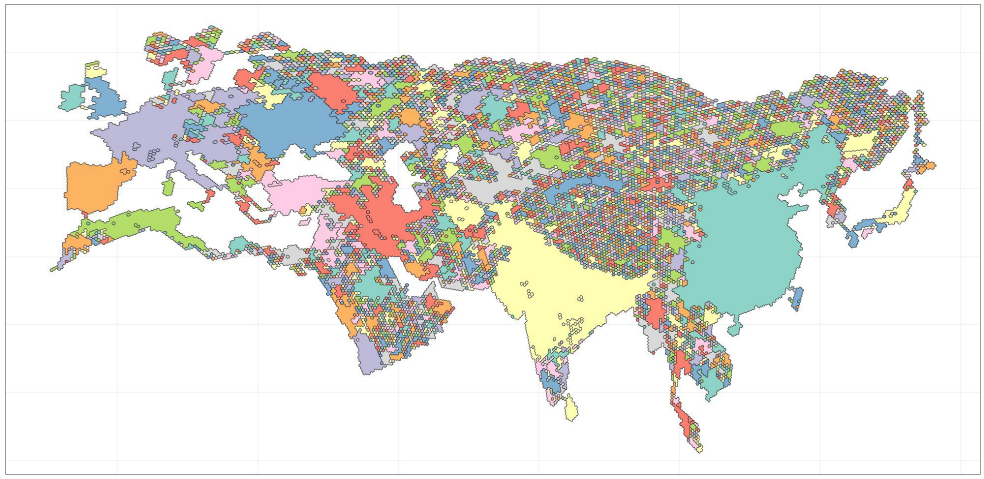


Figure 9: Period 500

the historical period we consider. Russians did not start the conquest of Siberia in earnest until the 16th century. What our simulation misses, however, is anything resembling the Mongol Empire and its successor states in the area, such as the Golden Horde, even if these processes of unification were rather transient.

5.1 Chinese Unification, European Fragmentation

The findings from the simulation above are not an anomaly. The first robust quantitative result of our model is that larger polities emerge early in China and that this part of the world tends to become unified under a single state. In contrast, political fragmentation is persistent over time in Europe.

Figures 10 and 11 depict the evolution of the Herfindahl index of political unification for China and Europe (defined as the red and green cells in Figure 2) over 500 periods for the 19 simulations under the baseline calibration. The heat map plots the estimated density of the index in each period.[15] Across all simulations, not only does political centralization occur in China, rather than in Europe, but political centralization in China emerges relatively quickly.

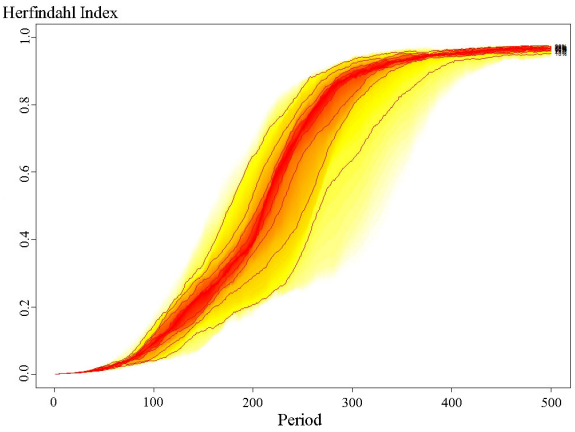


Figure 10: (China) Fan chart for 19 simulations of the benchmark model.

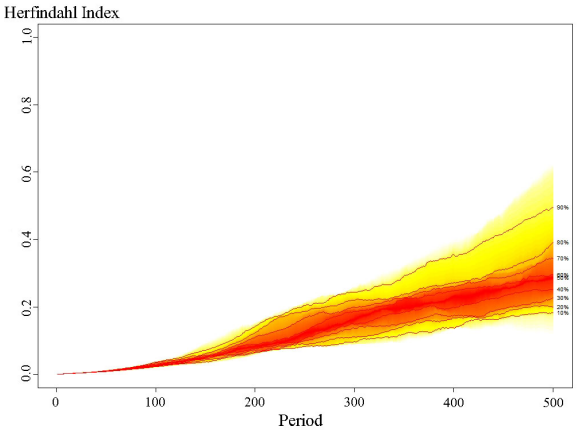


Figure 11: (Europe) Fan chart for 19 simulations of the benchmark model.

More concretely, in the case of China, the Herfindahl index crosses 0.75 after around 250 periods. In history, China was first unified by 221 BCE when the armies of Qin Shi Huang conquered the state of Qi, the last independent kingdom outside Qin's empire. Since we are

[15] In our context, the Herfindahl index of political unification of a region is defined as $H_{pc} = \sum_{i=1}^{N} s_i^2$, where N is the number of polities existing in the region, and s_i is the percentage of the cells in the region controlled by polity i.

taking time zero to be the start of the Iron Age (c. 1200–1000 BCE), 250 periods of our model is around 200 BCE–0 CE. In other words, not only does our model capture the consolidation of political power in China, but also its speed. In the case of Europe, the Herfindahl index stays as low as 0.25 in its median as late as 500 periods into the simulation.

5.2 Robustness of Findings

We now report a battery of experiments to assess the robustness of our results to the key assumptions of the model and to understand which forces drive our findings. In particular, we consider nine variations with respect to the baseline calibration. These variations are summarized in Table 2. For reference, our baseline calibration is included as specification (1). As a summary of our results, Figure 12 plots the median Herfindahl index of political unification for China and Europe in our 19 simulations for each of these 9 robustness exercises. For easy comparison, Panel (a) plots our baseline calibration.

Our first robustness exercise, which we call "minimum set of obstacles," eliminates the role of climate and forests in difficulting conquest by setting $\theta_{frigid} = \theta_{torrid} = \theta_{forest} = 0$ (column 2 in Table 2). This experiment is motivated by Diamond (1997), who focuses on mountain ranges and seas as barriers to conquest. Panel (b) shows that China unifies more rapidly than Europe. The main difference is that Europe unifies somewhat more by the end of the simulation. This is due to the absence of the barrier created by the central European forest. Because of its location at the very core of Europe, this densely forested area plays a bigger role in the baseline calibration than the frigid areas of Asia, which are more peripheral.[16]

Our second robustness exercise, which we call "no obstacles," pushes the argument in the previous exercise to its limit by setting the value of all parameters related to geographical and climatic barriers to conquest to zero: $\theta_{rugged} = \theta_{sea} = \theta_{frigid} = \theta_{torrid} = \theta_{forest} = 0$ (column 3 in Table 2). Panel (c) shows that, absent geographical and climatic barriers to conquest, Europe will still unify later than China, but will end up in a similar situation a few centuries later. Unification is more sluggish in Europe because China's core areas are more compact, facilitating early consolidation.

[16] This is not to say that forests did not slow down, in real history, China's conquest of Vietnam or Manchuria. Instead, we are just claiming those forests were located outside the main geographical center of political unification in East Asia.

Table 2: Summary of Specifications

	Baseline	Min. Obstacles	No Obstacles	Eur. Pop.×2	Uniform Pop.	No Obstacles + Uniform Pop.	Alternative Y (Ramankutty et al.)	Alternative Y (Galor and Ozak)	Alternative Y (Pop. 1000 BCE)	Alternative Y (Pop. 500 CE)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$y_{k \in China}$	Pop.×1	Pop.×1	Pop.×1	Pop.×1	0.5	0.5	Agri. Suitability	Caloric Yield	Pop.×1	Pop.×1
$y_{k \in Europe}$	Pop.×1	Pop.×1	Pop.×1	Pop.×2	0.5	0.5	Agri. Suitability	Caloric Yield	Pop.×1	Pop.×1
$y_{k \in RoW}$	Pop.×1	Pop.×1	Pop.×1	Pop.×1	0.5	0.5	Agri. Suitability	Caloric Yield	Pop.×1	Pop.×1
$\theta_{rugged} \cdot x_{rugged}$ (90th pctl)	2	2	0	2	2	0	2	2	2	2
θ_{sea}	2	2	0	2	2	0	2	2	2	2
θ_{frigid}	2	0	0	2	2	0	2	2	2	2
θ_{torrid}	2	0	0	2	2	0	2	2	2	2
θ_{forest}	1	0	0	1	1	0	1	1	1	1
Resource Comparison (y_k, median value)										
–China	0.35	0.35	0.35	0.31	0.50	0.50	0.71	0.72	0.08	0.23
–Europe	0.17	0.17	0.17	0.31	0.50	0.50	0.63	0.74	0.07	0.17
Fracturedness Comparison ($\Theta \mathbf{x}_k$, median value)										
–China	0.86	0.80	0.00	0.86	0.86	0.00	0.86	0.86	0.86	0.86
–Europe	0.79	0.45	0.00	0.79	0.79	0.00	0.79	0.79	0.79	0.79

All y_k values normalized to 0–1.

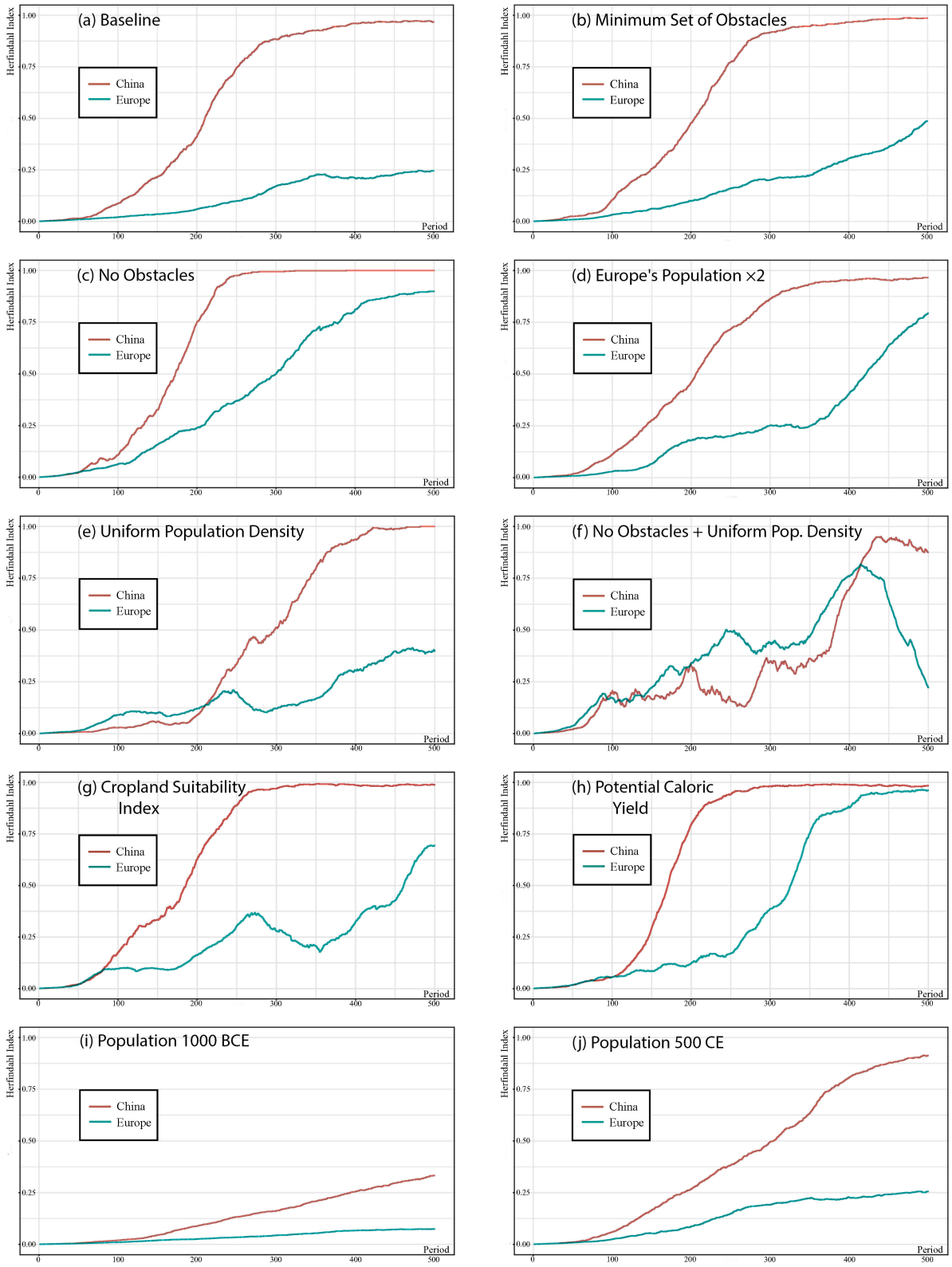


Figure 12: (Sensitivity Analysis) For each specification, we conduct the simulation exercise for 19 times and display the 'median" simulations for China (red) and Europe (green). The median simulation is the simulation with an average Herfindahl (Unification) index over 500 periods that ranks 10th out of 19.

Our third robustness exercise, which we call "doubling Europe's population," assesses whether a more populated Western Europe would have induced greater political consolidation. According to the data in Goldewijk et al. (2017), China's population was twice that of Western Europe in 0 CE. Perhaps the higher population in China (evidence, under a Malthusian logic, of higher resources and better technology, in this case probably rice and a climate suitable for its cultivation) creates more incentives to consolidate through the probability of a conflict $\alpha \cdot y_k$. To check this hypothesis, we double the population in every European cell (column 4 in Table 2). Panel (d) of Figure 12 shows that, indeed, population drives part of our results, with Europe achieving a much higher degree of unification by the end of the simulation. However, China still unifies earlier and remains more centralized than Europe, even at period 500.

Our fourth robustness exercise, with we call "uniform population density" is, essentially, the extreme converse exercise to "doubling Europe's population." Now we assume that population is uniformly distributed across our study area, with $y = 0.5$ for all cells. Thus, every cell is equally likely to engage in conflict (column 5 in Table 2). Panel (e) of Figure 12 is nearly the same as Panel (a), our baseline calibration.

The fifth robustness exercise combines "no obstacles" and "uniform population density." We assume that there are no geographical and climatic barriers to conquest, and population is uniformly distributed across our geographical space (column 6 in Table 2). In this counterfactual, our geographical space is neither "fractured" by geographical and climatic obstacles, nor separated into land clusters of varying productivity levels. Panel (f) of Figure 12 shows that once we neutralize both aspects of "fractured-land," China and Europe now unify at a comparable pace.

Next, we perform four robustness checks to ensure that the results are not sensitive to our choice of using world population in 0 CE as a proxy for historical resource availability. In Figures 12g–12j, we reinstate the role of geographical obstacles but replace population density in 0 CE with cropland suitability (based on Ramankutty et al., 2002), potential caloric yield (based on Galor and Özak, 2016), and population density in 1000 BCE and 500 CE (based on Goldewijk et al., 2017), respectively, as our y variable. Using these alternative measures, we continue to observe political unification in China taking place faster than in Europe.[17]

We also perform sensitivity tests with respect to the values of the parameter vector Θ, which

[17]We have also checked using population density in 1000 CE and 1500 CE to measure historical resource availability. The results are qualitatively similar and are available upon request.

measures the influence of the geographical and climatic characteristics on war outcomes. In the baseline, we set $\theta_{frigid} = \theta_{torrid} = \theta_{sea} = 2 \times \theta_{forest} = \theta_{rugged} \cdot x_{rugged\ (at\ 90th\ pctl)} = \theta = 2$. To ensure that the results are not determined by the precise values of these parameters, we set θ to different integer values between 0–10. For each integer value, we repeat the simulation 19 times. For each simulation, we compute the average Herfindahl index over 500 periods (H_{mean}) and the Herfindahl index at $t = 500$ (H_{final}) for China and Europe, respectively.

Figures 13 and 14 report the median H_{mean} and the average H_{final} of the 19 simulations at each integer value of θ. When $\theta = 0$, geographical obstacles have no influence on war outcomes (which corresponds to the case in Figure 12c). As θ increases, the likelihood of war ending with no victor or conquest increases with the presence of geographical obstacles. Indeed, as the two figures illustrate, the Herfindahl indices of China and Europe and, therefore, their respective likelihood toward political unification decrease with θ. However, at all values of θ, China displays a stronger tendency toward political unification.

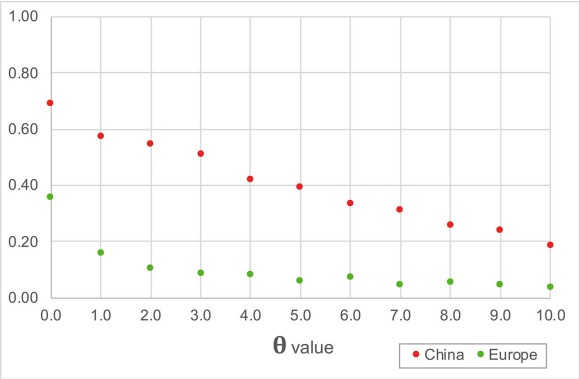


Figure 13: Average Herfindahl index over 500 periods at different values of θ.

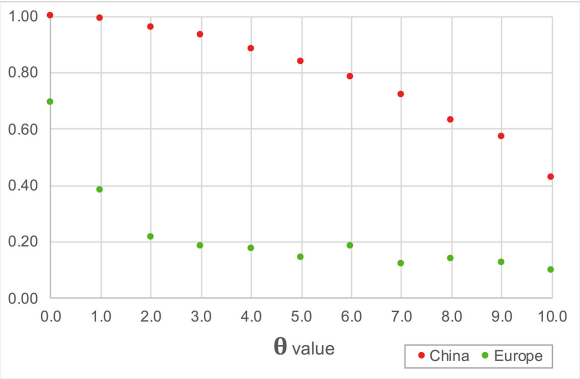


Figure 14: Herfindahl index at $t = 500$ at different values of θ.

In sum, all these sensitivity results suggest that it is insufficient to simply compare average levels of ruggedness between China and Europe. What mattered was the *distribution* of mountains and other geographical obstacles. While China is, indeed, more rugged than Europe, the location of geographical obstacles worked in a way that promoted faster political unification in China. Furthermore, while topography alone is a sufficient condition to explain China's recurring unification and Europe's persistent fragmentation, it is not necessary. Take away topography, and we continue to observe more rapid unification in China. Only removing *both* geographical barriers and land productivity ensures that China and Europe unify at a comparable pace.

6 Extensions

In the Online Appendix (A.1), we consider a number of additional robustness checks. In the interest of space, we summarize our main results here here.

The Eurasian Steppe Our first extension is allows polities proximate to the steppe to have an edge in conflict. This extension allows us to engage with the argument of Turchin et al. (2013), who note that the Eurasian steppe influenced state building both directly, because steppe nomads eliminated weaker and less cohesive polities, and indirectly, by developing and spreading technologies that intensified warfare.[18] Another factor that made the steppe historically critical is that it served as an undifferentiated "highway of grass," to which nomads, who were not dependent on holding land, could retreat in the face of attack by agrarian states or take advantage of to travel from Mongolia to the Black Sea in a matter of weeks (Frachetti, 2008, 7).

To capture these ideas, we modify our model in Appendix A.1 to give a premium in war and consolidation to cells located close to the steppe. Such extension does not qualitatively impact our main findings on the rate of political unification in Europe or China because it does affect the productivity and barriers to conquest of the core areas of state formation.

Major Rivers Our second extension allows rivers to play a role in our model. Scholars have argued that riverine connectivity contributed to patterns of political unification and fragmentation. Diamond (1997, 331) noticed that "China's long east-west rivers (the Yellow River in the north, the Yangtze River in the south) facilitated diffusion of crops and technology between the coast and inland." The role of rivers in England's early development is widely discussed by medievalists and historical geographers (see Langdon, 1993; Edwards and Hindle, 1993; Jones, 2000). Historically, numerous battles occurred took place either side of an important river.[19] Armies used major rivers as a source of supply. During antiquity, for instance, Roman

[18] According to Lattimore (1940), the struggle between the pastoral herders in the steppe and the settled populations in China was, first and foremost, an ecological one. The geography of Eurasia created a natural divide between the river basins of China and the Eurasian steppe. In the Chinese river basins, fertile alluvial soil, sufficient rainfall, and moderate temperature encouraged the early development of intensive agriculture. In the steppe, pastoralism emerged as an adaptation to the arid environment. Given the fragile ecology of the steppe, where droughts often led to extensive and catastrophic deaths among animal herds, the steppe nomads were impelled to invade their settled neighbors for food during periods of cold temperature.

[19] Some notable examples include the Battle of Granicus (334 BCE), the Battle of Rhone Crossing (218 BCE), the Battle of the Medway (43), the Battle of Red Cliffs (208), the Battle of the Milvian Bridge (312), the Battle

invasions of Persia frequently followed the course of the Euphrates. Even as recently as during the U.S. Civil War, most operations in the West followed rivers (the Mississippi, the Cumberland, etc.) But, at the same time, rivers separate basins and could impede movement between left and right banks.

We capture the role of rivers by increasing the probability of conquest when cells along the same river come into conflict and decreasing the probability of conquest when a riverine cell fights a non-riverine cell. The extension yields similar results as our baseline calibration, with only slightly slower political unification in China.

Exogenous Shocks and Dynastic cycles We introduce exogenous shocks into our model to observe whether they generate the rise and fall of particular polities –a phenomenon in Chinese history often interpreted through the lens of dynastic cycles (Usher, 1989; Chu and Lee, 1994). Recently, scholars have pointed to climate change as a cause of these dynastic cycles (see Zhang et al., 2006; Fan, 2010). Climatic factors have also been adduced as important in the rise and fall of the Roman empire (see Section 7.2) and widespread social upheavals in late medieval and early modern Europe (Lamb, 1982; Parker, 2013; Campbell, 2016).

We incorporate climatic and other random sociopolitical shocks by distinguishing between general system-wide crises, such as the collapse of Bronze Age empires c. 1177 BCE or the Little Ice Age of the 17th century, and regime specific crises that affect one polity, such as the Twenty Years' Anarchy in the Byzantine Empire (695–717), the An Lushan Rebellion of Tang China (755–763), or the War of Roses of 15th-century England. In our extension, these two kinds of crises occur randomly given exogenous probabilities.

In this version of the model, political cycles are muted in Europe, which never achieves full unification despite short periods of a hegemonic state. By contrast, China displays a pattern of periods of sustained unification interrupted by periods of disunity, resembling the successive dynasties of Chinese history (which motivated the opening quote of this paper). The result echoes Root (2017), who contrasts patterns of network stability in China and Europe and argues that China's organization as a hub-and-spoke system was less resilient than Europe's polycentricity, and Ko et al. (2018), who show that the Chinese empire displayed greater volatility of population and economic output than Europe after the collapse of the Roman Empire.

of Fei River (383), the Battle of Stamford Bridge (1066), and the Battle of Stirling Bridge (1297).

The Mediterranean Sea In our benchmark analysis, sea cells that are adjacent to land cells are traversable, but seas or other large bodies of water are barriers across which conflicts cannot take place. This assumption is consistent with the difficulties of large-scale maritime invasions in the premodern period and the capital-intensive nature of naval technology. However, this may not be the best way to model the Mediterranean sea, which, due to its geography, is calmer and less dangerous to traverse than other major seas. Indeed, scholars such as Braudel (1949) and Horden and Purcell (2000) have pointed out that the Mediterranean facilitated the spread of shared cultural values and institutions throughout history. Importantly, the most notable European state in the premodern period, the Roman empire, was based on the control of "Mare Nostrom" (*Our Sea*). It is therefore natural to ask what assumptions are required to generate an empire like the Roman empire forming around the Mediterranean sea.

Historians such as Harper (2017) have pointed to the confluence of favorable climatic conditions that facilitated the rise of the Roman empire, while Scheidel (2019) observes that Rome's success hinged on it obtaining early control of the Mediterranean before many competing powers could appear.[20]

Following these observations, we extend the model by making the Mediterranean traversable. This alone does not increase the likelihood of either a European empire or a Mediterranean empire emerging. It is only when we impose the additional condition that only polities spanning 100 cells (i.e., roughly the size of England) or more can attempt to control the Mediterranean Sea (to account for the fact that building a navy is capital intensive) that we observe the *occasional* formation of a Mediterranean empire. Furthermore, any likelihood of a Mediterranean empire emerging disappears once we replace population distribution in 0 CE with population distribution after 500 CE as our measure of resource availability, corroborating the accounts of Harper (2017) and Scheidel (2019) of why the Rome empire emerged when it did.

[20] Scheidel (2019, 74) notes that "Roman mastery of the Mediterranean was unique: never again in history would one power exercise lasting control over its entire coastline, and its effective naval supremacy was not renewed until the days of Admiral Nelson, if not the Second World War. Moreover, the Roman dominions were unusual simply for being centered on the Mediterranean: among later sizeable empires, only Habsburg Spain and the Ottomans shared this distinction, although on a smaller scale, especially the former. Neither one of them enjoyed anything like Roman hegemony. This is easy to explain. Even though a unified Mediterranean may have been a highly suitable core for an empire that already dominated it, later history documents the difficulties of reaching the requisite position of preeminence. This happened only once, at a time when lack of competition made it less challenging to establish hegemony over the less developed western half of the Mediterranean. Considering how much Rome struggled against just a single opponent during its first war with Carthage, a more crowded naval environment might well have prevented naval expansion".

State Formation Across Eurasia Consistent with what we observe historically, in our simulations the formation of large states is most pronounced in East Asia.[21] By contrast, large states in Europe are rare and transient. To what extent can our model also explain broader patterns of political fragmentation across Eurasia beyond East Asia and Europe?

To investigate, we compute the probability of a large state –defined as a polity controlling 600 cells or more (approximately the combined size of Turkey and Iraq)– arising in Europe, East Asia, South Asia, Southeast Asia, or the Middle East.

In our simulations, large states also emerge regularly in northern India. However, upon visual observation of the simulations, a single, huge polity does not always conquer the entire Indian subcontinent for several reasons. First, the Himalayas and the Hindu Kush in the north, the Thar Desert in the west, and the thick jungles of Burma and Gondwana in the east presented significant impediments in terms of either rugged terrain or low agricultural productivity that discourage state expansion in these directions.[22] Second, the rugged Deccan Plateau in southern India presented a formidable barrier to empire-building.[23] Third, the tropical climate of southern India, which historically posed difficulties in gathering and moving armies (see Lieberman, 2003, 2009), further impeded the conquest of the south by the north.

The tropical climate, together with the rugged terrain and low population density of Indochina, also explain why, in our simulations, large empires do not arise in Southeast Asia.[24]

The probability of large states arising in the Middle East is also very low in our baseline

[21] As Scheidel (2019) notes the "easternmost macro-region, East Asia, has been characterized by much stronger dominance of hegemonic empire than any of the others".

[22] Historically, while the thick jungles of Burma and Gondwana created unbridgeable, outer limits to Mughal expansion (Gommans, 2002, 198), the mountains were not insurmountable to armies and the Mughals conducted mountainous expeditions into places like Kashmir (1561, 1585, 1588), Garhwal (1635, 1656), Baltistan (1637), Ladakh and Tibet (1679–84). However, the lack of forage and food impeded all attempts to extend political authority permanently north of India. As Gommans (2002, p. 23) puts it: "Indian armies were faced by tremendous logistical problems. One mid-eighteenth-century source considered the Kabul area a land of snow: 'Men and cattle from India are not able to withstand the icy cold winds of that area. That is why it is difficult for the people of India to capture and occupy the Muslim countries of that area."' See also Nath (2019) for how diverse natural environment of South Asia interacted with Mughal warfare.

[23] The Deccan Plateau rises to over 1000 meters. It was the site of numerous conflicts between states from northern India and those from southern India. Multiple Hindu states in the Deccan were able to resist the expansion of Muslim empires such as the Mughals.

[24] Southeast Asia was less populous than other major regions of Eurasia until the 19th century and state formation took place later and under less favorable conditions there than elsewhere (Lieberman, 2003, 2009). There were periods that saw the formation of states with a considerable geographical scope such as the Khmer empire in the 9th century, the Taoungoo empire in the 16th century, or the Kingdom of Siam in the 18th and 19th centuries. But these larger states only retained regional hegemony for brief periods of time and the more common pattern was political fragmentation and polycentricity.

simulations, which may appear inconsistent with the historical record, but it increases substantially when we extend the model to give an advantage in war to cells located close to the steppe. This corroborates the arguments of historians that the steppe played a key role in driving state formation in the Middle East (Tapper, 1997; Kulikowski, 2019).

7 Historical Discussion Informed by the Model

Our model generates several insights that help us understand patterns of state formation across Eurasia. In this section, we use our model to inform a discussion of China and Europe.

7.1 China

At first glance, it is not obvious why geography should contribute to China's recurring political unification. China's terrain is significantly more rugged than Europe's (Hoffman, 2015b). Also, the climatic distinction between China's temperate north and subtropical south is stark. Different climatic conditions divided China into two agricultural zones: historically, sorghum and wheat were the staple crops in northern China, while rice was dominant in the south. Different crops, in turn, encouraged the development of different social organizations and cultural norms in each zone (Talhelm et al., 2014).

China's tendency toward political unification intrigued the Chinese themselves. During the late Warring States period, Lü Buwei, the chancellor of the Qin kingdom, noted that the number of states in China had decreased from tens of thousands c. 2200 BCE to three thousand c. 1600 BCE to only a handful in his time and asked why (Sellmann, 2002). Not long after Lü death, all but one of the remaining surviving states would perish as Qin built China's first unified empire in 221 BCE. While the Qin dynasty lasted only 15 years, it marked a watershed. From 221 BCE to the founding of the Chinese Republic in 1911, China was unified for 1142 years out of 2132 years (Ko and Sng, 2013). The record is unparalleled in world history.

Our model cast light on these phenomena by highlighting the salient role of North China in fostering unification. While North China is only one of several macroregions of China (Figure 15), it played an outsized role in Chinese history and was referred to as the "Central Plain" in historical records. The silty and flood-prone Yellow River, China's "mother river," runs

through the region. Regularly inundated by flooding, which replenished the soil, north China was agriculturally precocious and productive even with primitive agricultural tools (Huang, 1988). In 1943, Sha Xuejun, one of the first modern scholars of political geography in China, used the term "the hub of China" to describe North China and Lower Yangzi (Sha, 1972). Possibly paraphrasing Mackinder (1942), he remarked that:

> To control China, one needs to first control its heartland; To control China's heartland, one needs to first control its hub.

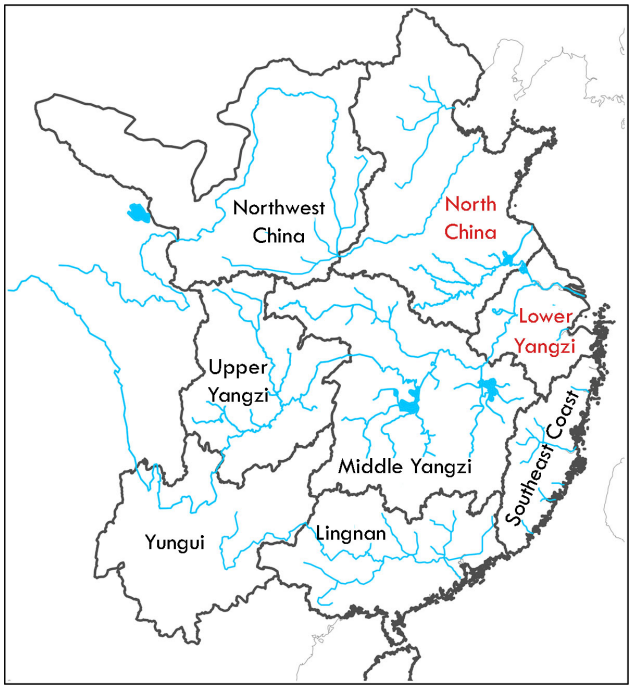


Figure 15: China's macroregions (Skinner, 1977).

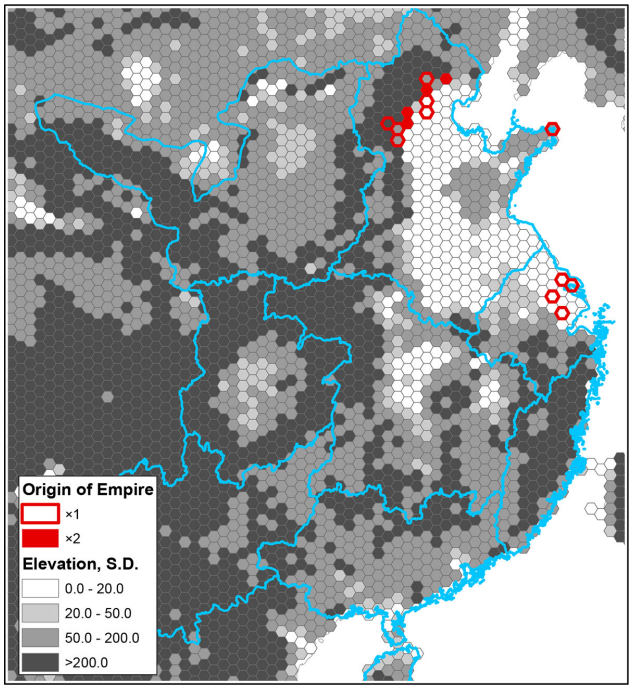


Figure 16: Flatness & Centrality of North China.

Indeed, in our simulations, North China stands out for its flat terrain and high agricultural productivity. The flatness of North China facilitates military conquest and political consolidation within itself (see in Figure 16 the cells from where the empire appears in each simulation). And once a unified state emerges in North China, the wealth of resources at the region's disposal makes it difficult, if not impossible, for the other Chinese regions –which find internal unification harder to achieve due to their rugged terrains– to resist being absorbed. The proximity and lack of significant natural barriers between north China and Lower Yangzi, another agriculturally productive region, accelerate these processes. Furthermore, due to the presence of the steppe

and deserts north of China, the Tibetan Plateau, and the Pacific Ocean, the north Chinese state generally expands in a southerly direction until it hits the tropical rainforests of Indochina and further expansion is hindered by an increased probability of secession. Thus, it is not a surprise that the resulting empire in our simulations often approximates the shape of China proper.

Our model also accounts for why the Qin dynasty could unify China despite attempts to maintain a "balance of power" by its adversaries during the Warring States period. The historical literature points to the reforms enacted by the Qin state, notably by Shang Yang. These included conscription, large-scale irrigation projects, and a system of land registration (Hui, 2005). However, these reforms were also pursued, partially or fully, by the other warring states. Why did "balancing" fail in China but succeed in Europe? Our model points to the lack of geographical barriers in North China and adds an emphasis on the potential for cumulative conquests in this core region of historical China.

As Figure 16 illustrates, political unifications of China in all of our 19 steppe-enriched simulations originate from cells in North China or Lower Yangzi (Section 6). The simulations we run suggest that the geographical characteristics of the northern Chinese plain made it possible for a single powerful state to overcome its rivals and to build a centralized state. This is broadly consistent with the historical record. All but one of the nine dynasties that controlled most or all of China proper at their peaks originated from the north (Turchin, 2009; Scheidel, 2019) (Table 3).[25] The exception was Ming dynasty, which came from the Lower Yangzi.

Importantly, our model is consistent with episodes of fragmentation interspersing periods of unification. Historically, there were long periods of political fragmentation in China: the Warring States period (475 BCE–221 BCE), the Three Kingdoms period (220–280), the Five Dynasties and Ten Kingdoms period (907–960) and the Southern Song period (1127–1279). However, if a powerful Chinese state did arise to gain control of North China and Lower Yangzi, it would often go on to subdue rival kingdoms and unify "all under heaven."

[25] In Scheidel's words "In China, empire emanated almost exclusively from the northern frontier" (Scheidel, 2019, 169). This is true of the formation of the First Chinese Empire under the Qin Dynasty in 221 BCE, but also of the Sui, Tang, Yuan, and Qing dynasties. Taizu, the first Song emperor, was a general from the Later Zhou, the most northern state in China.

7.2 Europe

Our model also illustrates the mechanisms that ensured that Europe remained politically fragmented for much of its history. The closest that Europe came to being ruled by a single polity was during the Roman empire. This was a unique development in European history; at no other point did a single polity come close to establishing stable rule over the majority of the European landmass or population.

Numerous factors were important in explaining the rise of Rome: the Mediterranean Sea as a conduit to empire, Rome's peripheral position on the edge of Eastern Mediterranean state system based around the Fertile Crescent, its early ability to incorporate the nearby population of Latium and to build an alliance system of nearby Italian cities, its unusual bellicosity (Harris, 1979, 1984), and favorable climatic conditions during the classical period (Braudel, 1949; Horden and Purcell, 2000; Harper, 2017; Scheidel, 2019). As we discussed in Section 6, making the Mediterranean traversable leads to the *occasional* formation of a Mediterranean empire, but only when we measure resource distribution using population data based on 0 CE. The result goes away when we use population distribution after 500 CE as our measure of resource availability. This observation lends credence to Harper and McCormick (2018)'s account that the Roman warm period increased the agricultural productivity of southern Europe and North Africa, favoring the expansion of a Mediterranean-based empire into the rest of Europe and the Near East.

Despite their remarkable successes, the Romans could not permanently incorporate Germany and Eastern Europe into their empire. Our model points to the dense northern European forest as one contributing factor that impeded the consolidation of a single European hegemonic state. When we remove the northern European forest in our model, the probability of Europe unifying under a hegemonic state increases (Figure 12b). The critical role of the northern European forest in deterring Roman expansion is attested to in the historical sources. See, for example, Tacitus (1877), Begle (1900), and Howorth (1909).

In this way, our model elucidates the role played by mountain and forest barriers in European history. Mountains and forests did not pose an insurmountable impediment to armies. Hannibal crossed the Pyrenees and Alps in 218 BCE. Frederick Barbarossa (r. 1155–1190) invaded Italy from Germany from the 1150s to the 1170s. Peter II of Aragon (r. 1196–1213) crossed the

Pyrenees to confront the army of Simon de Montfort in 1213 in the Battle of Muret. Nevertheless, the failure of most of these expeditions illustrates all these barriers substantially raised the cost of military interventions among polities in modern-day Spain, France, Germany, and Italy.[26]

After the fall of the Roman empire, the closest Europe came to being unified by a single ruler was during the sixteenth century under the Habsburg emperor Charles V.[27] Importantly, for our purposes, however, Charles V did not acquire this empire by conquest, but by generations of successful marriages and dynastic luck. He also did not create a unified state, but ruled his disparate kingdoms as separate entities. As such, our model does not directly speak to how the Habsburgs chanced on a European-wide empire. It does, however, speak to the difficulties Charles V had in managing his domains. Geography prevented Charles V from focusing attention on either facing the Ottomans in the Mediterranean, driving France out of Italy, or subduing Protestant German princes in the Holy Roman Empire. The final admission of the power of geography came when, at his abdication, Charles divided his territories between his son Philip II of Spain (r. 1556–1598) and his brother Ferdinand I (r. 1556–July 1564).

Consistent with our model, Habsburg hegemony depended on the possession of the rich and densely populated Low Countries and the productive lands of the Duchy of Milan. Habsburg dominance began to unravel once they lost control of the Netherlands as a result of the Dutch Revolt.[28] Following this defeat, and the consolidation of the Reformation in Germany, the Habsburg's lost their preeminence among European powers despite the wealth of the New World acquired by Spain.

Beyond accounting for the Habsburg's experience, our model also speaks to the failures of Lous XIV and Napoleon to successfully build a hegemonic state in Europe. The emergence of several medium-sized states in Europe is a common feature of our model. This captures the argument that a "balance of power" was crucial to preventing either Spain or later France from building a long-lasting continent-spanning state.[29] The prominence of several medium-size states

[26]Hannibal had to retreat back to Carthage after 15 years and witnessed Rome's ultimate triumph. Barbarossa's Italian campaigns did not yield much to the Empire. Peter perished in the Battle of Muret.

[27]Charles V ruled "a greater number of realms than had ever before been accumulated by any European ruler" and his territories spanned much of Europe with the result that "his duties took him everywhere.'; It is estimated that he spent 25% of his reign traveling. He himself described his life one long journey (Kamen, 2002, 50).

[28]Supplying the Spanish Tercios in the Low Countries became a logistical challenge of first magnitude that eventually doomed the Habsburg's efforts at defeating the United Provinces. These difficulties speak directly to the key importance of geographical barriers. See, for details, the classic account of Parker (2004).

[29]The concept of balancing is the subject of an extensive literature in International Relations, see Waltz (1979) and Mearscheimer (2001).

in our model is driven by Europe's geography. It reflects not only mountain barriers, stressed by Diamond (1997, 1998), but also the fact that the most productive agricultural land in Europe is dispersed, rather than concentrated, as is the case in China.

8 Conclusion

In this paper, we develop a model that allows us to adjudicate among competing explanations of Europe's political fragmentation and China's political centralization.

Our analysis takes seriously Jared Diamond's argument that Europe's mountain barriers and the shape of its coastline were responsible for its political fragmentation, whereas Chinese geography encouraged political centralization. By developing an explicit model of state formation that quantitatively incorporates the role of both topography and agricultural productivity, we provide a rigorous formulation of the fractured-land hypothesis. We demonstrate, through our simulations, that either topography or the location of productive land can generate political unification in China and persistent political fragmentation in Europe.

Furthermore, our model can be a starting point for numerous additional explorations. For example, in the line of Hoffman (2015a), we could incorporate military technological change. We could also add time-varying agricultural technology, or cultural aspects that feedback into the creation of states. For instance, after a state has been in existence for many periods, its inhabitants may have developed an "imagined community," which makes it harder to conquest but easier to maintain unified (Anderson, 1991). Also, some cells may share a religion, which makes unification easier, or by separated by it, which makes conflict more likely. Our methodological approach is quite flexible in allowing for these and many other quantitative exercises and in generating probability distributions of historical outcomes.

References

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics 117*(4), 1231–1294.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a fundamental cause of long-run growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 6, pp. 385–472. Elsevier.

Acharya, A. and A. Lee (2018). Economic foundations of the territorial state system. *American Journal of Political Science 62*(4), 954–966.

Anderson, B. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso.

Ashraf, Q. and O. Galor (2011). Dynamics and stagnation in the Malthusian epoch. *American Economic Review 101*(5), 2003–41.

Baechler, J. (1975). *The Origins of Capitalism*. Oxford: Basil Blackwell.

Barfield, T. (1989). *The Perilous Frontier: Nomadic Empires and China*. Cambridge, MA: Basil Blackwell.

Begle, G. G. (1900). Caesar's account of the animals in the Hercynian forest (de Bello Gallico, vi, 25–28). *The School Review 8*(8), 457–465.

Braudel, F. (1972, 1949). *The Mediterranean, and the Mediterranean World in the Age of Phillip II, Vol. II*. Berkeley: University of California Press.

Campbell, B. M. (2016). *The Great Transition: Climate, Disease and Society in the Late-Medieval World*. Cambridge: Cambridge University Press.

Campbell, R. (2018). *Violence, Kinship and the Early Chinese State: The Shang and their World*. Cambridge: Cambridge University Press.

Carneiro, R. L. (1970). A theory of the origin of the state. *Science 169*(3947), 733–738.

Chaudhry, A. and P. Garner (2006). Political competition between countries and economic growth. *Review of Development Economics 10*(4), 666–682.

Chen, Q. (2015). Climate shocks, dynastic cycles and nomadic conquests: Evidence from

historical China. *Oxford Economic Papers 67*(2), 185–204.

Chu, A. C. (2010). Nation states vs. united empire: Effects of political competition on economic growth. *Public Choice 145*(1/2), 181–195.

Chu, C. Y. C. and R. D. Lee (1994). Famine, revolt, and the dynastic cycle: Population dynamics in historic China. *Journal of Population Economics 7*(4), 351–378.

Cline, E. H. (2014). *1177 B.C.: The Year Civilization Collapsed*. Princeton, NJ: Princeton University Press.

Cowen, T. (1990). Economic effects of a conflict-prone world order. *Public Choice 64*, 121–134.

Diamond, J. (1997). *Guns, Germs, and Steel*. New York: W.W. Norton & Company.

Diamond, J. M. (1998). Peeling the Chinese onion. *Nature 391*, 433–434.

Downing, B. M. (1992). *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton, NJ: Princeton University Press.

Drews, R. (1993). *The End of the Bronze Age: Changes in Warfare and the Catastrophe Ca. 1200 B.C.* Princeton, NJ: Princeton University Press.

Edwards, J. F. and B. P. Hindle (1993). Comment: inland water transportation in medieval England. *Journal of Historical Geography 19*(1), 12 – 14.

Fan, K. (2010). Climatic change and dynastic cycles in Chinese history: a review essay. *Climatic Change 101*(3), 565–573.

Frachetti, M. (2008). *Pastoralist Landscapes and Social Interaction in Bronze Age Eurasia*. Berkeley: University of California Press.

Galor, O. (2005). From stagnation to growth: Unified growth theory. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*. Elsevier.

Galor, O. (2011). *Unified Growth Theory*. Princeton, NJ: Princeton University Press.

Galor, O. and O. Özak (2016). The agricultural origins of time preference. *American Economic Review 106*(10), 3064–3103.

Galor, O. and D. N. Weil (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review 90*, 806–828.

Gat, A. (2006). *War in Human Civilization*. Oxford: Oxford University Press.

Gennaioli, N. and H.-J. Voth (2015). State capacity and military conflict. *Review of Economic Studies 82*(4), 1409–1448.

Giráldez, A. (2015). *The Age of Trade: The Manila Galleons and the Dawn of the Global Economy*. Rowman & Littlefield.

Goldewijk, K., A. Beusen, J. Doelman, and E. Stehfest (2017). Anthropogenic land use estimates for the Holocene–HYDE 3.2. *Earth System Science Data 9*(2), 927–953.

Gommans, J. (2002). *Mughal Warfare*. London: Routledge.

Grousset, R. (1970). *The Empire of the Steppes: a History of Central Asia*. New Brunswick: Rutgers University Press. Translated by Naomi Walford.

Hajnal, J. (1965). European marriage pattern in historical perspective. In D. V. Glass and D. E. C. Eversley (Eds.), *Population in History*, pp. 201–213. London: Arnold.

Hall, J. A. (1985). *Power and Liberties*. London: Penguin Books.

Harper, K. (2017). *The Fate of Rome*. Princeton, NJ: Princeton University Press.

Harper, K. and M. McCormick (2018). Reconstructing the Roman climate. In W. Scheidel (Ed.), *The Science of Roman History*, Princeton, NJ, pp. 11–52. Princeton University Press.

Harris, W. V. (1979). *War and Imperialism in Republican Rome, 327–70 B.C.* Oxford: Clarendon Press.

Harris, W. V. (Ed.) (1984). *The Imperialism of Mid-Republican Rome*, University Park, PA. Pennsylvania State University Press.

Headquarters, Department of the Army (2017). *Foot Marches*. Department of Defense.

Hechter, M. and W. Brustein (1980). Regional modes of production and patterns of state formation in Western Europe. *American Journal of Sociology 85*(5), 1061–1094.

Hicks, J. (1969). *A Theory of Economic History*. Oxford: Oxford University Press.

Ho, P. (1975). *The Cradle of the East: An Inquiry into the Indigenous Origins of Techniques and Ideas of Neolithic and Early Historic China, 5000–1000 B.C.* Chicago: University of Chicago Press.

Hodgson, M. G. S. (1954). Hemispheric interregional history as an approach to world history. *Journal of World History 1*(3), 715–23.

Hoffman, P. T. (2015a). What do states do? politics and economic history. *Journal of Economic History 75*, 303–332.

Hoffman, P. T. (2015b). *Why Did Europe Conquer the World?* Princeton, NJ: Princeton University Press.

Horden, P. and N. Purcell (2000). *The Corrupting Sea*. Oxford: Wiley Blackwell.

Howorth, H. H. (1909). The Germans of Caesar. *The English Historical Review 24*(96), 625–658.

Huang, R. (1988). *China: A Macro History*. Armonk, NY: M.E. Sharpe.

Hui, V. T.-b. (2005). *War and State Formation in Ancient China and Early Modern Europe*. Cambridge: Cambridge University Press.

Huntley, B. and H. J. B. Birks (1983). *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 Years Ago*. Cambridge: Cambridge University Press.

Jiang, W., Y. Cheng, X. Yang, and S. Yang (2013). Chinese loess plateau vegetation since the last glacial maximum and its implications for vegetation restoration. *Journal of Applied Ecology 50*(2), 440–448.

Jones, E. L. (1981, 2003). *The European Miracle* (3rd ed.). Cambridge: Cambridge University Press.

Jones, E. T. (2000). River navigation in medieval England. *Journal of Historical Geography 26*(1), 60 – 75.

Kamen, H. (2002). *Spain's Road to Empire*. London: Penguin.

Karayalcin, C. (2008). Divided we stand united we fall: The Hume-North-John mechanism for the rise of Europe. *International Economic Review 49*, 973–997.

Kennedy, H. (2001). *The Armies of the Caliphs: Miltary and Society in the Early Islamic State, Warfare and History*. New York: Routledge.

Kennedy, H. (2008). *The Great Arab Conquests: How Islam Changed the World We Live in*. Boston, MA: Da Capo Press.

Kennedy, P. (1987). *The Rise and Fall of the Great Powers, 1500–1980*. New York: Vintage Books.

Kitamura, S. and N.-P. Lagerlöf (2019). Geography and state fragmentation. *Journal of the European Economic Association*.

Ko, C. Y., M. Koyama, and T.-H. Sng (2018). Unified China; Divided Europe. *International Economic Review 59*(1), 285–327.

Ko, C. Y. and T.-H. Sng (2013). Regional dependence and political centralization in imperial China. *Eurasian Geography and Economics 54*(5–6), 470–483.

Kowarik, I. and S. Körner (2005). *Wild Urban Woodlands: New Perspectives for Urban Forestry*. Springer.

Kremer, M. (1993). Population growth and technological change: One million B.C. to 1990.

Quarterly Journal of Economics 108(3), 681–716.

Kulikowski, M. (2019). *The Tragedy of Empire: From Constantine to the Destruction of Roman Italy*. Cambridge MA: Belknap Press of Harvard University Press.

Lagerlöf, N.-P. (2014). Population, technology and fragmentation: the European miracle revisited. *Journal of Development Economics 108*, 87–105.

Laitin, D. D., J. Moortgat, and A. L. Robinson (2012). Geographic axes and the persistence of cultural diversity. *Proceedings of the National Academy of Sciences 109*(26), 10263–10268.

Lamb, H. (1982). *Climate, History, and the Modern World*. London: Methuen.

Langdon, J. (1993). Inland water transport in medieval England. *Journal of Historical Geography 19*(1), 1 – 11.

Lattimore, O. (1940). *Inner Asian frontiers of China*. New York: American Geographical Society.

Leuschner, C. and H. Ellenberg (2017). *Ecology of Central European Forests: Vegetation Ecology of Central Europe, Volume I*. Cham, Switzerland: Springer.

Liberman, P. (1998). *Does Conquest Pay?: The Exploitation of Occupied Industrial Societies*. Princeton, NJ: Princeton University Press.

Lieberman, V. (2003). *Strange Parallels: Southeast Asia in Global Context, C. 800–1830. vol. 1: Integration on the Mainland*. Cambridge: Cambridge University Press.

Lieberman, V. (2009). *Strange Parallels: Southeast Asia in Global Context, C. 800–1830. Vol. 2, Mainland Mirrors: Europe, Japan, China, South Asia, and the Islands*. Cambridge: Cambridge University Press.

Linn, J. (2012). The Roman grain supply, 442–455. *Journal of Late Antiquity 5*(2), 298–321.

Mackinder, J. H. (1942). The geographical pivot of history. In *Democratic Ideals and Reality*. New York: H. Holt and Company.

Maddison, A. (2001). *The World Economy: A Millenial Perspective*. OECD.

Mayshar, J., O. Moav, and Z. Neeman (2017). Geography, transparency and institutions. *American Political Science Review 111*(3), 622–636.

Mayshar, J., O. Moav, Z. Neeman, and L. Pascali (2016). Cereals, Appropriability and Hierarchy. The Warwick Economics Research Paper Series (TWERPS) 1130, University of Warwick, Department of Economics.

McLaughlin, R. (2010). *Rome and the Distant East: Trade Routes to the Ancient lands of*

Arabia, India and China. London: Continuum.

Mearscheimer, J. J. (2001). *The Tragedy of Great Power Politics*. New York: W. W. Norton & Company.

Mokyr, J. (2007). The market for ideas and the origins of economic growth in eighteenth century Europe. *Tijdshrift voor Sociale en Economische Geschidenis 4* (1), 3–38.

Mokyr, J. (2016). *Culture of Growth*. Princeton, NJ: Princeton University Press.

Montesquieu, C. (1748, 1989). *The Spirit of the Laws*. Cambridge: Cambridge University Press. translated by Anne M. Cohler, Basia C. Miller, and Harold S. Stone.

Murphey, R. (1951). The decline of North Africa Since the Roman occupation: Climatic or human? *Annals of the Association of American Geographers 41* (2), 116–132.

Nath, P. (2019). *Climate of Conquest: War, Environment, and Empire in Mughal North India*. Oxford University Press India.

Norkus, Z. (2019). The economic output growth of Baltic countries in 1913–1938: a quantitative cross-country comparison. *Journal of Baltic Studies 50* (2), 183–203.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics 94* (1), 20–36.

Olsson, O. and G. Hansson (2011). Country size and the rule of law: Resuscitating Montesquieu. *European Economic Review 55* (5), 613–629.

Parker, G. (1988). *The Military Revolution: Military Innovation and the Rise of the West, 1500–1800*. Cambridge: Cambridge University Press.

Parker, G. (2004). *The Army of Flanders and the Spanish Road, 1567-1659: The Logistics of Spanish Victory and Defeat in the Low Countries' Wars* (2nd ed.). Cambridge University Press.

Parker, G. (2013). *Global Crises: War, Climate Change and Castrastrophe in the Seventeenth Century*. New Haven: Yale University Press.

Perdue, P. C. (2005). *China Marches West: The Qing Conquest of Central Eurasia*. Belknap Press of Harvard University Press.

Pirenne, H. (1925). *Medieval Cities: Their Origins and the Revival of Trade*. Princeton, NJ: Princeton University Press. Translated by Frank D. Halsey.

Pounds, N. J. G. and S. S. Ball (1964). Core-areas and the development of the European states system. *Annals of the Association of American Geographers 54* (1), 24–40.

Ramankutty, N., J. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: Current patters and sensitivity to possible climate change. *Global Ecology and Biogeography 11*, 377–392.

Reale, O. and P. Dirmeyer (2000). Modeling the effects of vegetation on Mediterranean climate during the Roman classical period: Part i: Climate history and model sensitivity. *Global and Planetary Change 25*(3), 163–184.

Rickman, G. E. (1980). The grain trade under the Roman Empire. *Memoirs of the American Academy in Rome 36*, 261–275.

Root, H. L. (2017). Network assemblage of regime stability and resilience: Comparing Europe and China. *Journal of Institutional Economics 13*(3), 523–548.

Ropa, A. and T. Dawson (Eds.) (2020). *The Horse in Premodern European Culture*, New York. De Gruyter.

Rosenberg, N. and L. J. Birdzell (1986). *How the West Grew Rich: The Economic Transformation of the Industrial World*. New York: Basic Books.

Sachs, J. D. (2001). Tropical Underdevelopment. NBER Working Papers 8119, National Bureau of Economic Research.

Scheidel, W. (2019). *Escape from Rome*. Princeton, NJ: Princeton University Press.

Scott, J. C. (2017). *Against the Grain*. Princeton, NJ: Princeton University Press.

Sellmann, J. (2002). *Timing and Rulership in Master Lü's Spring and Autumn Annals*. Albany: State University of New York Press.

Sha, X. (1972). *Essays in Geography [dili xue lunwen ji]*. Taipei: Commerical Press.

Skinner, G. W. (1977). *The City in Late Imperial China*. Stanford: Stanford University Press.

Sng, T.-H., P. Z. Chia, C.-C. Feng, and Y.-C. Wang (2018). Are China's provincial boundaries misaligned? *Applied Geography 98*, 52–65.

Tacitus (1877). *The Agricola and Germania*. London: Macmillan. Translated by J. Church and W. J. Brodribb.

Talhelm, T., X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, and S. Kitayama (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science 344*(6184), 603–608.

Tapper, R. (1997). *Frontier Nomads of Iran: A Political and Social History of the Shahsevan*. Cambridge: Cambridge University Press.

Tilly, C. (1990). *Coercion, Capital, and European States, AD 990–1990*. Oxford: Blackwell.

Turchin, P. (2009). A theory for formation of large empires. *Journal of Global History 4*, 191–217.

Turchin, P., J. Adams, and T. Hall (2006). East-west orientation of historical empires and modern states. *Journal of World-Systems Research 12*(2), 219–229.

Turchin, P., T. E. Currie, E. A. L. Turner, and S. Gavrilets (2013). War, space, and the evolution of old world complex societies. *Proceedings of the National Academy of Sciences 110*(41), 16384–16389.

Usher, D. (1989). The dynastic cycle and the stationary state. *American Economic Review 79*(5), 1031–1044.

Voigtländer, N. and H.-J. Voth (2013). The three horsemen of riches: Plague, war, and urbanization in early modern Europe. *Review of Economic Studies 80*, 774–811.

Waltz, K. N. (1979). *Theory of International Politics*. Long Grove, , IL: Waveland Press In.

Whittlesey, D. (1944). *The Earth and the State: A study of Political Geography*. New York: Henry Holt and Company.

Wilson, J. K. (2012). *The German Forest: Nature, Identity, and the Contestation of a National Symbol, 1871–1914*. Toronto: University of Toronto Press.

Zhang, D. D., C. Jim, G. Lin, Y.-Q. He, J. J. Wang, and H. F. Lee (2006). Climatic change, wars and dynastic cycles in China over the last millennium. *Climatic Change 76*(3–4), 459–477.

Zheng, P. (1984). Livestock breeds of china. Animal Production and Health Papers 46, Food and Agriculture Organization.

Zon, R. (1920). Forests and human progress. *Geographical Review 10*(3), 139–166.

A Online Appendix: Extensions

We include now additional details about the several extensions of our baseline model that we described in Section 6 of the main text.

A.1 The Eurasian Steppe

In premodern times, horses were an invaluable military asset, a powerful war machine likened to modern tanks (Ropa and Dawson, 2020). Control of horses allowed some states to develop highly mobile cavalry troops that easily outflank and break infantry units.

Horses were a location-specific resource. Equine domestication began in the Eurasian steppe and steppe horses were especially stocky and vigorous (Zheng, 1984). Moreover, only an extensive expanse of grassland could support a dense concentration of horses and the high concentration of skills on breeding and riding them. Thus, states close to the Eurasian steppe fielded larger and better cavalry forces, thereby giving them a distinctive advantage in war (Barfield, 1989; Gat, 2006; Turchin et al., 2013).

To capture this idea, we introduce a parameter ψ into the contest function (3) and rewrite it as:

$$\pi_{i,win} = \frac{\psi_i \cdot Y_{i,t}}{(\psi_i \cdot Y_{i,t} + \psi_j \cdot Y_{j,t}) \times (1 + max\{\Theta \cdot \mathbf{x}_k, \Theta \cdot \mathbf{x}_{\overline{k}}\}} \tag{5}$$

where $\psi_i = 3$ if regime i originates as a steppe cell and $\psi_i = 1$ otherwise. We define a steppe cell as one that is within 100 kilometers of the Eurasian steppe. The extension assumes that a regime that originates from a steppe cell is more proficient in war because they have better access to horses and the associated organizational and military technologies. In addition, to account for the role of the steppe as a "highway of grass" that facilitated movement (Frachetti, 2008, 7), we reduce the obstacle value $\Theta \cdot x_k$ by two if cell k is a steppe cell.

We simulate this extension for 19 times and plot the Herfindahl index of the median simulation in Figure 17. We can see that even by making steppe cells three times better at conflict (a difference that is as big as one can plausibly assume) and lowering the geographical barriers by half, the main results of the paper remain unchanged.

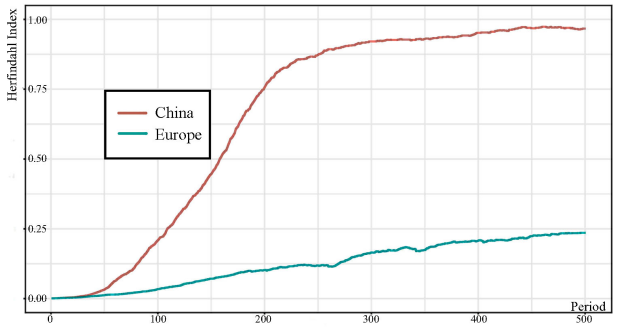


Figure 17: Median Plot for 19 simulations with steppe.

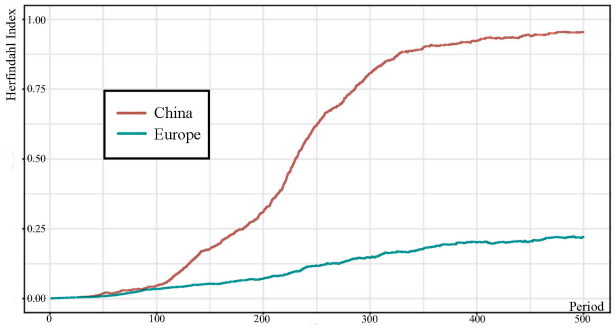


Figure 18: Median Plot for 19 simulations with rivers.

A.2 Major Rivers

A river connects its upstream and downstream areas and fosters their interdependence. In particular, China was a riverine civilization that depended upon its rivers to serve as its primary means of transportation until the early twentieth century (Skinner, 1977). This natural water system was complemented by the Great Canal, built in the Sui dynasty (581-618 CE), which connected the Yangzi and Yellow rivers. However, a wide river also impedes movement between its two banks, especially during times of war. In some episodes of Chinese history, regimes in southern China successfully built defense lines along the Huai and Yangzi rivers to deter invasions from the north for prolonged periods (Sng et al., 2018).

We capture these dual roles of rivers by setting the obstacle value $\Theta \cdot x$ to zero when cells along the same river came into contact and increasing the obstacle value $\Theta \cdot x$ by two when a riverine cell comes into conflict with a non-riverine cell. We simulate this extension for 19 times and plot the Herfindahl index of the median simulation in Figure 18. The main results remain unchanged, with only a slightly slower unification of China.

Historians have also observed that regular flooding along the Yellow River gave North China a head start in state development through at least two channels. First, flooding replenished the soil and allowed agriculture to remain sustainable even with limited farming knowledge (Ho, 1975). Second, flood management problems increased tensions between upstream and downstream states along the Yellow River and accelerated the emergence of a unified regime through intense warfare (Huang, 1988). Our model does not consider these extensions. Doing so is likely to increase the pace of unification in China further.

A.3 Climatic Shocks and Dynastic Cycles

History is contingent. An independent event could interact with existing conditions to trigger a chain of unanticipated consequences. Absent the event or change an existing condition, and history may develop in a completely different direction. Our model allows for contingency as the outbreak and outcome of wars in the simulations are not predetermined. They are random events with probabilities assigned based on existing conditions. If and when a cell emerges to dominate its neighbors is neither a fluke nor destiny. That is why we report a distribution of events coming from our simulations.

To further account for contingency in history and to study the role of large exogenous events such as natural disasters or incompetent leaders on the rise and fall of polities, we modify the model by extending our simulation range to 4000 periods and allowing random shocks to occur each period. Specifically, we consider two kinds of negative shocks: a general shock, which directly affects all polities (e.g., such as the Little Ice Age), and a regime-specific shock (e.g., the ascension of a weak ruler such as Charles II of Spain, r. 1665-1700, or the Chongzhen Emperor, r. 1627-1644). Thus, we call these events "climatic shocks" and "dynastic cycles."

We set the probability of a general shock occurring at $\frac{1}{1000}$ and the probability of a regime specific shock occurring at $\frac{1}{300}$. These values imply that, on average, a general shock will occur once every 1000 periods, and each polity will independently experience a specific shock once every 300 periods. When a shock occurs, the regime disintegrates into its constituent cells. These two frequencies are somewhat irrelevant, since nearly all our parameter values are time-invariant, and they just determine how often we will observe a collapse of existing state systems. We could also have desintegrations into larger polities, without changing much the results.

Figure 19 depicts the Herfindahl index from a single simulation for both China and Europe. For China, we observe periods of sustained unification interrupted by periods of disunity, resembling the patterns of dynastic rise and fall so often depicted in Chinese historiography. Some periods of unified rule are short-lived; others persist for many periods. By contrast, political cycles are relatively muted in Europe. Europe never achieves full unification in this realization of the model. There are periods of heightened military conquests that rest on one state becoming hegemonic in Europe, but these are always transitory; political fragmentation remains persistent.

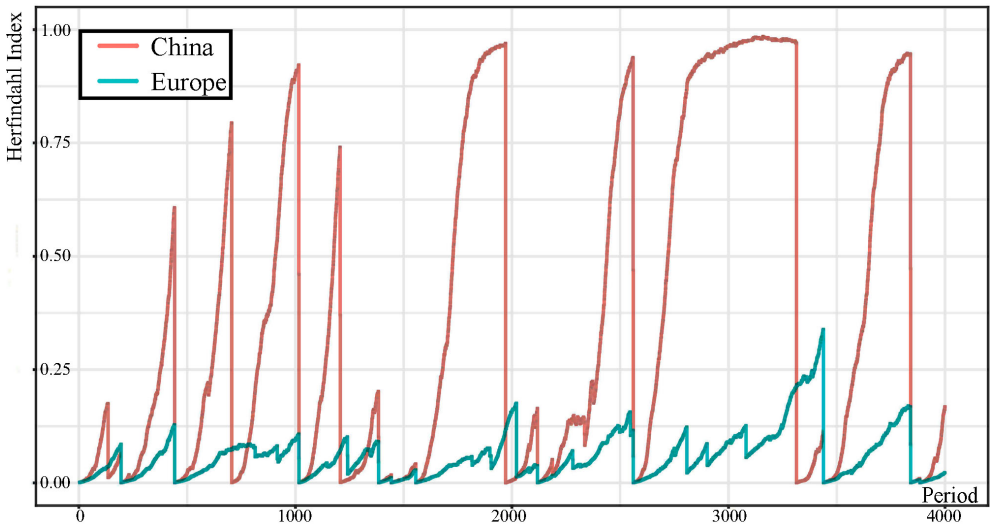


Figure 19: This figure depicts one realization of a 4,000 periods simulation where we allow for both regime-specific shocks (prob. $0.0\dot{3}$) and general shocks (prob. 0.001) occurring.

A.4 The Mediterranean Sea

In our benchmark model, we do not allow conflicts to take place across oceans and wide seas. Only land cells and narrow sea channels are traversable. This ensures that conflicts can take place, for example, between the British Isles and mainland Europe or across the Korea Strait. But it prevents polities separated by wide bodies of water from coming into direct conflict.

In general, this is a realistic depiction of the world before 1500. However, it may not the best way to model the Mediterranean Sea. As the Mediterranean is nearly landlocked, its tides are comparatively weak. Consequently, although shipwrecks remained common in the Mediterranean even as late as the 1500s (Braudel, 1949), it was probably easier to transverse across the Mediterranean than across other equivalent seas.

This is potentially a concern for our analysis because at least one European state, the Roman Empire, was built upon control of the Mediterranean. To address this concern, we modify the model to allow the Mediterranean to be traversable and simulate the extension for 19 times. We find that simply making the Mediterranean traversable does not increase the likelihood of either a European empire or a Mediterranean empire emerging.

However, if we account for the fact that building a navy is capital intensive and allow only polities spanning 100 cells or more to bid for control of the Mediterranean Sea, then we do observe the rise of a Mediterranean empire in the simulations (Figure 20). Yet even in this case,

we do not detect a perceptible difference in the pace of unification in Europe ex-Mediterranean.

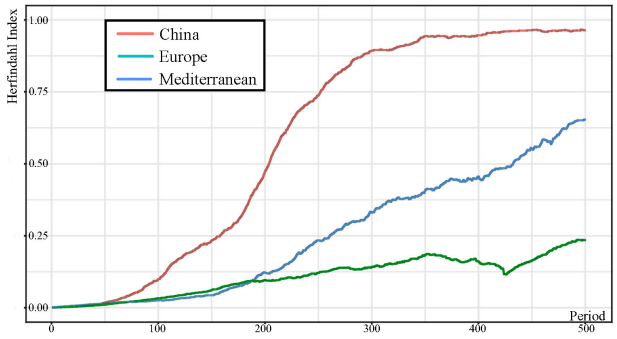


Figure 20: "Median" Simulations for China, Europe, and the Mediterranean region, based on population in 0 CE.

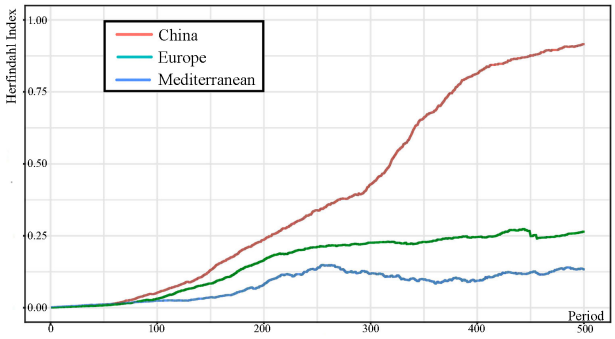


Figure 21: "Median" Simulations for China, Europe, and the Mediterranean region, based on population in 1000 CE.

Ancient historians have also suggested that the Roman economy was built on its access to the extremely productive North Africa agriculture. Climatic conditions during the Classical period ensure that North Africa was wetter than today (Murphey, 1951; Reale and Dirmeyer, 2000). As a consequence, the provinces of Egypt and Africa (modern-day Algeria, Morocco, and Tunisia) were the "bread baskets" of the empire.[30]

Ecological degradation then led to a significant fall in agricultural activities in North Africa. To verify the effect of ecological degradation on empire formation in the Mediterranean region, we replace population density in 0 CE with population density in 1000 CE as our measure of resource availability. As the median plot in Figure 21 illustrates, while we continue to observe China unifying faster than Europe, we no longer see a heightened likelihood of empire formation around the Mediterranean. This suggests that the rise of the Roman empire might have been predicated on a temporal confluence of factors and explains why no other power has followed Rome's steps.

A.5 State Formation Across Eurasia

Beyond China and Europe, to what extent can our model explain broader patterns of political fragmentation across Eurasia? In Figure 22, we use the benchmark model to compute the probability of a large state –defined as a polity controlling 600 cells or more (approximately the

[30] See Rickman (1980). According to Linn (2012, 305–306), "[S]ince the first century BCE, whenever Rome was shut off from North African grain, a shortage typically had ensued . . . All these instances demonstrate two facts about the relationship between North Africa and the city of Rome: (1) North Africa was the lifeline for the city of Rome; (2) warfare commonly led to a food crisis in Rome because of transport blockages".

combined size of Turkey and Iraq)– *originating from* Europe, East Asia, South Asia, Southeast Asia, or the Middle East.

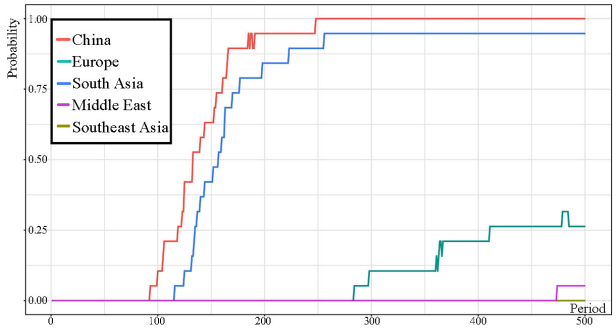


Figure 22: Probability of a large state arising in China, Europe, India, Middle East, or Southeast Asia (Basic Model).

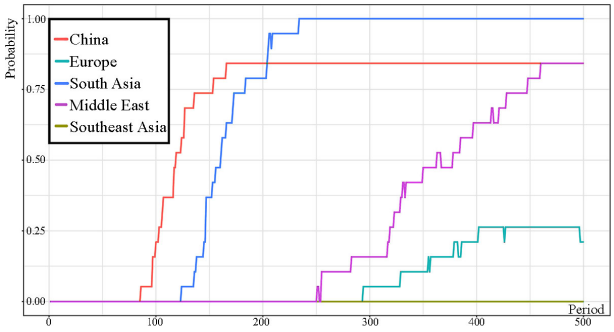


Figure 23: Probability of a large state arising in China, Europe, India, Middle East, or Southeast Asia (Incorporating the Steppe Effect).

Figure 22 shows that the formation of large states is most pronounced in China. We also see large states emerging regularly in India. In the case of India, closer scrutiny reveals that these large states always originate from the north. This is consistent with what we observe historically. Large polities or empires often emerged in North India. But until the Mughal empire and the British Raj, they did not come close to unifying the Indian sub-continent. Consistent with historical observations, in our simulations, large states seldom arise in Europe, and they never do so in Southeast Asia. However, the emergence of large states in the Middle East is rare in our baseline simulations, which is not consistent with history.

Historians suggest that the threat of invasion by, and continuous contact with, the steppe has been a constant factor in Middle Eastern state formation since the domestication of horses for use as calvary. The first empire to span the entire Middle East, the Achaemenid Empire founded by Cyrus the Great c. 550 BCE, was created by a military that relied on steppe horsemen and horse archers. Through its history, Persian state formation has been shaped by the steppe: "Nomadism and transhumant pastoralism have been a consistent element in Iranian history over millennia, and recent scholarship has documented long continuities in the way nomads coexisted with settled urban and agricultural communities in the more populous regions of the Iranian plateau" (Kulikowski, 2019, 236). In fact, this pattern persists into the 20th century (Tapper, 1997).

Thus, in Figure 23, we incorporate the steppe extension discussed in Section A.1. With this extension, the Middle East now experiences an intermediate level of empire formation. This is

more or less in line with what we observe historically. The Middle East was rarely dominated by a single centralized empire like China nor comprising a relatively stable set of medium-sized states like Europe. Specifically, during the past 2500 years, a single empire ruled the majority of the Middle Eastern region for approximately 400 years or 18.5% of the time. For approximately 41% of the time, the Middle East was fragmented. For the remaining 40% of the last 2500 years, the Middle East had been dominated by two powers (i.e., the Roman and Parthian/Sassanian empires or the Ottoman/Safavid/Qajar Persia).[31]

One episode that is not easily generated by our model is the rapid Arab conquests of the 7th century. This is not surprising, as the episode has long been regarded as *sui generis* by historians.[32] The success of the Arab conquest of much of the Byzantine and the entire Persian empires in the mid-7th century has been attributed to three factors: (i) the weaknesses of the Byzantine and Persian empires due to war and plague-induced population losses; (ii) religious divisions with the Roman empire (between the dominant Chalcedonian Church and Monophysitism); (iii) the religious cohesion or *asabiyyah* of the first Muslims. These factors, particularly the role of religion are not in our model, but could be added in future research.[33]

As a final observation, in the steppe-enriched simulations, the speed of a large state emerging in China is faster than what Figure 23 suggests. In some simulations, we observe polities originating from Mongolia or Manchuria expanding into China and eventually unifying it. These events are not reflected in Figure 23, which counts only large states originating from China and the other four regions. In this regard, the steppe extension adds realism to our model, as two of the nine Chinese dynasties in Table 3, Yuan and Qing, originated north of the Great Wall.[34]

[31] This calculation is obtained as follows. Between 500 BCE and 2000 CE, the following empires have ruled the majority of the Middle Eastern region: Achaemenid Empire/Alexander the Great, 500–323 BCE; Umayyad Caliphate and Abbasid Caliphates, 640–880. The following periods have seen two major powers dominate much of the region: 30 BCE–640, Rome and Parthian/Sassanian Persia; 1520–1914, Ottoman Empire and Safavid/Qajar Persia. The remaining periods were characterized by a varying degree of political fragmentation.

[32] For example, Kennedy (2001, 2) observes that "Despite the mass of words, the full explanation for Muslim victory still eludes us."

[33] The importance of religion in inspiring these conquests was highlighted by the earliest Islamic tradition (see the discussion in Kennedy, 2008). The impact of religious dissension in the Eastern territories of the Byzantine empire is likewise discussed in at length in the historiography.

[34] We report unifications of China proper. Hence we exclude the Shang and Western Zhou and unifications of northern China under the Wei and Jin dynasties. These also all stemmed from Northern China.

Table 3: Major Unifications of China

Dynasty	Period	Capital Region
Qin	221–206 BCE	Northwest
Han	202 BCE–220	Northwest, North China
Western Jin	280–316	North China
Sui	581–618	Northwest, North China
Tang	618–907	Northwest, North China
Northern Song	960–1127	North China
Yuan	1206–1368	North China
Ming	1368–1644	Lower Yangzi, North China
Qing	1644–1912	North China

Source: Skinner (1977); Huang (1988); Turchin (2009); Scheidel (2019).