Unified China and Divided Europe

BY CHIU YU KO, MARK KOYAMA, TUAN-HWEE SNG[1]

National University of Singapore, Singapore; George Mason University, U.S.A.;

National University of Singapore, Singapore

Abstract

This paper studies the causes and consequences of political centralization and fragmentation in China and Europe. We argue that a severe and unidirectional threat of external invasion fostered centralization in China while Europe faced a wider variety of smaller external threats and remained fragmented. Political centralization in China led to lower taxation and hence faster population growth during peacetime compared to Europe. But it also meant that China was more vulnerable to occasional negative population shocks. Our results are consistent with historical evidence of warfare, capital city location, tax levels, and population growth in both China and Europe.

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

---

[1]This paper has benefited from presentations at Academia Sinica, Shandong University, Shanghai University of Finance and Economics, Chinese University of Hong Kong, Hitotsubashi University, American University, Hong Kong University, Carlos III Madrid, the Osaka Workshop on Economics of Institutions and Organizations, the Washington Area Economic History Workshop at George Mason University, the 2014 EHS Annual Conference, ISNIE 2014, the 2014 EHA Annual Meeting, ESNIE 2015, and the 2015 ASSA meetings. We are grateful for comments from Jesús Fernández-Villaverde of the Editorial Board, three anonymous referees, Warren Anderson, Christophe Chamley, Jiahua Che, Mark Dincecco, Rue Esteves, James Fenske, Nicola Gennaioli, Boris Gershman, Noel Johnson, Yi Lu, Mona Luan, Debin Ma, Andrea Matranga, Joel Mokyr, Michael Powell, Nancy Qian, Jared Rubin, Noah Smith, Yannay Spitzer, Alex Tabarrok, Alex Teytelboym, Zhigang Tao, Denis Tkachenko, Melanie Meng Xue, Se Yan, Helen Yang, and many others. We thank Peter Brecke for sharing his Conflict Catalog Dataset and Ryan Budny, Pei Zhi Chia, Jane Perry, and Jenisa Rumdech for research support. Financial support from Singapore Ministry of Education Academic Research Fund Tier 1 FY2014-FRC3-002 and George Mason University Summer Research Grant is gratefully acknowledged. Please address correspondence to: Tuan-Hwee Sng, Department of Economics, National University of Singapore, AS2 Level 6, 1 Arts Link, Singapore 117570. E-mail: tsng@nus.edu.sg.

JEL Classification: H2, H4, H56; N30; N33; N35; N40; N43; N45

Running Head: Unified China and Divided Europe

1 Introduction

Since Montesquieu, scholars have attributed Europe's success to its political fragmentation (Montesquieu, 1989; Jones, 2003; Mokyr, 1990; Diamond, 1997). Nevertheless, throughout much of history, the most economically developed region of the world was China, which was typically a unified empire. This contrast poses a puzzle that has important implications for our understanding of the origins of modern economic growth: Why was Europe perennially fragmented after the collapse of Rome? Why was political centralization an equilibrium for most of Chinese history? Can this fundamental difference in political institutions account for important disparities in Chinese and European growth patterns?

This paper proposes a unified framework based on Eurasian geography to (a) help explain the different political equilibria in China and Europe, and (b) explore the economic consequences of political centralization and fragmentation. Historically, Europe faced periodic invasions from Scandinavia, Central Asia, the Middle East, and North Africa. By contrast, while China was relatively isolated from the rest of Eurasia, it had to confront a severe recurring threat on its northern frontier due to its relative proximity to the Eurasian steppe. We develop a Hoteling-style model to show that a severe unidirectional external threat undermines the fiscal viability of small states and thus provides an impetus towards political centralization. Meanwhile, multi-sided external threats favor a more decentralized approach to defense and reduce the likelihood for a centralized empire to survive and prosper. We argue that China's perennial steppe problem was an important driver of its recurring unification, while the presence of multi-sided threats in Europe, especially in the first millennium AD, doomed the Roman empire and helped thwart subsequent efforts to resuscitate political unification in Europe.

Our model also suggests that the different political paths that China and Europe took had important economic consequences. Political centralization allowed China to avoid wasteful interstate competition. This enabled it to enjoy more rapid economic and population growth during peacetime. Meanwhile, while taxes were higher in Europe than in China, the presence of multiple states to protect different parts of the continent meant that Europe was more robust to both known threats and unexpected negative shocks, and therefore less susceptible to the kind of growth reversals that Aiyar et al. (2008) have highlighted.

To test the mechanisms identified in our model, we use time series analysis to show that an increase in the frequency of nomadic attacks on China is associated with more political centralization in historical China. Our estimates suggest that each additional nomadic attack per decade was associated with a 6.3–8.3 percentage point higher probability of political unification in the long run. Given that China experienced an average of 2.5 nomadic attacks per decade, this effect is substantial. We also use our theory in conjunction with narrative and qualitative evidence to discuss the disintegration of Rome and why the Carolingians and the Ottonians failed in their attempts to rebuild a Europe-wide empire. Finally, we provide evidence supporting the predictions of the model concerning the location of capital cities, taxation, and population growth.

Our paper relates to several strands of literature. Our theoretical framework builds on the research on the size of nations originated by Friedman (1977) and Alesina and Spolaore (1997, 2003). In particular, our emphasis on the importance of external threats is related to the insights of Alesina and Spolaore (2005), who study the role of war in shaping political boundaries. It is also related to Levine and Modica (2013), who propose a theory on the emergence (or absence) of hegemonic rule.[2] In examining the causes of political fragmentation and centralization in China and Europe, we

[2]In their canonical model, Alesina and Spolaore (1997) explain the size of nations in terms of a trade-off between economics of scope and heterogeneous preferences. One insight of the model is that external threats lead to the consolidation of countries (Alesina and Spolaore, 2003). Using an evolutionary setting, Levine and Modica (2013) argue that the presence of strong outsiders would instead weaken the tendency toward hegemony (i.e., empire). Our model suggests that external threats can indeed foster political centralization in some situations and political fragmentation in others depending on the threat nature (magnitude and direction).

build on earlier work that points to the role of geography, such as Diamond (1997), and on the work of many historians who stress how the threat of nomadic invasion from the steppe shaped Chinese history (Lattimore, 1940; Grousset, 1970; Huang, 1988; Barfield, 1989; Gat, 2006; Turchin, 2009).

By developing a new framework to help explain why Europe was persistently fragmented, we complement the literature that emphasizes the positive economic consequences of European political fragmentation, which include promoting economic and political freedom (Montesquieu, 1989; Pirenne, 1925; Hicks, 1969; Jones, 2003); encouraging experiments in political structures and investments in state capacity (Baechler, 1975; Cowen, 1990; Tilly, 1990; Hoffman, 2012; Gennaioli and Voth, 2015);[3] intensifying interstate conflicts and thereby promoting urbanization (Voigtländer and Voth, 2013b);[4] and fostering innovation and scientific development (Diamond, 1997; Mokyr, 2007; Lagerlof, 2014).[5] Our analysis is also related to the rise of state capacity in Europe and the weakening of the Chinese state after 1750 (Dincecco, 2009; Dincecco and Katz, 2016; Johnson and Koyama, 2013, 2014a,b; Sng, 2014; Sng and Moriguchi, 2014), and to recent research that emphasizes other aspects of Europe's possible advantages in the Great Divergence such as the higher age at first marriage than the rest of the world (Voigtländer and Voth, 2013a); public provision of poor relief versus reliance on clans as was the case in China (Greif et al., 2012); institutions that were less reliant on religion (Rubin, 2011);

[3]Baechler (1975, 74) observes that 'political anarchy' in Europe gave rise to experimentation in different state forms. Cowen (1990) argues that interstate competition in Europe provided an incentive for early modern states to develop capital markets and pro-market policies. Tilly (1990) studies the role capital-intensive city states played in shaping the emergence of nation states in Europe. Hoffman (2012) uses a tournament model to explain how interstate competition led to military innovation in early modern Europe. Gennaioli and Voth (2015) show that the military revolution induced investments in state capacity in some, but not all, European states.

[4]Voigtländer and Voth (2013b) argue that political fragmentation interacted with the Black Death so as to shift Europe into a higher income steady-state Malthusian equilibrium.

[5]Diamond (1997, 414) argues that 'Europe's geographic balkanization resulted in dozens or hundreds of independent, competing statelets and centers of innovation' whereas in China 'a decision by one despot could and repeatedly did halt innovation.' Mokyr (2007, 24) notes that 'many of the most influential and innovative intellectuals took advantage of . . . the competitive "states system."' Lagerlof (2014) develops a growth model that emphasizes the benefits to scale in innovation under political unification and a greater incentive to innovate under political fragmentation.

greater human capital (Kelly et al., 2014); and higher social status for entrepreneurs and inventors (McCloskey, 2010).

Perhaps the closest argument to ours is that of Rosenthal and Wong (2011) who argue that political fragmentation led to more frequent warfare in medieval and early modern Europe, which imposed high costs but also lent an urban bias to the development of manufacturing and more capital-intensive forms of production. Like them, we emphasize that political fragmentation was costly for Europe, but we develop a different argument based on the observation that the costs of political collapse and external invasion were particularly high in China. Theoretically and empirically, we show that the Chinese empire could indeed have been more conducive to Smithian economic expansion during stable periods as Rosenthal and Wong claim, but we also note that it was less robust to negative shocks, and this greater volatility of population and economic output was a major barrier to sustained economic growth in China before 1800.

Clearly, the political development of China and Europe over the past two millennia was subject to numerous complex forces. The mechanism that we highlight, while important, was not the only one at work. A more complete examination of China's tendency toward unification and Europe's enduring fragmentation must incorporate other explanations such as topology, culture, and institutions. While consideration of space and focus prevents us from conducting such an exercise, we discuss alternative and complementary hypotheses in greater detail in Section 6.

The rest of the paper is structured as follows. Section 2 provides historical evidence that characterizes (i) the extent to which China was politically unified and Europe fragmented throughout their respective histories, and (ii) the degree to which both China and Europe were threatened by external invasions. In Section 3 we introduce a model of political centralization and decentralization. Section 4 provides empirical evidence to support our hypothesis that a severe threat from the Eurasian steppe discouraged political fragmentation in China. In Section 5, we show that our model provides a coherent framework that can help to explain the choice of capital cities, differential levels of taxation, and population growth patterns in historical China and Europe. Section 6 presents alternative hypotheses, and Section 7 concludes.

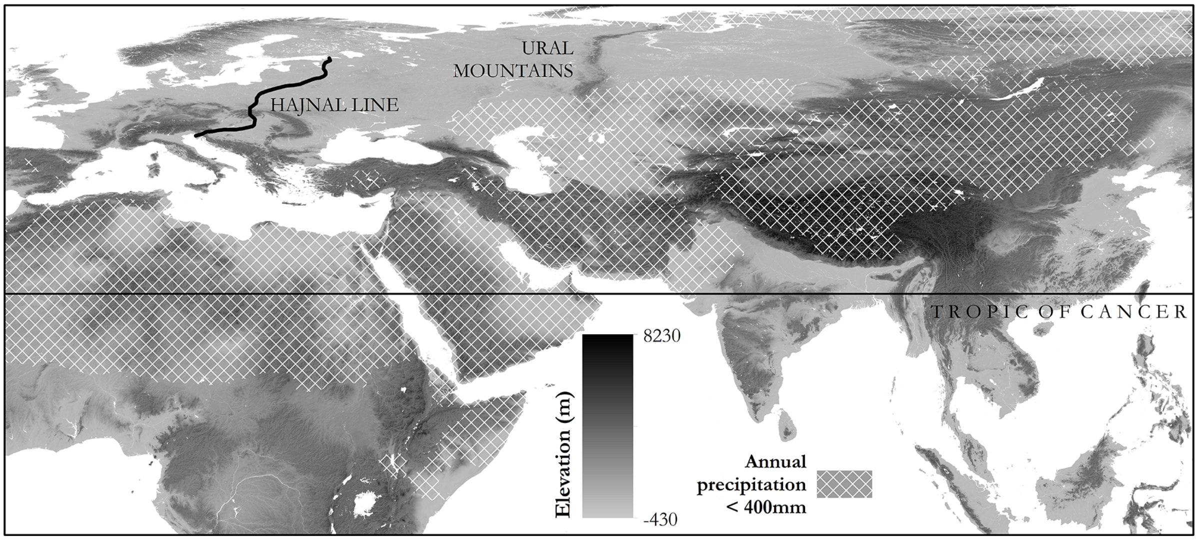


Figure 1: Two ends of Eurasia: Western Europe (i.e., west of the Hajnal line) and China proper (i.e., the agricultural zone bounded by the 400mm isohyet line in the north, the Himalayas and other mountain ranges in the west, tropical rainforests in the south, and the Pacific Ocean in the east). Notice that China is relatively isolated except for its northern frontier. By contrast, Europe is connected to the rest of Eurasia and Africa in multiple directions.

2 The Puzzle: Unified China and Divided Europe

Unit of analysis. States and state systems first emerged in areas suitable for settled agriculture where cereal grain surpluses were available to form the basis of taxation (Childe, 1936; Carneiro, 1970; Mayshar et al., 2015). In this paper, we focus on the two continuous agricultural zones at either end of Eurasia, China and Europe (Figure 1). For Europe, we focus on its western portion, or the area west of the Hajnal line.[6] Meanwhile, we equate China with China proper, an area bounded by the Pacific Ocean to its east, the thick tropical rainforests of Indochina to its south, huge mountain ranges—including the Himalayas—to its west, and the Great Wall to its north. Although the Great Wall was manmade, it overlaps largely with the 400mm isohyet line, which approximates the northern limit of rainfed agriculture (Brouwer and Heibloem, 1986). In other words, the Great Wall delineates

[6]Our analysis is unchanged if we consider instead the Ural mountains as the eastern boundary of 'Europe.' Indeed, our framework provides a potential explanation as to why empires were more frequent in Eastern Europe than in Western Europe.

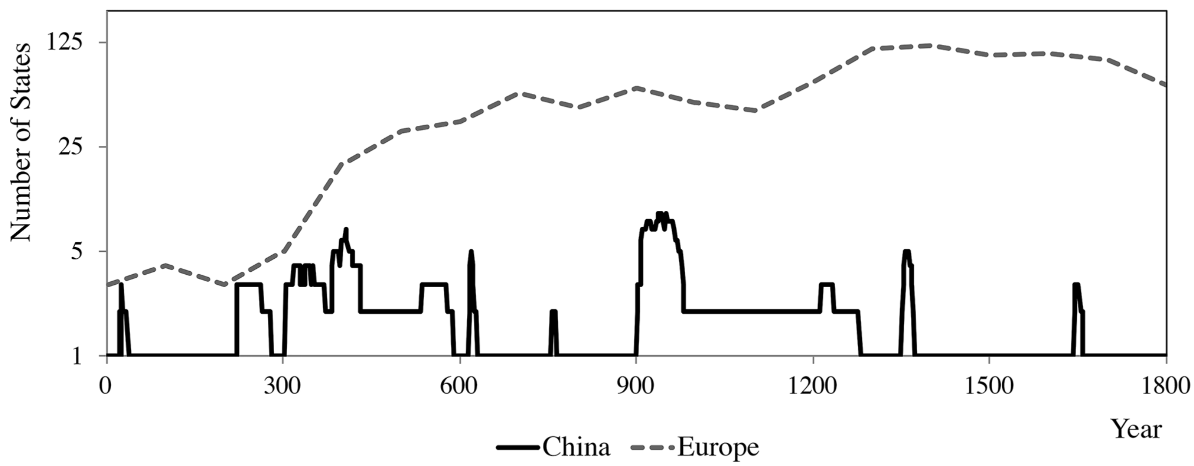


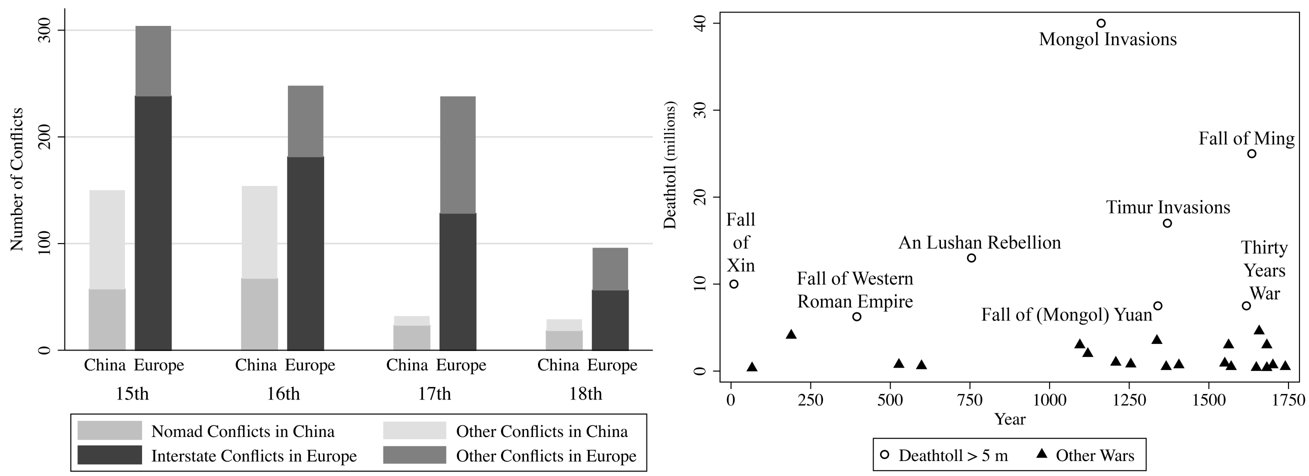
Figure 2: Number of states in China and Europe, AD 0–1800 (Nussli, 2010; Wei, 2011).

the ecological divide between the steppe nomads of Central Asia and the agricultural population in the river basins of China. 'China' and 'Europe' are comparable in size: China proper covers a land area of 2.8 million square kilometers, while Western Europe has slightly more than 2.5 million square kilometers.

Patterns of unification and fragmentation. Chinese historical records indicate that fewer than 80 states ruled over parts or all of China between AD 0 and 1800 (Wilkinson, 2012). Nussli (2010) provides data on the sovereign states in existence at hundred year intervals in Europe. Figure 2 plots the number of sovereign states in China and in Europe for the preindustrial period. There have always been more states in Europe than in China throughout the past two millennia; in fact, since the Middle Ages, there have been an order of magnitude more states in Europe than in China.[7]

The Chinese first established a unitary empire in the third century BC, before Rome's dominance of the Mediterranean (Elvin, 1973; Fukuyama, 2011). Moreover, the Chinese empire outlasted Rome. Although individual dynasties rose and fell, China as an empire survived until 1912. Between AD 0

[7]The Nussli (2010) data does not capture all political entities in Europe since that number is unknown—there may have been as many as 1000 sovereign states within the Holy Roman Empire alone—but it does record the majority of them (Abramson, 2013). By contrast, the Chinese dynastic tables are well known and the potential for disagreement is immaterial for our purposes. We count only sovereign states. Including vassal states would further strengthen the argument.



(a) Number of Violent Conflicts (Brecke, 1999) (b) Largest Wars By Number of Deaths (White, 2013)

Figure 3: The Nature, Frequency, and Intensity of Warfare in China and Europe

and 1800, the landmass between the Mongolian steppe and the South China Sea was ruled by one single authority for 1008 years (Ko and Sng, 2013).

In comparison, after the fall of the Roman Empire, Europe was characterized by persistent political fragmentation—no subsequent empire was able to unify a large part of the continent for more than a few decades. The number of states in Europe increased from 37 in AD 600 to 61 in 900, and by 1300 there were 114 independent political entities. The level of political fragmentation in Europe remained high during the early modern period.

Patterns of Warfare. It is well established that interstate warfare, or military conflicts between sedentary societies, was more common in Europe, while military conflicts with nomads from the Eurasian steppe featured more prominently in China (e.g., Rosenthal and Wong, 2011; Hoffman, 2015). Figure 3a, derived from Brecke (1999), lends further support to this observation.[8] According to Chaliand (2005), out of the seven major waves of nomadic invasions witnessed in Eurasia since the first century, China was involved in six, while Europe was affected only twice (See Appendix A.1).

Figure 3b provides another intriguing—and hitherto overlooked—observation: the most violent

[8]This dataset is widely used by researchers in political science and economics (e.g., Iyigun, 2008; Besley and Reynal-Querol, 2014; Iyigun et al., 2015).

wars of the preindustrial period occurred in Asia, and particularly in China. While warfare might have been less common in China, it was more costly than in Europe. Only two wars with estimated death tolls in excess of five million are recorded for Europe before 1750, compared with five for China.[9] Wars in China such as the An Lushan Rebellion, the Mongol invasions, and the Ming-Manchu transition were extremely costly because they involved the collapse or near collapse of entire empires. Notably, each of these wars had a nomadic dimension.[10] By contrast, warfare in Europe was endemic, but rarely resulted in large-scale socio-economic collapse. The only European war that matched the death tolls of the worst conflicts in Chinese history was the Thirty Years War.

We argue that the patterns in Figures 2 and 3 are connected: while the immediate effect of a nomadic invasion was to create chaos and weaken sedentary regimes, in the long run the presence of a severe steppe threat along China's northern border constituted a centripetal force that regularly pressed the constituent regions of China toward unification; meanwhile, the foremost concerns of European regions were the idiosyncratic threats and problems that they individually faced, which in turn discouraged the rise of empires in Europe.

The Eurasian Steppe Throughout its history, China was repeatedly invaded by the nomadic and semi-nomadic people north of its borders: Hu, Xiongnu, Xianbei, Juan-juan, Uyghurs, Khitan, Jurchen, Mongols, and Manchus (Grousset, 1970; Barfield, 1989; Di Cosmo, 2002; Chaliand, 2005). This was an inevitable outcome of China's proximity to the grasslands of Central Asia. Figure 4 illustrates the distance of cities in China and Europe from the Eurasian steppe. As it makes clear,

[9]All data on deaths from warfare in the preindustrial period are highly speculative, but for our purposes what is important is the order of magnitude instead of the precise numbers reported. The high death tolls reported for conflicts such as the Mongol Invasions, the Ming-Qing transition, and the Taiping Rebellion are all borne out by recent research. Note that the majority of deaths did not occur on the battlefield but were the result of disease and pressure on food supplies (see Voigtländer and Voth, 2013b, 781 for a discussion).

[10]The Mongols and Manchu were nomadic or semi-nomadic. An Lushan was a general of nomadic origins. The Xin dynasty (9–23 AD) collapsed after a costly military campaign against the nomads coupled with massive flooding along the Yellow River triggered a civil war (China's Military History Editorial Committee, 2003).

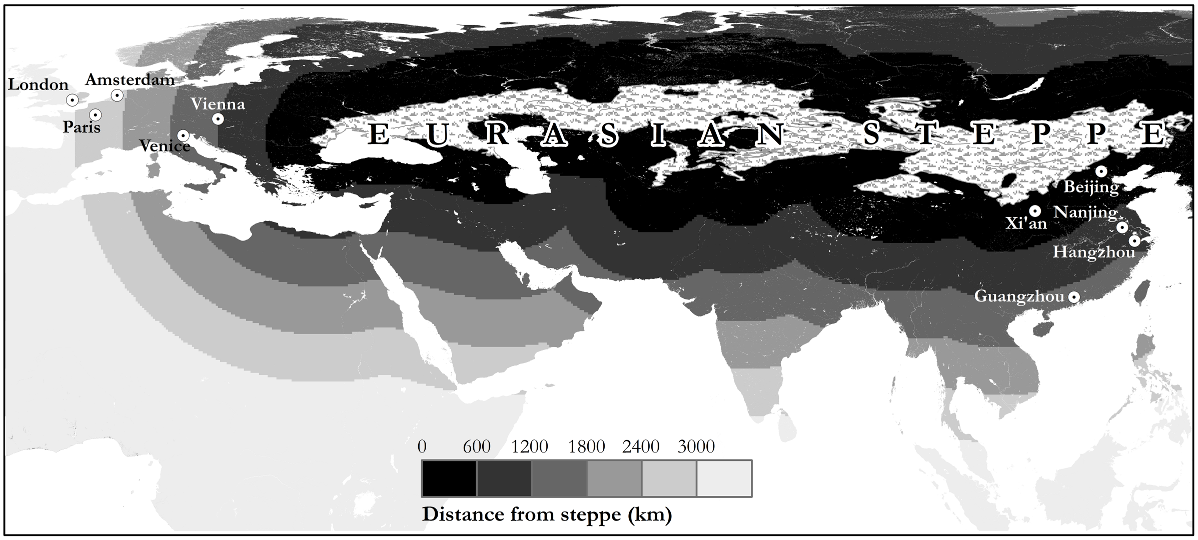


Figure 4: The Eurasian Steppe and Major Cities in China and Europe. Each shade represents 600 kilometers from the steppe.

Guangzhou, the southernmost major Chinese city, is almost as close to the steppe as Vienna, the easternmost major western European city.

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.

Three characteristics of the recurring conflicts between the steppe nomads and the agrarian Chinese differentiate them from typical interstate wars. First, as observed by Central Asian specialists (Lattimore, 1940; Barfield, 1989) and demonstrated empirically by Bai and Kung (2011) and Zhang et al. (2015), nomadic-agrarian conflicts were often climate-driven and therefore largely exogenous.

Second, warfare between the steppe and China was asymmetric in ways that favored the steppe. Although the sedentary Chinese were more populous by far, the expertise of the steppe nomads on horseback allowed them to develop mobile and powerful cavalry units that could easily outflank and outmaneuver infantry-based armies (Barfield, 1989; Gat, 2006). Importantly, horses were a location-specific asset. Horses bred in the steppe were hardy and had greater vigor as they were raised in an environment similar to that of wild horses (Zheng, 1984).

The third characteristic that sets nomadic-agrarian conflicts apart from typical interstate wars is the absence of towns or cities in the steppe for the sedentary people to capture in times of war. Since the main properties of the steppe pastoralists were their animal herds, which could be moved readily, nomads needed not defend their land against the enemy. When the odds were not in their favor, they could simply retreat into the safe haven of the steppe, where the undifferentiated 'highway of grass' allowed them to reach the Black Sea from Mongolia in a matter of weeks (Frachetti, 2008, 7). Hence, the nomads enjoyed an 'indefinite margin of retreat'—no matter how badly they were defeated in battle, they could never be conquered in war (Lattimore, 1940).

Until Russia's expansion into Central Asia in the seventeenth and eighteenth centuries denied the nomads their traditional escape route, the steppe threat was a recurring problem that the Chinese could not permanently resolve (Perdue, 2005).[11] Their best hope for security was the successful containment of the nomadic threat—hence the construction of the first Great Wall immediately after the first unification of China under the Qin dynasty in 221 BC.[12] The project was repeated time and again by successive dynasties at great cost to keep the 'barbarians' at bay.

Unidirectional versus Multidirectional Threats Many scholars have recognized the importance of the steppe nomads to state formation in ancient China (Lattimore, 1940; Huang, 1988; Turchin, 2009; Ma, 2012). In particular, Turchin (2009) observes that most historical empires were

[11]The Russian factor made possible the Qing dynasty's conquest of the Zunghar khanate, the last major nomadic empire in Asia, in 1755 (Perdue, 2005). From then onward, Qing China went into a prolonged period of military decline as its real military expenditures contracted steadily over time until the 1850s (Sng, 2014; Vries, 2015).

[12]During the Warring States period (475–221 BC) when China was divided into several competing kingdoms, the three that bordered the steppe—Qin, Zhao, and Yan—built long walls that were later linked up to form the first Great Wall of China after Qin successfully unified China.

situated on the fringes of the Eurasian steppe and identifies steppe raiding as a driver of state formation in China. We build on this literature by highlighting another important element in the nature of this threat: the external threats confronting China were unidirectional. There were no major threats from other fronts that would have increased the appeal of a more flexible politically decentralized system.

Before 1800, all major invasions of China came from the north. We argue that this was geographically determined as major geographical obstacles shielded China's eastern, western, and southern flanks (Figure 1). In the mid-1500s, coastal China did face extensive raiding by pirates (Kung and Ma, 2014). However, the problem was short-lived and in no way comparable to the perennial threat posed by the Eurasian steppe.

By contrast, Europe's external environment was different in two important ways. First, while Europe was also threatened by invasions from the steppe from Goths, Sarmatians, Vandals, Huns, Avars, Bulgars, Magyars, Pechenegs, Cumans, Mongols, and Turks,[13] the threat was less severe as Western Europe was relatively protected along its eastern flank by its forests and mountain ranges, and because it was relatively far from the steppe (Figure 4) and was buffered by the semi-pastoral lands of Hungary and Ukraine (Gat, 2006).

Second, Europe was more exposed to the rest of Eurasia and Africa. Consequently, prospective European empires typically faced enemies on multiple fronts: Vikings from the north; Arabs, Berbers, and Turks from the south and south-east; Magyars, Mongols, and others from the east (Appendix A.1 Table 7). These security challenges were particularly substantial in the first millennium. In Section 4.2, we discuss how this contributed to the collapse of the Roman empire and thwarted the attempts of Rome's successors, such as the Carolingian empire, to reunify Europe.

[13]See Appendix A.1 Table 6 for a list of all major nomadic invasions of both China and Europe.

3 Model

Building on the preceding discussion, we develop a model to explore the consequences of the severe one-sided threat that China faced in contrast with the weaker multi-sided threat faced by states in Europe. We consider a continent, which may represent China or Europe, as a Hoteling's linear city of unit length.[14] The continent faces external threats that can be one or two-sided. The continent contains one or more political regimes. Each regime (a) chooses its capital city, represented by a point along the linear line, (b) taxes its population, and (c) builds a military to resist the external threat and to compete with other regimes for territory and population. Our central concern is the fiscal viability of the regime(s) under political centralization and fragmentation, given the external threats that the continent confronts.[15] For illustrative purposes, we employ parametric forms for the functions in our analysis. The validity of our results is not tied to these parametric forms; in the Appendix (A.3), we provide the proofs of the results with more general functional forms.

3.1 Setup

We model a continent as a line $[0, 1]$ with a unit mass of individuals uniformly distributed along this line. An individual at $x \in [0, 1]$ is endowed with income $\underline{y} + y$ where y is taxable. For now, we fix the level of taxation at y and will endogenize it later.

The continent faces threats from outside. An external threat of magnitude Λ, if realized, causes gross damage Λ at the frontier(s). The damage can spread further into the continent: if a point is t distance away from the frontier, the **gross damage** is $\max\{\Lambda - \alpha t, 0\}$ where $\alpha > 0$ is a scaling constant. Moreover, a threat may emanate either from one frontier (at $x = 0$ only, without loss of generality) or from both frontiers (Figures 5 and 6). Whether it is one-sided or two-sided, and the value of Λ, depends on the continent's geographical environment, which is exogenously determined.

The continent is divided into $S \in \mathbb{N}_+$ connected, mutually exclusive intervals each ruled by a

[14] We refer to both Europe and China as 'continents' for convenience.

[15] Our theory builds on an extensive literature on modeling conflict. See Garfinkel and Skaperdas (2007) for a survey.

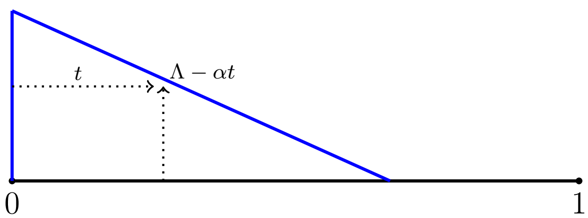


Figure 5: A severe one-sided threat.

Figure 6: A smaller, two-sided threat.

separate political authority or regime. We take S as given and do not model how regimes arise.[16] Instead, we focus on the fiscal viability of these regimes: we ask, for a given S, are the regimes fiscally viable given the continent's external environment?

For ease of exposition, we focus on $S \in \{1, 2\}$.[17] When $S = 1$ (political centralization), one regime or empire, e, rules the entire continent. When $S = 2$ (political fragmentation), two regimes, l and r, coexist. Regime l is on the left of regime r. For tractability, and because we are only interested in analyzing comparable regimes, we treat l and r as identical and focus on the symmetric equilibrium.[18]

A regime may invest in the military to (a) block the external threat, and (b) compete with other regimes for territory. The cost of military investment is convex; for regime $i \in \{e, l, r\}$ to provide a military investment of $M_i \geq 0$, it costs $c(M_i) = \theta M_i^2$.

A regime's military is strongest at its center of deployment, G, referred to here as its capital city. Like the external threat, military effectiveness deteriorates over distance. As Figure 7 illustrates, for a location that is t distance away from G_i, regime i's **military strength** on that location is given by $\max\{M_i - \beta t^2, 0\}$, where $\beta > 0$ measures the loss of military strength over distance. We expect the value of β to be relatively large in premodern times (compared to the present day) given the constraints imposed by premodern transportation and organizational technologies.

We assume that each regime can only maintain one capital city. Alternatively, we may assume that the regime can set up multiple auxiliary military bases (i.e., regional capitals), but the central

[16]Historically, the emergence of a regime is often associated with stochastic elements—the birth of a military genius; policy errors made by the incumbent ruler; climate change; and so on—that are difficult to capture in a model.

[17]A model extension that reproduces the results for $S > 2$ is available upon request.

[18]If one regime rules a much larger interval than the other one, the continent is effectively politically centralized.

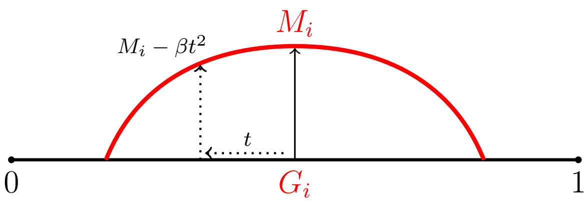


Figure 7: Regime i decides the location of its capital city (G_i) and its military investment (M_i).

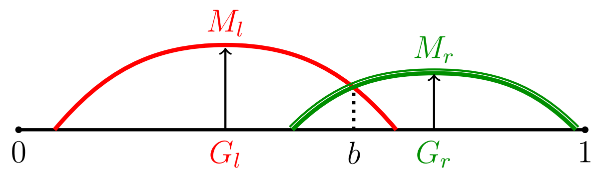


Figure 8: The border (b) between two regimes is determined by the locations of their capital cities and their relative military investments.

army must be dominant to prevent the regional armies from breaking away. In Appendix A.4, we show that these two assumptions are effectively equivalent; we also discuss how, historically, empires that maintained two or more comparable political-military centers (with none being dominant) either behaved like multiple states or would fragment into multiple states.

As Figure 8 illustrates, under political fragmentation, regime l controls $[0, b]$ and regime r controls $[b, 1]$. The border b is the location between the two capitals at which the military strength of the regimes are equal. Specifically, b is defined by the following equation:

$$M_l - \beta (b - G_l)^2 = M_r - \beta ((1 - G_r) - b)^2 . \tag{1}$$

Besides helping to define the border, the military also acts as a defense against the external threat by blocking it from spreading inland. Let $\kappa_i(x) = (\Lambda - \alpha\, x) - (M_i - \beta(G_i - x)^2)$. A location $x \in [0, 1]$ is **protected** by regime i from the external threat originating from 0 if there exists $0 \leq \hat{x} \leq x$ such that $\kappa_i(\hat{x}) \leq 0$. Otherwise, the external threat inflicts a **net damage** of $\kappa_i(x)$ at x. In a symmetric fashion, a location $x \in [0, 1]$ is **protected** by regime i from the external threat originating from 1 if there exists $x \leq \hat{x} \leq 1$ such that $\kappa_i(\hat{x}) \leq 0$. Let $\mathbb{D}_i$ denote the set of protected locations under regime i's control.

If less than δ fraction of the continent is protected, then a revolution occurs and all regimes in the continent receive negative payoffs. This assumption, common in models of political economy, captures the idea that regimes that disregard the welfare of the population risk being overthrown by revolutions, but revolutions involve overcoming collective-action problems and therefore require support from a threshold population of $1-\delta$ to be successful (see Alesina and Spolaore, 2003; Acemoglu and Robinson, 2006, for similar formulations).[19] If the revolution constraint is not violated, the net revenue of regime e under empire is $V_e = y - c(M_e)$ while the net revenues of regimes l and r under interstate competition are $V_l = by - c(M_l)$ and $V_r = (1-b)y - c(M_r)$, respectively.

3.2 Equilibrium

Under political centralization ($S = 1$), regime e first decides the location of its capital $G_e \in [0,1]$ and then decides its military investment $M_e \geq 0$ to maximize its net revenue V_e. Since this is a two-stage decision process, we employ backward induction to derive the optimal solution.

Proposition 1 (Empire)**.** *Under a two-sided threat of size* Λ*,*

1. *There exists* Λ_I *such that for all* $\Lambda \leq \Lambda_I$*,* $M_e^* = 0$*,* $G_e^* \in [0,1]$*, and* $|\mathbb{D}_e| \geq \delta$*;*

2. *There exists* $\Lambda_{II} > \Lambda_I$ *such that for all* $\Lambda_I < \Lambda \leq \Lambda_{II}$*,* $G_e^* = 1 - \frac{\Lambda}{\alpha} - \delta$*,* $M_e^* > 0$*, and* $|\mathbb{D}_e| = \delta$*;*

3. *For all* $\Lambda > \Lambda_{II}$*,* $G_e^* = \frac{1}{2}$*,* $M_e^* > 0$*, and* $|\mathbb{D}_e| = \delta$*;*

Under a one-sided threat of size Λ*,*

4. *There exists* $\bar{\Lambda}_I$ *such that for all* $\Lambda \leq \bar{\Lambda}_I$*,* $M_e^* = 0$*,* $G_e^* \in [0,1]$*, and* $|\mathbb{D}_e| \geq \delta$*;*

5. *For all* $\Lambda > \bar{\Lambda}_I$*,* $G_e^* = 1 - \delta$*,* $M_e^* > 0$*, and* $|\mathbb{D}_e| = \delta$*.*[20]

Case 1 of Proposition 1 implies that, if the external threat is very weak, the revolution constraint never binds. As such, the empire optimally makes zero military investment. Case 2 is the intermediate case in which the two-sided threat remains weak enough that the empire only needs to focus on

[19]It is also consistent with the Confucian belief that the legitimacy of a government is contingent upon its ability to protect the people from harm and tax reasonably so that the people can maintain a constant means of livelihood. A government that loses this ability loses its 'mandate from heaven,' and the people would therefore be entitled to depose it (Mencius, 2004).

[20]The closed-form expressions of Λ_I, Λ_{II}, and $\bar{\Lambda}_I$ are $\Lambda_I = \frac{1}{2}\alpha(1-\delta)$, $\Lambda_{II} = min\{\frac{1}{4}\beta\delta^2 + \Lambda_I, 2\Lambda_I\}$, and $\bar{\Lambda}_I = 2\Lambda_I$.

building up its military on one frontier while ignoring the other frontier to protect δ fraction of its population. In Case 3, the empire locates its capital at the center of the continent and builds its military to defend both frontiers against a threat that is now **nontrivial** in that it cannot be fully or partially ignored. Cases 4 and 5 depict the empire's optimal responses under a one-sided threat and are analogous to Cases 1 and 3 respectively.

Next, consider a two-stage game with interstate competition ($S = 2$). Regimes l and r simultaneously choose their capital cities $G_l \in [0,1]$ and $G_r \in [0,1]$. After observing the capital city locations, the regimes simultaneously make military investments $M_l \geq 0$ and $M_r \geq 0$. Again, we employ subgame-perfect equilibrium as the solution concept.

Proposition 2 (Interstate Competition)**.** *Consider the symmetric equilibrium where $M_l^* = M_r^*$ and $G_l^* = 1 - G_r^*$. Under a two-sided [one-sided] threat of size Λ,*

6. *There exists Λ_{III} [$\bar{\Lambda}_{III}$] such that, if $\Lambda \leq \Lambda_{III}$ [$\Lambda \leq \bar{\Lambda}_{III}$], the revolution constraint does not bind, $|\mathbb{D}_l| + |\mathbb{D}_r| \geq \delta$, and $M_l^* = M_r^* > 0$. The equilibrium military investments and location of capitals are the same as in the case in which $\Lambda = 0$;*

7. *If $\Lambda > \Lambda_{III}$ [$\Lambda > \bar{\Lambda}_{III}$], the revolution constraint binds, $|\mathbb{D}_l| + |\mathbb{D}_r| = \delta$, and $M_l^* = M_r^* > 0$.*[21]

In contrast to the case of empire, in which the optimal military investment is zero when the external threat is trivial, regimes in a competitive state system have to invest in the military to compete for territory with or without the external threat. Proposition 2 states that, unless the external threat is **severe** (Case 7), regimes l and r do not make additional military investments to protect their populations, as the military capacity built to compete between themselves already meets the need of defending against the external threat.

[21] The closed-form expressions of Λ_{III} and $\bar{\Lambda}_{III}$ are $\Lambda_{III} = \frac{1}{2}\alpha(1-\delta) - \frac{\beta}{4}\left(\delta - \left(\frac{y}{4\theta\beta^2}\right)^{\frac{1}{3}}\right)^2 + (\frac{y^2}{16\theta^2\beta})^{\frac{1}{3}}$ and $\bar{\Lambda}_{III} = \alpha(1-\delta) - \frac{\beta}{4}\left(2\delta - 1 - \left(\frac{y}{4\theta\beta^2}\right)^{\frac{1}{3}}\right)^2 + (\frac{y^2}{16\theta^2\beta})^{\frac{1}{3}}$.

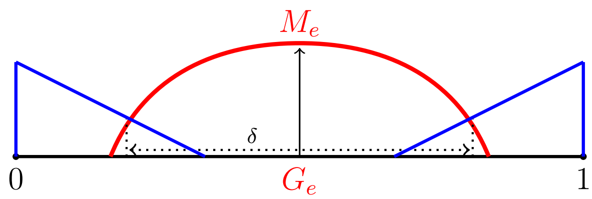


Figure 9: Optimal military investment under political centralization.

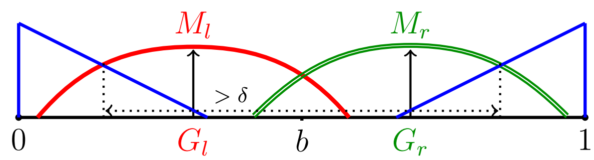


Figure 10: Optimal military investment under political fragmentation.

3.3 Implications for Political Centralization or Fragmentation

Together, Propositions 1 and 2 indicate that political centralization and fragmentation have different strengths and weaknesses. First, in the absence of external threats, political fragmentation is wasteful from a static perspective and there are Pareto gains to be reaped if competitive regimes coordinate to reduce their military spending. Hence:

Implication 1 (**Wastefulness of interstate competition**). *If $\Lambda = 0$, military investment is zero under an empire but strictly positive under interstate competition.*

When a nontrivial external threat is present, an empire will only protect up to δ fraction of the population to satisfy the revolution constraint (Figure 9). By contrast, in a competitive state system, the competition-induced over-investment in the military may result in a larger-than-δ fraction of the continent being protected (Figure 10). Hence:

Implication 2 (**Robustness of interstate competition**). *If $\Lambda > 0$, interstate competition protects a weakly bigger interval of the continent than an empire does.*

Proposition 1 also suggests that the choice of an empire's capital city is influenced by the nature of the external threats that it confronts. In particular, if the empire faces a nontrivial one-sided threat, it will locate its capital city at $G_e^* = 1 - \delta$ to contain the threat. The higher is δ, the closer the capital city is to the frontier where the threat originates. Hence:

Implication 3 (**Locational choice of capital city**). *Under a one-sided external threat of size $\Lambda > \bar{\Lambda}_I$, it is not optimal for an empire to locate its capital city at the center, i.e., $G_e^* \neq 0.5$.*

Theoretical and empirical studies generally argue that capital cities should be centrally located to maximize tax revenue or improve governance (Alesina and Spolaore, 2003; Olsson and Hansson, 2011; Campante et al., 2016). However, history is replete with examples of peripheral centers being chosen as capitals. Implication 3 offers an explanation as to why, in some of these cases, it may indeed be optimal to separate the political center of the empire from its economic or population center. In Section 5.1, we further provide a historical discussion on Roman and Chinese capital cities in light of this prediction.

3.4 Resilience of Political Centralization or Fragmentation under Different Threat Scenarios

Our central concern is the resilience of political centralization and political fragmentation under (a) one-sided and (b) two-sided external threats. We measure resilience by the fiscal viability of the regimes in question. A regime is (fiscally) **viable** if its equilibrium net revenue is non-negative. When $S = 1$ and the empire is nonviable, political centralization cannot last. Likewise, when $S = 2$ and one or both regimes are nonviable, political fragmentation is unsustainable.

Implication 4 (**Resilience under one-sided threat**). *Political centralization is more resilient than political fragmentation to a one-sided threat.*

Implication 4 highlights the resource advantage of political centralization. Compare two continents, one politically centralized and the other politically fragmented, that are otherwise identical. Each faces a one-sided external threat. As the threat level increases, the fiscal viability of regimes in both continents decreases as they respond by making (weakly) larger investments in the military. In Figure 11, regime l will become nonviable if the external threat is sufficiently severe. However, at this threshold threat level, the empire in the parallel world can remain viable simply by replicating regime l's equilibrium decisions on G and M because it commands the resources of the entire continent. In fact, it is easy to see that $V_e^* > V_l^* + V_r^*$ under a one-sided threat. Hence, while a one-sided threat weakens both empires and competitive states, the empire is more likely to survive unscathed.

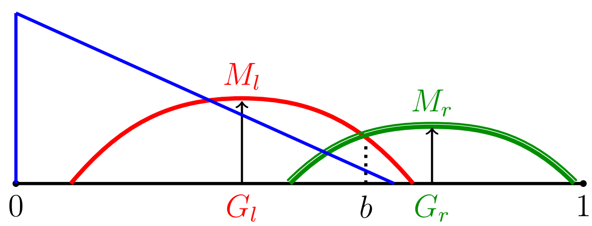


Figure 11: A severe one-sided threat jeopardizes the viability of regime l, which has to compete with regime r for the continent's resources.

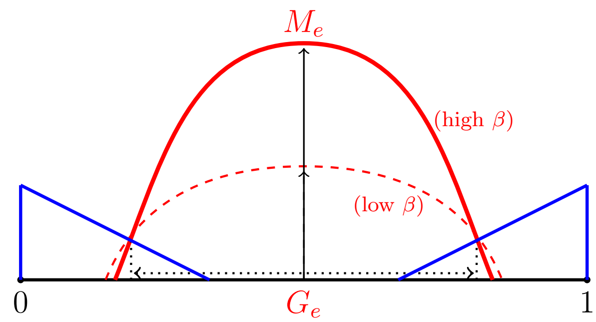


Figure 12: Even when a two-sided threat is nonsevere, the empire has to invest heavily to secure two borders from one center if β is large.

Implication 5 (**Resilience under two-sided threat**). *Political fragmentation is more resilient than political centralization to a two-sided threat if the efficiency of projecting military power is low.*

Implication 5 highlights a potential drawback of political centralization: when confronted with a nontrivial two-sided threat, an empire has to simultaneously protect two frontiers that are far apart, while competitive states are collectively capable of managing both frontiers from close proximity at once. As Figure 12 shows, the empire may need to spend heavily on the military to deal with a two-sided threat, even if the threat is nonsevere and competitive states would effectively ignore it.

Generally, under a nontrivial two-sided threat, $V_e^* < V_l^* + V_r^*$ if β, which measures the inefficiency of military projection over long distances, is sufficiently large (Appendix A.3.4). Hence, political fragmentation is more resilient than political centralization as long as military effectiveness deteriorates relatively quickly over distance, which is likely to be the case given premodern technological constraints.[22]

In Section 4.2, we discuss the decline of the Roman and Carolingian empires in light of Implication 5.

[22]Historically, a key constraint on the projection of military power was logistics (van Creveld, 2004). According to the Chinese polymath Shen Kuo (1031–1095), in Song China, a soldier would need one porter for supplies to march 18 days. Extending the campaign to 31 days would require a tripling of porters as the existing porter would need help for his supplies too (Shen, 2011).

3.5 Other Results: Taxation

Importantly, the two scenarios depicted in Figures 11 and 12—a severe one-sided threat that squeezes out small regimes and a nonsevere two-sided threat that renders empires inefficient—are analogous to China's and Europe's external environments, respectively. Would the levels of taxation differ in these scenarios?

So far, we assume that the tax on an individual is fixed at y. To endogenize taxation, suppose regime i has the option of reducing the tax burden of its people by $R_i \geq 0$ so that the effective rate of taxation becomes $y - R_i$. By keeping the population content, lowering taxes eases the revolution constraint. Specifically, consider a location x taxed by regime i that is not protected from the external threat. The individual at x does not rebel as long as

$$\underbrace{R_i}_{\text{tax reimbursement}} + \underbrace{M_i - \beta(G_i - x)^2}_{\text{military protection}} - \underbrace{(\Lambda - \alpha x)}_{\text{damage from threat}} \geq 0. \tag{2}$$

To satisfy this revolution constraint, an empire may find it cheaper to offer a policy mix of low taxes (i.e., set $R_e > 0$) and some military investment than to rely on military investment alone. By contrast, if interstate competition is intense, the revolution constraint does not bind as a consequence of heavy military investments and there is therefore no need to lower taxes (i.e., $R_l = R_r = 0$).[23] Hence:

Implication 6 (Taxation). *Taxation is weakly lower in a politically centralized continent confronting a one-sided threat than in a politically fragmented continent confronting a two-sided threat.*

In keeping with political economy models of autocratic states (North, 1981; Olson, 1993; Mayshar et al., 2015), we assume that rulers aim to maximize tax revenue. Despite this, however, we show that it could be in the interest of an empire to impose comparatively low taxes in order to relax the revolution constraint while no such incentive exists for competitive states.

[23] See Appendix A.3.5 for a more formal treatment.

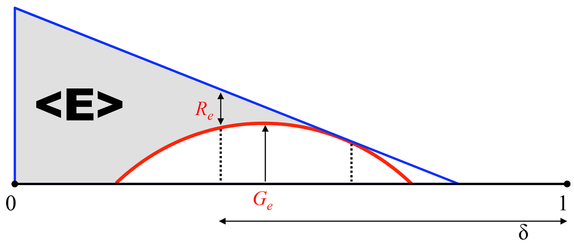


Figure 13: Under political centralization, there is a positive level of tax reduction, i.e. $R_e^* > 0$. Under political fragmentation, tax reduction is zero.

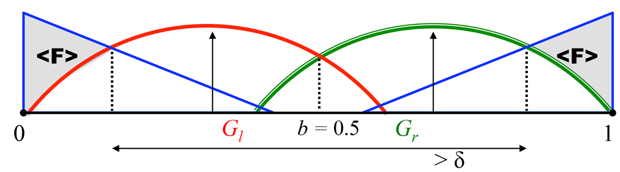


Figure 14: Under political fragmentation, the fraction of the population protected from invasion is at least δ. Under political centralization, it is at most δ.

3.6 Population Dynamics and Long-run Growth

Until now, we have assumed that external threats are always present. Suppose instead that the external threat is realized with some positive probability. Suppose also that each individual inelastically supplies labor to produce $y + \underline{y}$, where $\underline{y}$ is not taxable. For individual x under regime i, her disposable income is $\bar{y} = \underline{y} + R_i - \kappa_i(x)$, where R_i is the tax reimbursement and $\kappa_i(x)$ is the net damage caused by the stochastic shock. Each individual chooses between private consumption c and producing n offspring to maximize her utility $u(c,n) = c^{1-\gamma}n^{\gamma}$ subject to the budget constraint $\rho n + c \leq \bar{y}$, where ρ represents the cost of raising a child. We assume that c and n are complements and u is increasing and concave in both arguments. Standard optimization implies that the optimal number of children is $n = \frac{\gamma}{\rho} \cdot \bar{y}$. The continent's population will therefore grow to

$$N = \int_0^1 n dx = \frac{\gamma}{\rho} \cdot \int_0^1 \bar{y} \, dx \, . \tag{3}$$

Let N_E and N_F denote future population levels in continents E and F, respectively. The two continents are identical except that E is ruled by an empire ($S = 1$) and faces a one-sided threat of size Λ_E, while F is politically fragmented ($S = 2$) and faces a two-sided threat of size Λ_F.

When the external threat is not realized, the populations in the two continents are $N_E = \frac{\gamma}{\rho} \cdot (\underline{y} + R_e)$ and $N_F = \frac{\gamma}{\rho} \cdot (\underline{y})$ respectively. Since $N_E > N_F$, population grows faster under the empire.

However, if the external threat is realized,

$$N_E = \frac{\gamma}{\rho} \cdot \left\{ (\underline{y} + R_e) - \underbrace{\int_{x \notin \mathbb{D}_e} (\Lambda_E - \alpha x) - (M_e - \beta (G_e - x)^2) \; dx}_{\text{Area}\langle E \rangle} \right\}; \tag{4}$$

$$N_F = \frac{\gamma}{\rho} \cdot \left\{ \underline{y} - \underbrace{2 \cdot \int_{x<b, x \notin \mathbb{D}_l} (\Lambda_F - \alpha x) - (M_l - \beta \; (G_l - x)^2) \; dx}_{\text{Area}\langle F \rangle} \right\}, \tag{5}$$

where Area $\langle E \rangle$ and Area $\langle F \rangle$ are illustrated in Figures 13 and 14.

If $\Lambda_E > \Lambda_F$, Area $\langle E \rangle$ is likely to be larger than Area $\langle F \rangle$ not only because continent E confronts a more severe external threat, but also because the empire protects less than δ fraction of the continent (or exactly δ if tax reimbursement is zero), while the protected fraction of continent F is always weakly larger than δ (Implication 2). If Area $\langle F \rangle <$ Area $\langle E \rangle - R_E$, it follows that $N_E < N_F$.

Hence, when the external threat is not realized, population grows faster in continent E, but when the threat is realized, a population contraction is also more likely there:

Implication 7 (**Population Change**). *Population change displays a higher variance in a politically centralized continent confronting a severe one-sided threat than in a politically fragmented continent confronting a nonsevere two-sided threat.*

In interpreting our model, we have focused on external invasions. More generally, however, negative shocks could also stem from unforeseen political collapses and peasant rebellions in addition to invasions from outside. The central point we emphasize is that interstate competition results in a greater proportion of territory being protected than is the case under political centralization.[24]

[24]The outline of a model extension that incorporates internal rebellions is available upon request.

4 External Threats and Political Unification or Fragmentation: Empirical Evidence

The model predicts that an environment with nontrivial external threats originating from multiple fronts favors interstate competition, while a unidirectional threat promotes political unification. To test these predictions, we first investigate the empirical relationship between the frequency of nomadic attacks and political unification in China. Subsequently, we examine historical evidence from Europe.

4.1 Empirical Evidence from China

To test our hypothesis that a severe and unidirectional threat from the steppe provided a recurring impetus for unification in China, we exploit time series variation in political unification and fragmentation in Chinese history. We show that periods of more conflict with steppe nomads were positively associated with periods of political unification, and vice versa.

Data Sources and Definition of Variables. In a recent paper, Bai and Kung (2011) show that nomadic incursions into China were correlated with exogenous variations in rainfall as subsistence crises triggered by droughts and other climatic disasters often drove inhabitants of the ecologically fragile steppe to invade their settled neighbors. Their dataset and empirical strategy provide us with a convenient tool to test whether there was a relationship between the frequency of nomadic attacks and political unification in China. It also helps to ensure that our empirical evidence is robust and is not selectively adopted to suit our purpose.

Bai and Kung's data span 2,060 years (from 220 BC to AD 1839) and are drawn from four sources: *A Chronology of Warfare in Dynastic China* (China's Military History Editorial Committee, 2003), *A Compendium of Historical Materials on Natural Disasters in Chinese Agriculture* (Zhang et al., 1994), *A Concise Narrative of Irrigation History of the Yellow River* (Editorial Committee of Irrigation History of the Yellow River, 1982), and the *Handbook of the Annals of China's Dynasties* (Gu, 1995). Of these sources, the first three have been widely used in related research and are considered reliable

Table 1: List of Variables and Summary Statistics

Variable		Description	mean	s.d.
Unification	y_t	=1 if only 1 regime ruled China in decade t	0.59	0.49
#Regimes	y_t^{alt}	Average number of regimes in China proper in decade t (*log*)	0.39	0.54
Nomad attacks	x_t	Number of attacks initiated by the nomads in decade t	2.53	3.50
Lower precipitation	z_{1t}	Share of years with records of drought disasters on the Central Plain in decade t	0.50	0.30
Higher precipitation	z_{2t}	Share of years with records of Yellow River levee breaches in decade t	0.18	0.21
Snow disasters	w_{1t}	Share of years with records of heavy snow on the Central Plain in decade t	0.12	0.14
Low temperature disasters	w_{2t}	Share of years with records of low-temperature calamities (e.g., frost) on the Central Plain in decade t	0.16	0.19
Temperature	w_{3t}	Average temperature in decade t	9.46	0.89
Nomadic conquest 1	w_{4t}	=1 if the Central Plain was governed by the nomads (317–589)	0.13	0.33
Nomadic conquest 2	w_{5t}	=1 if the Central Plain was governed by the nomads (1126–1368)	0.12	0.32
Nomadic conquest 3	w_{6t}	=1 if the Central Plain was governed by the nomads (1644–1839)	0.10	0.29
Time trend	w_{7t}	Decade: -22–183 (219 BC–1839)	80.5	59.6
Steppe empire	w_{8t}	Presence of nomadic empire in the steppe in decade t	0.64	0.48

Sources: Bai and Kung (2011), Wei (2011), and Barfield (1989).

sources while the fourth contains general historical information that can be easily verified.

As Table 1 shows, the decadal variables Bai and Kung constructed include: (i) the frequency of nomadic attacks on China's Central Plain (x_t); (ii) two precipitation variables that measure the extent of severe droughts and floods in the Central Plain (z_{1t}, z_{2t}); (iii) other climatic control variables (snow and other low temperature disasters, temperature; w_{1t}–w_{3t}); (iv) dummy variables that denote the three periods in Chinese history when the Central Plain was governed by the nomads (w_{4t}–w_{6t}); and (v) a time trend (w_{7t}). Of these variables, we are primarily interested in the frequency of nomadic attacks, which constitutes our main explanatory variable. But we will use the other variables as controls to check the robustness of our benchmark results later.

We add three new variables to the dataset: our dependent variable, *Unification* (y_t), which takes the value of 1 if China was politically unified in a given decade and 0 otherwise; the log number of Chinese regimes in decade t (y_t^{alt}); and a dummy variable w_{8t} which takes the value of 1 if there existed a steppe empire in a given decade, and 0 otherwise. The first two variables are drawn from Wei (2011) and the last variable is based on Barfield (1989). The correlation matrices of the variables are provided in Appendix Table 8.

ADL Estimation. We first employ a simple autoregressive distributed lag (ADL) model:

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{q} \mu_i x_{t-i} + \epsilon_t \ , \tag{6}$$

where y_t is the dummy variable *Unification* and x_t is the number of nomadic incursions in decade t.

The ADL model is appropriate for our purpose because of its flexibility. Furthermore, it generates unbiased long-run estimates and valid t-statistics even in the presence of endogeneity (Harris and Sollis, 2003). To validate our use of the ADL methodology, we use the Augmented Dickey-Fuller test to ensure that all variables are stationary. To determine the appropriate number of lags, we follow the general-to-specific approach proposed by Ng and Perron (1995) to seek the values of p and q in Equation 6 that minimize the Akaike Information Criterion (AIC), which occurs at $p = 3$ and $q = 1$.[25]

According to Implication 4, an increase in the severity of nomadic threat favors political unification. This effect may not be immediate because, in the short run, a spike in nomadic attacks would decrease the resilience of both political unification and fragmentation. If an established state collapses, a host of competing successors would emerge and scramble to fill up the political vacuum so that more political fragmentation is observed initially. Nonetheless, Implication 4 predicts that increased nomadic attacks should have a long-run positive effect on political unification. In other words, we expect $\mu_0 + \mu_1 + \mu_2 + ... + \mu_q > 0$.

Indeed, in the ADL estimation reported in Table 2 Column (a), the nomadic invasion variable and its lagged value are statistically significant and carry opposite signs: an additional nomadic attack in decade t is associated with a 1.2 percentage point decrease in the probability of political unification in China in the same decade, but an attack in the previous decade (at $t - 1$) is associated with a larger 1.96 percentage point increase in the probability of political unification in decade t.[26] Their joint F statistic is 5.32. Hence one can reject the null hypothesis that the two coefficients are jointly zero at

[25] When implementing the general-to-specific approach, we choose $p = q = 10$ as the cut-off and check every combination of $p \leq 10$ and $q \leq 10$.

[26] The Durbin's h-test indicates that the errors are serially independent. In addition, the roots of the characteristic equation are all smaller than 1 and therefore the estimation model is 'dynamically stable.'

Table 2: ADL Model

Dependent variable:	(a) Unification	(b) Unification	(c) Unification	(d) Unification	(e) Unification
Unification: Lag1	0.906***	0.915***	0.916***	0.875***	0.847***
	(0.0651)	(0.0649)	(0.0660)	(0.0668)	(0.0672)
Unification: Lag2	-0.283***	-0.272***	-0.272***	-0.277***	-0.289***
	(0.0843)	(0.0838)	(0.0846)	(0.0837)	(0.0830)
Unification: Lag3	0.256***	0.239***	0.239***	0.202***	0.199***
	(0.0630)	(0.0630)	(0.0641)	(0.0648)	(0.0641)
Nomad attacks	-0.0120**	-0.0104*	-0.0102*	-0.0108*	-0.0136**
	(0.00605)	(0.00605)	(0.00615)	(0.00628)	(0.00633)
Nomad attacks: Lag1	0.0196***	0.0170***	0.0171***	0.0182***	0.0174***
	(0.00604)	(0.00606)	(0.00616)	(0.00621)	(0.00614)
Steppe empire					0.104**
					(0.0448)
Additional controls:					
Droughts & floods	No	Yes	Yes	Yes	Yes
Climatic controls	No	No	Yes	Yes	Yes
Nomad conquest	No	No	No	Yes	Yes
Observations	203	203	203	203	203
R-squared	0.743	0.753	0.753	0.765	0.772
AIC	0.122	0.120	0.150	0.141	0.122

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Constant terms not reported.

1 percent significance level. In line with Implication 4, the relationship between nomadic invasions and political unification is positive in the long run: each additional nomadic attack is associated with an increase in the probability of political unification of 6.3 ($= \frac{-0.012+0.0196}{1-0.906+0.283-0.256}$) percentage points. Given that China experienced an average of 2.5 nomadic attacks per decade, the relationship between nomadic invasions and its political unity is a substantial one.

Next, to check that the analysis is not distorted by omitted variable bias, we deviate from the classic ADL model and introduce control variables incrementally. Since Bai and Kung (2011) find that nomadic incursions into China were positively correlated with less rainfall and negatively correlated with more rainfall, it is possible that in our regression analysis, nomadic attacks are merely acting as a proxy for bad climatic conditions, which affects unification in China through other channels (e.g., by causing harvest failures and therefore weakening state finances). To check, we add the frequency of droughts and floods (z_{1t} and z_{2t}) as control variables in Table 2 Column (b). In Column (c), we add three more climatic variables that may be correlated with our main explanatory variable of nomadic invasions: decadal share of years with records of snow disasters (w_{1t}), decadal share of years with records of frost and other low temperature disasters (w_{2t}), and average temperature (w_{3t}). The inclusion of these controls leads to multicollinearity and increases the standard errors of the estimates.

However, we obtain coefficient estimates similar to the baseline results in Column (a).

Historically, there were periods when some nomadic groups settled in China. During these times, the odds of China suffering further attacks from the steppe might be lower. To check for omitted variable bias of this nature, we control for the three historical periods during which the central plains of China were ruled by nomadic regimes (w_{4t}, w_{5t}, w_{6t}) and add a decadal time trend (w_{7t}) in Column (d). In Column (e), we further include a dummy variable to control for periods when a nomadic empire dominated the steppe (w_{8t}). The results continue to remain stable. In fact, if we interpret the presence of steppe empires as an increased nomadic threat, the positive coefficient estimate of the steppe empire dummy in Column (e) provides further evidence to support our argument.

VAR Estimation. Nonetheless, the estimation above establishes correlation, not causation. Reverse causality is clearly a concern. For example, suppose nomadic attacks did not matter at all and it was in fact some other structural factor(s) that alone encouraged the rise of autocracy in China. Since autocratic states tend to "overgrow" (Alesina and Spolaore, 1997), nomadic invasions on the Chinese empire could merely be a product of the tendency for the autocratic Chinese state to expand till it bordered the steppe and encountered conflicts with the nomads. However, if this was the case, the coefficients of the nomadic invasion variables in our ADL analysis should be economically and statistically insignificant (since the Chinese border region would experience periodic attacks by the neighboring nomads regardless of whether China was unified or fragmented). We can therefore safely rule out this alternative explanation.

Reverse causation could also arise if the steppe nomads were more likely to attack China when it was politically divided and weakened. This source of endogeneity does not pose a problem for our argument. In fact, it suggests that the short-run negative estimated effect of nomadic attacks on political unification of 1.2 percentage points may be an overestimate. If so, the long-run effect of nomadic attacks on political unification would be larger than what we estimated above. To investigate further, we implement the following vector autoregression (VAR), which models the simultaneity of our dependent and main explanatory variables explicitly:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \phi_0 \\ \mu_0 \end{bmatrix} + \begin{pmatrix} \phi_1^1 & \mu_1^1 \\ \phi_1^2 & \mu_1^2 \end{pmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{pmatrix} \phi_2^1 & \mu_2^1 \\ \phi_2^2 & \mu_2^2 \end{pmatrix} \begin{bmatrix} y_{t-2} \\ x_{t-2} \end{bmatrix} + \begin{pmatrix} \phi_3^1 & \mu_3^1 \\ \phi_3^2 & \mu_3^2 \end{pmatrix} \begin{bmatrix} y_{t-3} \\ x_{t-3} \end{bmatrix} + \begin{bmatrix} \epsilon_{t-1} \\ \epsilon_{t-2} \end{bmatrix} . \quad (7)$$

As with the previous estimations, we select the lagged values that minimize the AIC. We also check for autocorrelation and that the eigenvalues lie inside the unit circle (hence the VAR model is 'dynamically stable'). As Table 3 illustrates, the estimates from the VAR model share the same order of magnitude as the results from the ADL estimations. The coefficient estimate of Lag-1 nomadic attack is 0.0176 in Column (a) of Table 3, compared with 0.0196 in Table 2 Column (a). In the long run, each additional nomadic attack is associated with an increase in the probability of political unification of 8.3 ($= \frac{0.0176-0.00701+0.00602}{1-0.893+0.317-0.225}$) percentage points. The Wald test statistic of the coefficients on the lags of nomadic attacks is 11.36, which allows us to reject the null hypothesis that nomadic attacks did not Granger-cause political unification at 1 percent significance level.[27]

Table 3: VAR

Dependent variable:	(a) Unification	(b) Nomad attacks
Unification: Lag1	0.893***	-2.075***
	(0.0665)	(0.733)
Unification: Lag2	-0.317***	1.631*
	(0.0848)	(0.935)
Unification: Lag3	0.225***	0.377
	(0.0656)	(0.723)
Nomad attacks: Lag1	0.0176***	0.321***
	(0.00626)	(0.0690)
Nomad attacks: Lag2	-0.00701	0.257***
	(0.00657)	(0.0724)
Nomad attacks: Lag3	0.00602	-0.0108
	(0.00638)	(0.0703)
Additional controls	Yes	Yes
Observations	203	203

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: IV Analysis

Dependent variable:	Unification
Unification: Lag1	0.743***
	(0.141)
Unification: Lag2	-0.0797
	(0.188)
Unification: Lag3	0.108
	(0.121)
Nomad attacks	-0.0874
	(0.0561)
Nomad attacks: Lag1	0.134**
	(0.0639)
Additional controls	Yes
Observations	203

First stage not reported.

[27]In Appendix A.6, we further test the robustness of the results by using the *log* number of regimes in China proper as an alternative dependent variable. Consistent with the above results, an increase in nomadic attacks is associated with a decrease in the number of regimes in China proper.

Instrument Variable Analysis. As a further check on reverse causality, instead of using the frequency of droughts and floods (z_{1t} and z_{2t}) as control variables, we use them as instruments to tease out the direction of causation. As Table 4 illustrates, the estimated effect of nomadic invasions on political unification in China remains statistically significant when the instruments are applied. In fact, its magnitude is much larger than before. In the long run, each additional nomadic attack is associated with an increase in the probability of political unification of 20.4 ($= \frac{-0.0874+0.134}{1-0.743+0.0797-0.108}$) percentage points. However, this estimate should be interpreted with caution. There are reasons to think that the exclusion restriction may not hold in our IV analysis as climate could have affected unification in China through other channels.

4.2 Historical Evidence from Europe

We are unable to replicate the above empirical exercise for Europe because data on the number of regimes in Europe only exists on a per century basis. However, European historical patterns do conform to predictions of our theoretical model.

Europe has historically been politically fragmented; the closest Europe came to be ruled by a unified political system was under the Roman Empire. The rise of Rome parallels the rise of the first empire in China (Scheidel, 2009). In terms of the model, one advantage Rome had over its rivals in the Hellenistic world was a relatively less convex cost function of military investment—Rome's ability to project power and increase its resources of manpower was unequaled among European states in antiquity. Thus, Rome was able to impose centralized rule upon much of Europe. Our model suggests that two factors can account for the decline of Rome: (1) over time, Rome's military advantage declined relative to the military capacities of its rivals such as the Persian empire or the Germanic confederacies; and (2) these rising threats came from multiple directions along Rome's long border (Heather, 2006). Like episodes of dynastic and imperial collapse in China, the fall of the western Empire was associated with political disintegration and economic collapse across Europe (Ward-Perkins, 2005).

After the Fall of the Western Empire, the Eastern Roman Empire continued for another millennium.

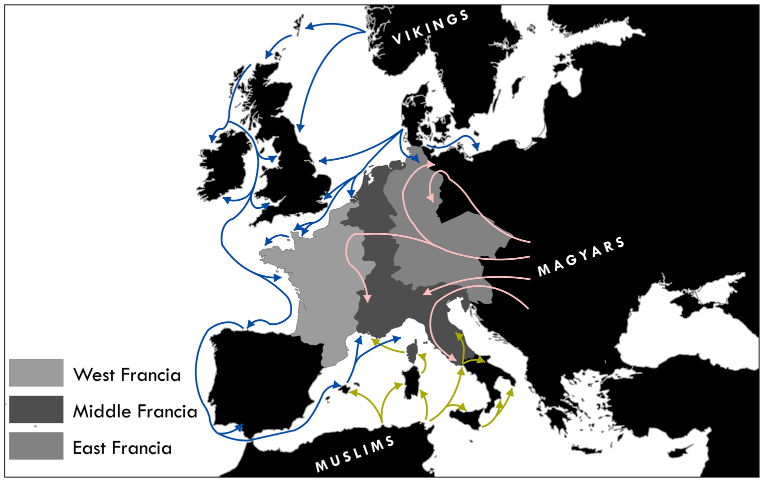


Figure 15: Viking, Magyar, and Muslim Invasions of Western Europe in the Ninth and Tenth Centuries; The Carolingian Empire after the partition of AD 843.

In the mid-sixth century Justinian I (r. 527–565) attempted to recreate the old empire by conquering north Africa and Italy. But this attempt was short-lived. In the early seventh century, the empire nearly collapsed under the two-sided threat of first the Avers and Persians and then the Arabs. The remnant of the Byzantine empire that survived was a substantially smaller state (Appendix A.5).

The creation of a Frankish empire by the Carolingians represents another attempt to build a Europe-wide empire. During the reign of Charlemagne (r. 768–814), the Carolingians came to control an empire that spanned France, parts of Spain, and much of Italy and central Europe (Collins, 1998; McKitterick, 2008). Consistent with our model, however, the Carolingian empire was not long-lasting. It went into decline as the successors of Charlemagne struggled to deal with the external threats posed by the Magyars, the Vikings, and the Muslims from different fronts (Morrissey, 1997). In East Francia, a different dynasty, the Ottonians, came to power as a response to the repeated Magyar invasions, and established the Holy Roman Empire. Increasingly, emperors based in Germany found it difficult to control their Italian provinces, and by the thirteenth century, the Holy Roman Empire was no more than a loose federation of German principalities.

Incidentally, the threats posed to the Europeans by the Vikings and the Muslims receded after the eleventh century. One could argue that from then on, Western Europe no longer experienced meaningful multi-sided external threats.[28] If this interpretation is correct, our model predicts that the status quo of political fragmentation would persist, and it did. The Mongol invasion of Europe in the thirteenth century was too brief to provide a sustained impetus toward European unification. However, the less dramatic but more sustained rise of the Ottoman empire after the fifteenth century serves as yet another test of our model, and it provides further supporting evidence that our mechanisms are relevant. Iyigun (2008) shows that the external threat of invasion from the Ottomans between 1410 and 1700 reduced the frequency of interstate warfare in Eastern Europe. Indeed, a comparison of the political maps of Central and Eastern Europe of the fourteenth century and seventeenth century indicates that 'a significant degree of political consolidation accompanied the Ottoman expansion in continental Europe' (Iyigun, 2008, 1470).

In Section 6, we further discuss how Europe's political fragmentation before 1100 provided the necessary precondition for political institutions that reinforced the centrifugal tendencies—such as the independent city states and a powerful Catholic Church—to emerge and take root.

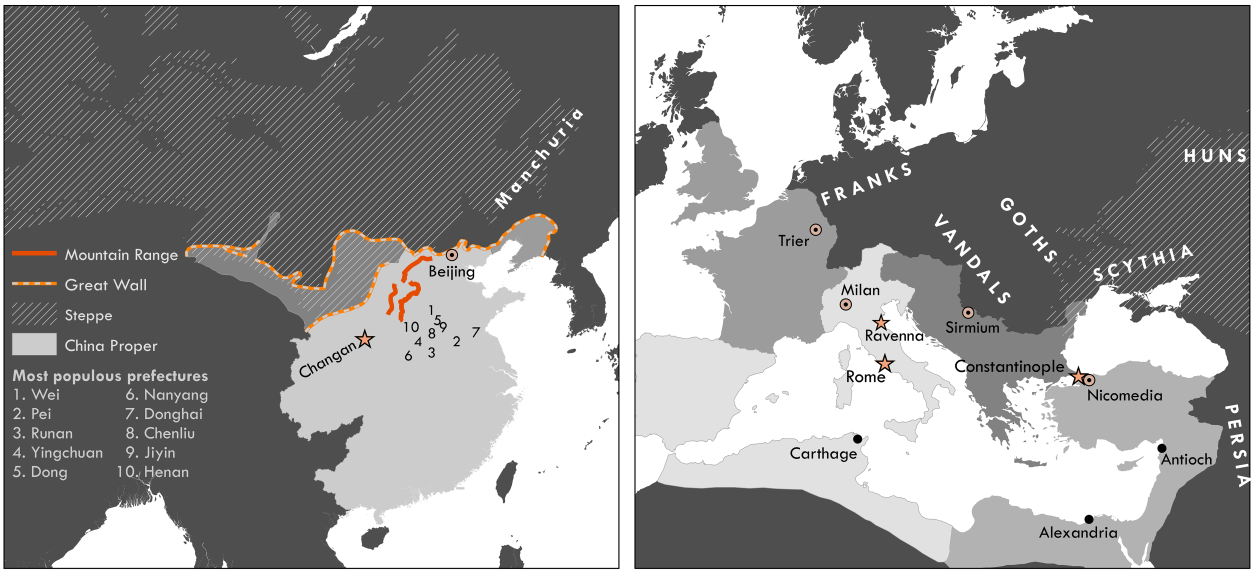
5 Applying the Model

While our model is evidently too simple to explain many aspects of historical China and Europe, it offers a useful framework to organize some salient comparative historical evidence. Below, we use our theory to offer new insights into the location choice of capital city, differential levels of taxation, and differential patterns of population change at the two ends of Eurasia.

[28]Nevertheless, Europe continued to face a potential invasion threat from the south and the east, and after 1300 it faced the threat of invasion from the south-east (Appendix Table 7). Portugal and Spain faced a threat from North Africa through the sixteenth and seventeenth centuries, but they typically dealt with it through offensive actions (e.g., Charles V's conquest of Tunis 1535—a response to raids by Hayreddin Barbarossa along the southern Italian coast—and Sebastian I's invasion of Morocco, which ended in his defeat at the Battle of Alcácer Quibir in 1578).

5.1 *Locations of Capital Cities*

Our model predicts that the threat of external invasion is an important determinant of the location of an empire's capital. There are numerous examples of empires changing capitals to confront their external enemies more effectively; we focus on examples from China and Europe.



(a) Han Dynasty (206 BC–220 AD)

(b) Roman Empire, Tetrarchy, circa 300 AD

Figure 16: Capital cities in China and the Roman Empire. Panel (a) depicts the Han dynasty's capital city, Changan, and its most populous prefectures. Beijing replaced Changan as China's preeminent political center in the second millennium. Panel (b) depicts the major cities of the Roman Empire. During the Tetrarchy period, there were four capitals: Trier, Milan, Sirmium, and Nicomedia. Constantinople and Ravenna were the capitals of the Eastern and Western Empires respectively. Carthage, Alexandria, and Antioch were the largest cities after Rome itself. The maps are adapted from Herrmann (1966) and Talbert (2000).

Consistent with our model, for most of its history, China's capital city was located in its northern or northwestern frontier instead of the populous Central Plain or Lower Yangzi Delta. For the 1,418 years between 221 BC and AD 1911 when China was under unified rule, Beijing and Changan (modern-day Xi-an) served as its national capital for 634 years and 553 years, respectively, or together 8.4 years out of every 10 years (Wilkinson, 2012).

Changan was China's preeminent political center in the first millennium AD. It was the capital city of the unified dynasties of Qin (221–206 BC), Former Han and Xin (202 BC–AD 23), Sui (581–618), and Tang (618–907). Figure 16a illustrates two salient characteristics of Changan's geographical location that buttress our argument: (1) it was not the population or economic center of the empire;[29] and (2) it was strategically placed to shield China's populous Central Plain from nomadic invasions.

In the second millennium, when China's main threat shifted eastward from Inner Asia to the semi-nomadic lands of Manchuria, Beijing replaced Changan as the new political center of China due to its proximity to the northeastern frontier.

For the European case, our evidence comes from the Roman Empire—the single long-lasting empire to span much of the continent in European history. The Roman Republic and Empire expanded symmetrically from the city of Rome over several centuries to encompass the entire Mediterranean and western Europe. Over time, therefore, the location of the capital became less and less convenient from the viewpoint of military operations. As the severity of the external threats facing the empire grew from the mid-second century onward, emperors spent less and less time in Rome, and they eventually set up other capital cities in which to reside.

Importantly, these new capitals were not the largest cities in the empire. After Rome, the most populous Roman cities were Alexandria with around 600,000 inhabitants, and Antioch and Carthage with between 300,000–500,000 people each (Scheidel, 2013, 78). However, with the exception of Antioch, these cities were far from the frontiers and were never chosen as capital cities.

When the emperor Constantine (r. 306–338) established a new permanent capital at the small Greek city of Byzantine, renamed as Constantinople, he chose this location not because it had any economic significance at the time, but because it was close to both the eastern frontier of the empire and to the important Danube front where the empire faced some of its most determined enemies.[30]

[29] During the Han dynasty, for example, only 3 million people, or around 4 percent of the Chinese population, resided in Guanzhong, the region in which Changan was located. By contrast, the Guandong region in central-eastern China was home to 60 percent of the empire's subjects.

[30] For example, see Odahl (2004, 232) and Goldsworthy (2009, 186).

	Per Capita Tax Revenue (silver grams)			Total Tax Revenue (silver tons)		
	1700	1750	1780	1700	1750	1780
England	91.9	109.1	172.3	559.4	821.1	1627.3
France	43.5	48.7	77.6	878.2	1081.2	1962
Dutch Republic	210.6	189.4	228.2	400.6	367.6	466.8
Spain	28.6	46.2	59.0	219.2	439.3	642.5
European average	52.1	58.0 (27%)	77.3	278.2	403.2	711.5
China	10.4	11.8 (6%)	9.2	1812.1	2633.3	2769.3

Table 5: Per capita tax revenue in grams of silver. European average tax revenue includes Venice, Austria, Russia, Prussia, and Poland-Lithuania in addition to England, France, Dutch Republic, and Spain. Sources: Karaman and Pamuk (2013) and Sng (2014). In parentheses, we include a comparison of per capita tax revenue as a proportion of 'bare-bones' subsistence in 1750 as measured by Allen et al. (2011).

5.2 *Taxation*

Our model predicts that taxation would be higher in Europe relative to China (Implication 6). This contradicts traditional comparative accounts of Europe and China, which complained that economic development in China was retarded by high taxes (e.g., Jones, 2003), but it is consistent with recent scholarship in economic history. Tax rates in Europe were high and especially so in the Dutch Republic and England after 1689 (Hoffman and Norberg, 1994; Bonney, 1999).[31] By contrast, taxes were comparatively low in China. Karaman and Pamuk (2013) provide data on tax revenues for a range of European countries. Table 5 depicts this data in conjunction with estimates of per capita and total tax revenue from China (Sng, 2014). The average European per capita level of taxation as measured in silver was roughly four times higher than that in China. As China was a net importer of silver, the value of silver in China might have been higher than in Europe. As a check, we use the bare-bones subsistence basket constructed in Allen et al. (2011) to estimate the tax burden in Europe and China, and obtain similar results. Clearly, as Implication 6 suggests, taxation was lighter under politically centralized China than it was in fragmented Europe.

[31] This increase in taxes obviously implies that tax revenues of the central state were lower before 1700. However, it does not mean that the total level of taxation was lower before 1700 when one includes feudal dues, local taxes, and tithes to the Church. Much of the increase in tax revenue collected by centralized states came at the expense of these forms of taxation, which were gradually abolished after the seventeenth century (Johnson and Koyama, 2014a).

5.3 *Population cycles in China and Europe.*

Implication 7 predicts that population growth should be more variable under political centralization because political centralization is associated with lower taxes during peacetime but also greater vulnerability to external shocks. We provide evidence in support of this prediction by drawing on population data from China and Europe.

Pre-modern population data is of variable quality. McEvedy and Jones (1978) provide imperfect but comparable population estimates for both China and Europe for the past two thousand years. Figure 17a presents these estimates. It shows that the population growth of China was more variable than that of Europe. Figure 17b, which shows the implied annual population growth rate, confirms this finding. It is evident that the time series of Chinese population are more scattered and display greater variance.[32] Interestingly, there is no visible difference in population variation at the two ends of Eurasia when they were ruled by empires (before AD 400) and when they were fragmented (400–600); it is only after the consolidation of political centralization in China and fragmentation in Europe that significant differences in population patterns emerged.

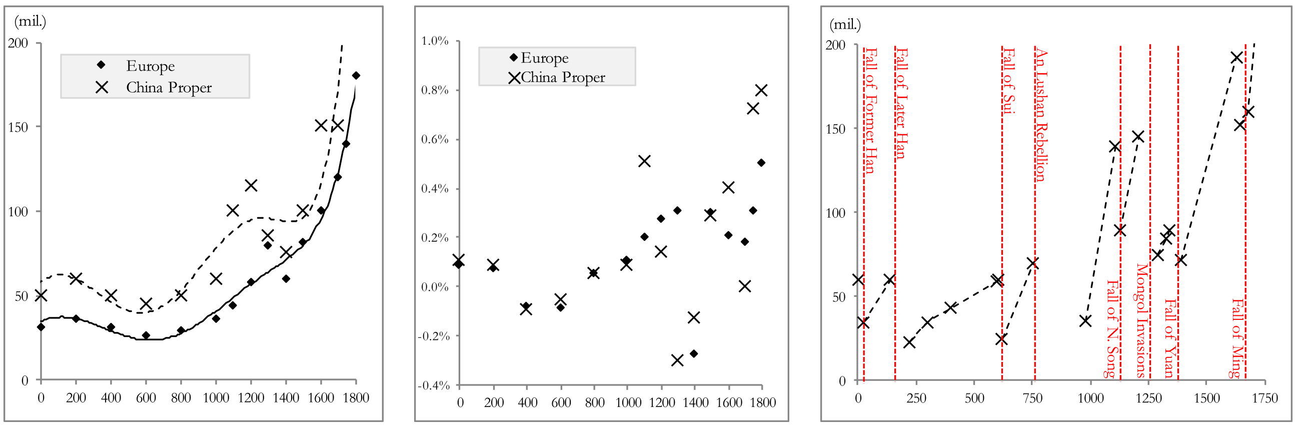
Since McEvedy and Jones (1978) report data for every 50, 100, or 200 years, the resulting time series is necessarily smoother than would be the case if data was available at a higher frequency. This potential problem biases us against finding a difference between the population fluctuations in China and Europe as there are several well-known sharp declines in Chinese population that are either absent or moderated in the McEvedy and Jones (1978) data.

Figure 17c displays a higher frequency population series from Cao (2000).[33] This data series

[32]In addition, in Appendix A.7, we fit the population estimates with polynomials up to the sixth order and find that (a) it is easier to fit the European population estimates than it is to fit the Chinese population estimates, and (b) even if we set aside differences in the degree of goodness of fit, Europe's fitted trend line is smoother than the Chinese one.

[33]We use the population estimates provided in Cao (2000) because of the coverage and relative accuracy. The plunges in China's population depicted in Figure 17c would be even more severe if we had used official historical statistics. For example, official historical records suggest that China's population fell from more than 50 million to 7 million in the third century after the collapse of the Han dynasty. A substantial amount of this population 'loss' was

Figure 17: Population Estimates in China and Europe (McEvedy and Jones, 1978; Cao, 2000)



(a) Population level (M&J, 1978) (b) Annual growth (M&J, 1978) (c) China's population (Cao, 2000)

highlights the devastating effect of external invasions and political collapses on the Chinese population. The population effects of the fall of the Xin, Han, Sui dynasties, the An Lushan Rebellion, the Jurchen and Mongol invasions, and the collapse of Yuan and Ming are all visible in the figure.

For example, the Mongol invasions are associated with a sharp population collapse. Kuhn (2009) observes that between 1223 and 1292, China's population 'decreased by roughly 30 million, or one third of the population, to 75 million.' The fall of the Yuan Dynasty is thought to have caused the population to fall again by up to 25 percent. In contrast, there was only one major Europe-wide collapse in population after the fall of Rome: the Black Death of the mid-fourteenth century.

6 Alternative Hypotheses

Although we find that nomadic attacks Granger-caused political unification in China, this does not mean that the nomadic factor was the only force driving China's recurring unification. In fact, the historical patterns we reviewed suggest that there were other important factors at work too.

A popular explanation posits that European (Chinese) geography was less (more) conducive to political centralization due to the presence (absence) of irregular coastlines and mountain barriers

likely due to the state's inability to keep accurate records during times of crises instead of actual deaths. By contrast, Cao (2000) puts the late third-century estimate at 23 million instead of 7 million.

such as the Alps or Pyrenees (Kennedy, 1987; Diamond, 1997).[34] Some scholars view the need for coordination to deal with frequent flooding along the silt-laden Yellow River as a critical factor that facilitated China's early political unification (Needham and Huang, 1974).[35] Others point to institutional and cultural factors such as the imperial examination system in China and the papacy in Europe to explain the patterns of political consolidation and fragmentation observed (Fairbank, 1992; Hoffman, 2015). On top of these persistent forces, historical contingency cannot be ignored either. The Investiture Controversy in medieval Europe and the wars of religion in the sixteen and seventeenth centuries are two examples of how unpredictable events could have further augmented the tendency towards political fragmentation in Europe. Likewise, the adoption of a logographic writing system in China and a phonographic one in Europe were random events that nonetheless reinforced their respective centripetal and centrifugal tendencies.

Our argument and these existing explanations are not mutually exclusive. In fact, they complement and reinforce each other. For example, while the presence of multi-sided threats could shed light on why Europe was fragmented in the Middle Ages (Section 4.2), our argument cannot explain why the number of states in Europe continued to rise after 1100 (as depicted in Figure 2), when external threats to the continent subsided considerably. Tilly (1990) addresses this, arguing that the presence of independent city states along the corridor between southern England and northern Italy prevented the emergence of large empires in Europe at the end of the Middle Ages.[36] More recently, Hoffman (2015) also suggests that the Catholic Church played a crucial role in preventing first the Holy Roman Emperor, and later the Habsburgs, from dominating Europe.

However, Tilly's theory does not explain the existence of independent city states in late medieval

[34]This argument was recently challenged by Hoffman (2015), who finds that the average elevation is in fact higher in China than in Europe as internal mountain barriers divide southern China from northern China, whereas northern Europe is flat and no natural barrier separates France from Germany.

[35]Our model can actually be applied to develop this argument if we expand the definition of threats to include recurring natural disasters and reinterpret military investment as investment in state capacity.

[36]Spruyt (1994) and Finer (1999) also develop similar arguments.

Europe, which was a legacy of the collapse of the Carolingian empire (see Pirenne, 1925, 1936). Similarly, Hoffman takes as given the papacy's rise to secular authority in medieval Europe—a phenomenon made possible by the existence of numerous rival kingdoms and the absence of a legitimate and powerful European hegemon in the eleventh and twelfth centuries (Morris, 1989; Miller, 2005; Heather, 2014). In this regard, our argument complements the above theories by highlighting how Europe's external threats contributed to its political fragmentation before the eleventh century. This "initial" fragmentation then proved persistent because the absence of a single long-lasting hegemonic state made it easier for the Catholic Church to entrench itself as a genuine rival to the secular powers of Europe and allowed small independent city states like Genoa, Florence, and Venice to emerge and contest the power of territorial states.

In the case of China, our argument also cannot explain why periods of disunity became significantly shorter in the second millennium, especially after 1300 (see Figure 2). Chinese historians often attribute the phenomenon to the gradual strengthening of political institutions that checked the centrifugal forces. An important example is the imperial examinations, which fundamentally undermined the local aristocratic families that were hugely influential in medieval China (Miyakawa, 1955; Fairbank, 1992). However, the Chinese imperial examination system did not emerge overnight. As Fairbank (1992) points out, its perfection as a tool for training compliant bureaucrats took a long time and involved the efforts of successive unified dynasties. The full development of the system in the early modern period therefore represents the fruits of a long-run process of political unification. Once this institution was in place, it became an important driver that reinforced the persistence of political centralization in China. This example, again, highlights the complementarity between our argument and existing theories.

7 Conclusion

The idea that Europe's political and economic success is related to its political fragmentation goes back to the Enlightenment. Montesquieu noted that, in contrast to Asia where strong nations are able to subdue their neighbors, in Europe 'strong nations are opposed to the strong; and those who join each other have nearly the same courage. This is the reason of the weakness of Asia and of the strength of Europe; of the liberty of Europe, and of the slavery of Asia' (Montesquieu, 1989, 266).

In this paper we have proposed a unified theory of the origins, persistence, and consequences of political centralization and fragmentation in China and Europe. We build on the argument that external threats were a powerful force for political unification in China, but were less of a factor in Europe. Our theory suggests that political centralization should indeed be stable in China, but not in Europe, and that this centralization was beneficial from a static perspective, as it minimized costly interstate competition. However, we also show that, over a longer period, a centralized empire such as China was less robust than a decentralized state system.

Although beyond the scope of our paper, this start-stop nature of population growth in empires that we highlight theoretically and empirically has the potential to be built upon to help reconcile a big puzzle in the history of economic growth: why China, the most populous economy in the world for much of recorded history, was capable of coming 'within a hair's breadth of industrializing in the fourteenth century' (Jones, 2003, 160), but swiftly and permanently lost its technological lead after the devastating wars of the Mongol conquests.

Since more people means more ideas, growth theory often contains a scale-effect that implies that the largest economy should be the first to experience modern economic growth (Kremer, 1993). However, as Aiyar et al. (2008) point out, in a (premodern) world in which technological knowledge is embodied in humans (instead of being stored in computers), the effect of population change on the stock of knowledge is asymmetric: technological knowledge grows slowly with population growth, but regresses swiftly when the population contracts. Pairing this insight with our theory suggests that, because China was more centralized and more vulnerable to negative population shocks, it experienced more frequent interruptions in cumulative innovation. In other words, China's higher variance of population growth could have diminished its chances of escaping the Malthusian trap, while the European population and economy were able to expand gradually to the point at which the transition from stagnation to growth was triggered (as in theories of unified growth. See Galor and

Weil, 2000; Galor, 2011).[37]

In sum, our theory provides a novel channel through which geography could have helped shape economic outcomes in Eurasia. Scholars have argued that decentralization gave Europe an edge in the Great Divergence because it led to greater innovation (Mokyr, 1990; Diamond, 1997; Lagerlof, 2014), support for merchants (Rosenberg and Birdzell, 1986), or political freedoms and representation (Hall, 1985). Recent work has also shown how the consequences of political fragmentation interacted with the Black Death to raise incomes and increase urbanization in Europe (Voigtländer and Voth, 2013b). Our theory complements these important arguments by highlighting the significance of a previously neglected consequence of political centralization in China. There were periods of economic expansion in China, but these were brought to a halt by external invasions and political crises. It was these population crises, we conjecture, that help to explain why China did not enter a period of sustained economic growth in the preindustrial era. By contrast, Europe's polycentric system of states gave it the institutional robustness that was one of the preconditions for modern economic growth to occur.

References

Abramson, S., "The Economic Origins of the Territorial State," mimeo, 2013.

Acemoglu, D. and J. Robinson, *The Economic Origins of Dictatorship and Democracy* (Cambridge: Cambridge University Press, 2006).

Aiyar, S., C.-J. Dalgaard, and O. Moav, "Technological Progress and Regress in Pre-industrial Times," *Journal of Economic Growth* 13 (2008), 125–144.

Alesina, A. and E. Spolaore, "On the Number and Size of Nations," *Quarterly Journal of Economics* 112 (1997), 1027–56.

Alesina, A. and E. Spolaore, *The Size of Nations* (Cambridge, MA: MIT Press, 2003).

[37] Theoretical and empirical work suggests that what distinguishes modern developed economies is the ability to sustain positive GDP growth for long periods of time (see Jones and Olken, 2008; Che et al., 2013).

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

Allen, R. C., J.-P. Bassino, D. Ma, C. Moll-Murata, and J. L. V. Zanden, "Wages, Prices, and Living Standards in China, 1738–1925: In Comparison with Europe, Japan, and India," *Economic History Review* 64 (2011), 8–38.

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

Bai, Y. and J. Kung, "Climate Shocks and Sino-nomadic Conflict," *Review of Economics and Statistics* 93 (2011), 970–981.

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

Besley, T. and M. Reynal-Querol, "The Legacy of Historical Conflict: Evidence from Africa," *American Political Science Review* 108 (2014), 319–336.

Bonney, R. (Ed.), *The Rise of the Fiscal State in Europe c.1200–1815* (Oxford: Oxford University Press, 1999).

Brecke, P., "Violent Conflicts 1400 A.D. to the Present in Different Regions of the World," Paper presented at the 1999 meeting of the Peace Science Society (1999).

Brouwer, C. and M. Heibloem, *Irrigation Water Management, FAO Training manual no. 3* (Rome: Food and Agriculture Organization, 1986).

Campante, F., Q.-A. Do, and B. Guimaraes, "Capital Cities, Conflict, and Misgovernance," mimeo, 2016.

Cao, S., *Zhongguo renkou shi [A Demographic History of China]* (Shanghai: Fudan University Press, 2000).

Carneiro, R. L., "A Theory of the Origin of the State," *Science* 169 (1970), 733–738.

Chaliand, G., *Nomadic Empires* (New Brunswick: Transaction Books, 2005).

Che, J., K.-S. Chung, and X. Qiao, "The good, the bad, and the civil society," *Journal of Public Economics* 106 (2013), 68–76.

Childe, V., *Man Makes Himself* (New York: Spokesman, 1936).

China's Military History Editorial Committee, *Zhongguo lidai zhanzheng nianbiao [A Chronology of Warfare in Dynastic China]* (Beijing: PLA Press, 2003).

Collins, R., *Charlemagne* (Toronto: Toronto University Press, 1998).

Cowen, T., "Economic Effects of a Conflict-prone World Order," *Public Choice* 64 (1990), 121–134.

Di Cosmo, N., *Ancient China and its Enemies* (Cambridge: Cambridge University Press, 2002).

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

Dincecco, M., "Fiscal Centralization, Limited Government, and Public Revenues in Europe, 1650–1913," *Journal of Economic History* 69 (2009), 48–103.

Dincecco, M. and G. Katz, "State Capacity and Long-run Economic Performance," *Economic Journal* 126 (2016), 189–218.

Elvin, M., *Patterns of the Chinese Past* (Stanford: Stanford University Press, 1973).

Fairbank, J., *China: A New History* (Cambridge, MA: Belknap, 1992).

Finer, S., *The History of Government from the Earliest Times*, Volume II (Oxford: Oxford University Press, 1999).

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

Friedman, D., "A Theory of the Size and Shape of Nations," *Journal of Political Economy* 85 (1977), 59–77.

Fukuyama, F., *The Origins of Political Order* (New York: Farrar, Straus and Giroux, 2011).

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

Galor, O. and D. Weil, "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review* 90 (2000), 806–828.

Garfinkel, M. and S. Skaperdas, "Economics of Conflict: An Overview (Chapter 22)," in T. Sandler and K. Hartley, eds., *Handbook of Defense Economics*, Volume II (Elsevier, 2007), 649–709.

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

Gennaioli, N. and H.-J. Voth, "State Capacity and Military Conflict," *Review of Economic Studies* 82 (2015), 1409–1448.

Goldsworthy, A., *How Rome Fell* (New Haven: Yale University Press, 2009).

Greif, A., M. Iyigun, and D. Sasson, "Social Institutions and Economic Growth: Why England and not China Became the First Modern Economy," mimeo, 2012.

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

Gu, J., *Zhongguo lidai jinian shouce [Handbook of the Annals of China's Dynasties]* (Shanghai: Shanghai Ancient Books Press, 1995).

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

Harris, R. and R. Sollis, *Applied Time Series Modelling and Forecasting* (West Sussex: Wiley, 2003).

Heather, P., *The Fall of the Roman Empire: A New History of Rome and the Barbarians* (Oxford: Oxford University Press, 2006).

Heather, P., *The Restoration of Rome* (Oxford: Oxford University Press, 2014).

Herrmann, A., *An Historical Atlas of China* (Chicago: Aldine, 1966).

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

Hoffman, P., "Why Was it Europeans Who Conquered the World?" *Journal of Economic History* 72 (2012), 601–633.

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

Hoffman, P. and K. Norberg (Eds.), *Fiscal Crises, Liberty, and Representative Government* (Stanford: Stanford University Press, 1994), 1450–1789.

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

Iyigun, M., "Luther and Suleyman," *Quarterly Journal of Economics* 123 (2008), 1465–1494.

Iyigun, M., N. Nunn, and N. Qian, "Agricultural Productivity and Conflict: Evidence from Potatoes, 1400–1900," mimeo, 2015.

Johnson, N. and M. Koyama, "Legal Centralization and the Birth of the Secular State," *Journal of Comparative Economics* 41 (2013), 959–978.

Johnson, N. and M. Koyama, "Tax Farming and the Origins of State Capacity in England and France," *Explorations in Economic History* 51 (2014a), 1–20.

Johnson, N. and M. Koyama, "Taxes, Lawyers, and the Decline of Witch Trials in France," *Journal of Law and Economics* 57 (2014b), 77–112.

Jones, B. and B. Olken, "The Anatomy of Start-stop Growth," *Review of Economics and Statistics* 90 (2008), 582–587.

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

Karaman, K. and Ş. Pamuk, "Different Paths to the Modern State in Europe: The Interaction Between Warfare, Economic Structure and Political Regime," *American Political Science Review* 107 (2013), 603–626.

Kelly, M., J. Mokyr, and C. Ó Gráda, "Precocious Albion: A New Interpretation of the British Industrial Revolution" *Annual Review of Economics* 6 (2014), 363–389.

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

Ko, C. Y. and T.-H. Sng, "Regional Dependence and Political Centralization in Imperial China," *Eurasian Geography and Economics* 54 (2013), 470–483.

Kremer, M., "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics* 108 (1993), 681–716.

Kuhn, D., *The Age of Confucian Rule* (Cambridge, MA: Harvard University Press, 2009).

Kung, J. and C. Ma, "Autarky and the Rise and Fall of Piracy in Ming China," *Journal of Economic History* 74 (2014), 509–534.

Lagerlof, N.-P., "Population, Technology and Fragmentation: the European Miracle Revisited," *Journal of Development Economics* 108 (2014), 87–105.

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

Levine, D. and S. Modica, "Conflict, Evolution, Hegemony and the Power of the State," Technical Report 19221, NBER Working Paper, 2013.

Luttwak, E., *The Grand Strategy of the Bzyantine Empire* (Cambridge, MA: Belknap, 2009).

Ma, D., "Political Institutions and Long Run Economic Trajectory: Some Lessons from Two Millennia of Chinese Civilization," CEPR Discussion Papers 8791, 2012.

Mayshar, J., O. Moav, Z. Neeman, and L. Pascali, "Cereals, Appropriability and Hierarchy," CEPR Discussion Papers 10742, 2015.

McCloskey, D., *Bourgeois Dignity: Why Economics Can't Explain the Modern World* (Chicago: University of Chicago Press, 2010).

McEvedy, C. and R. Jones, *Atlas of World Population History* (Harmondsworth: Penguin, 1978).

McKitterick, R., *Charlemagne: The Formation of a European Identity* (Cambridge: Cambridge University Press, 2008).

Mencius, *Mencius* (London: Penguin Books, 2004). Translated by D. C. Lau.

Miller, M., *Power and the Holy in the Age of the Investiture Conflict* (New York: Palgrave Macmillan, 2005).

Miyakawa, H., "An Outline of the Naito Hypothesis and its Effects on Japanese Studies of China," *Far Eastern Quarterly* 14 (1955), 533–552.

Mokyr, J., *The Lever of Riches* (Oxford: Oxford University Press, 1990).

Mokyr, J., "The Market for Ideas and the Origins of Economic Growth in Eighteenth Century Europe," *Tijdschrift voor Sociale en Economische Geschiedenis* 4 (2007), 3–38.

Montesquieu, C., *The Spirit of the Laws* (Cambridge: Cambridge University Press, 1989).

Morris, C., *The Papal Monarchy: the Western Church from 1050 to 1250* (Oxford: Clarendon, 1989).

Morrissey, R., *Charlemagne & France* (Notre Dame: University of Notre Dame Press, 1997).

Needham, J. and R. Huang, "The Nature of Chinese Society: A Technical Interpretation," *Journal of Oriental Studies* 12 (1974), 1–16.

Ng, S. and P. Perron, "Unit Root Tests in ARMA Models with Data-dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association* 90 (1995), 268–281.

North, D., *Structure and Change in Economic History* (New York: W. W. Norton, 1981).

North, D., J. Wallis, and B. Weingast, *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History* (Cambridge: Cambridge University Press, 2009).

Nussli, C., URL: http://www.euratlas.com/about.html (2010).

Odahl, C., *Constantine and the Christian Empire*, 2nd ed. (London: Routledge, 2004).

Olson, M., "Dictatorship, Democracy, and Development," *American Political Science Review* 87 (1993), 567–576.

Olsson, O. and G. Hansson, "Country Size and the Rule of Law: Resuscitating Montesquieu," *European Economic Review* 55 (2011), 613–629.

Perdue, P., *China Marches West: The Qing Conquest of Central Eurasia* (Cambridge: Belknap, 2005).

Pirenne, H., *Medieval Cities* (New York: Doubleday Anchor Books, 1925).

Pirenne, H., *Economic and Social History of Medieval Europe* (London: Routledge, 1936).

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

Rosenthal, J.-L. and B. Wong, *Before and Beyond Divergence* (Cambridge: Harvard University Press, 2011).

Rubin, J., "Institutions, the Rise of Commerce and the Persistence of Laws: Interest Restrictions in Islam and Christianity," *Economic Journal* 121 (2011), 1310–1339.

Scheidel, W., "From the 'Great Convergence' to the 'First Great Divergence': Roman and Qin-Han State Formation and its Aftermath," in W. Scheidel, eds., *Rome and China: Comparative Perspectives on Ancient World Empires* (Oxford: Oxford University Press, 2009), 11–23.

Scheidel, W., "Demography," in W. Scheidel, I. Morris, and R. Saller, eds., *Cambridge Economic History of the Greco-Roman World* (Cambridge: Cambridge University Press, 2013), 38–86.

Shen, K., *Mengxi bi tan [Brush talks from dream brook]* (London: Paths International, 2011).

Sng, T.-H., "Size and Dynastic Decline: The Principal-agent Problem in Late Imperial China, 1700–1850," *Explorations in Economic History* 54 (2014), 107–127.

Sng, T.-H. and C. Moriguchi, "Asia's Little Divergence: State Capacity in China and Japan Before 1850," *Journal of Economic Growth* 19 (2014), 439–470.

Spruyt, H., *The Sovereign State and Its Competitors* (Princeton: Princeton University Press, 1994).

Talbert, R., *Barrington Atlas of the Greek and Roman World* (Princeton: Princeton University Press, 2000).

Tilly, C. (1990). *Coercion, Capital, and European States, AD 990–1990* (Oxford: Blackwell, 2000).

Treadgold, W., *A History of the Byzantine State and Society* (Stanford: Stanford University Press, 1997).

Turchin, P., "A Theory for Formation of Large Empires," *Journal of Global History* 4 (2009), 191–217.

van Creveld, M., *Supplying War: Logistics from Wallenstein to Patton* (Cambridge: Cambridge University Press, 2004).

Voigtländer, N. and H.-J. Voth, "How the West Invented Fertility Restriction," *American Economic Review* 103 (2013a), 2227–2264.

Voigtländer, N. and H.-J. Voth, The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe. *Review of Economic Studies* 80 (2013b), 774–811.

Vries, P., *State, Economy, and the Great Divergence: Great Britain and China, 1680s–1850s* (London: Bloomsbury, 2015).

Ward-Perkins, B., *The Fall of Rome* (Oxford: Oxford University Press, 2005).

Wei, L., *Zhongguo wenshi jianbiao huibian [Concise Tables of Chinese Cultural and Historical Materials]* (Beijing: Commercial Press, 2011).

White, M., *Atrocities: The 100 Deadliest Episodes in Human History* (New York: W. W. Norton, 2013).

Whittow, M., *The Making of Byzanium, 600–1025* (Berkeley: University of California Press, 1996).

Wilkinson, E., *Chinese History: A New Manual* (Cambridge: Harvard University Asia Center, 2012).

Zhang, B., F. Feng, L. Zhang, and H. Li, *Zhongguo nongye ziran zaihai shiliao ji [Historical Materials on Natural Disasters in Chinese Agriculture]* (Xi'an: Shaanxi Technology Press, 1994).

Zhang, D., Q. Pei, H. Lee, J. Zhang, C. Q. Chang, B. Li, J. Li, and X. Zhang "The Pulse of Imperial China: A Quantitative Analysis of Long-term Geopolitical and Climatic Cycles," *Global Ecology and Biogeography* 24 (2015), 87–96.

Zheng, P., *Livestock breeds of China, Animal Production and Health Papers vol. 46* (Rome: Food and Agriculture Organization, 1984).

A Appendix

A.1 Eurasia's External Threats

Phase	Century	Nomadic Peoples	W. Europe	Russia	China
1st	8th–2nd BC	Scythians		✓	
2nd	2nd BC	Xiongnu (Hsiung-nu)			✓
	2nd BC	Yuezhi (Yüeh-chih)			✓
3rd	4th	To-pa (Toba Turks)			✓
	5th	Huns	✓	✓	
4th	6th	Tu-chueh (Göktũrks)			✓
	6th	Juan-juan			✓
	7th	Avars	✓	✓	
	8th	Bulgars	✓		
	7th	Khazar Turks		✓	
	9th	Magyars	✓		
5th	9th	Uyghurs			✓
	10th	Khitans			✓
	11th	Pechenegs and Kipchaks		✓	
	12th	Jurchens			✓
6th	13th	Mongols		✓	✓
	14th	Tatars		✓	✓
7th	15th	Oirots			✓
	17th	Manchus			✓

Table 6: Major waves of nomadic invasions. Source: Chaliand (2005). It is evident that China faced a greater threat from the steppe invaders than did Europe. See Section 2 for a historical discussion.

Invader	Date of Invasion	Location of Invasion	Direction of threat
Huns	c. 370–450	Italy, France, Balkans	East
Avars	580	South-Eastern Europe	East
Bulgars	c. 850	South-Eastern Europe	East
Arabs	711	Spain	South
Arabs	721	France	South
Arabs	732	France	South
Vikings*	793–1066	Britain	North
Vikings*	c. 810–1000	France	North
Vikings*	c. 810–1000	Low Countries	North
Arabs	831	Sicily	South
Arabs	840	Crete	South
Arabs	846	Italy	South
Magyars	907	Germany	East
Magyars	917	France	East
The Almohads	1172	Spain	South
Mongols	1240	Poland	East
Mongols	1241	Hungary	East
Mongols	1241	Croatia	East
The Marinids	1340	Gibraltar	South
Ottomans	1371	Serbia	South-East
Ottomans	1385	Albania	South-East
Ottomans	1463	Bosnia	South-East
Ottomans	1479	Hungary	South East
Ottomans	1480	Italy	South East
Crimean Tatars	1480–1507	Poland	East
Crimean Tatars	1507–1570	Russia	East
Ottomans	1526	Hungary	South-East
Ottomans	1529	Austria	South-East
Ottomans	1541	Hungary	South-East
Ottomans	1565	Malta	South-East
Ottomans	1573	Cyprus	South-East
Crimean Tatars	1571	Russia	East
Crimean Tatars	1577, 1584, 1590	Poland	East
Ottomans	1573	Cyprus	South-East
Crimean Tatars	1621	Poland	East
Crimean Tatars	1648	Poland	East
Crimean Tatars	1667	Poland	East
Ottomans	1669	Crete	South-East
Ottomans	1683	Austria	South-East

Table 7: The Multidirectional Threat in Europe. A list of invasions of Europe from the Fall of the Roman Empire onward. See Section 2 for a historical discussion. We list invasion attempts that failed as well as those that succeeded. *We count the Vikings as external invaders—due to their different religion and distinct culture they were seen as outsider invaders by contemporaries. But from the perspective of our model, we consider later Swedish or Danish campaigns in Europe as instances of interstate competition.

A.2 Data

A.2.1 Conflict Data

The Brecke data is available at `http://www.cgeh.nl/data#conflict`. The dataset reports 3708 conflicts for the period 1400–1800. It contains information on the parties involved in a conflict, the date the conflict began, and the date that it ended. It is a superset of previous datasets that combines all conflicts from a large number of existing datasets on conflict and is based on a large number of secondary sources. The definition of a conflict is 32 violent deaths per year. We remove political persecutions (such as persecutions of Jews) so as to focus on military conflicts. Further details of the dataset are provided in Brecke (1999).

A.2.2 Number of Deaths Data

Information on the number of casualties or deaths from premodern wars is highly speculative. Our source for this information is a recent compendium of military deaths (White, 2013). From our perspective, however, only the relative magnitudes of war deaths between Europe and East Asian wars matter. In this respect, the high death tolls reported for conflicts such as the Mongol Invasions of China, the Ming-Qing transition, and the Taiping Rebellion are all borne out by recent research.

A.3 Proofs of Propositions in Section 3

We now provide the proofs of Propositions 1 and 2 in Section 3. We use more general functional forms here to show that our results do not depend on the specific forms assumed in the main paper.

A.3.1 Setup Without Parametric Form Assumptions

The setup of the model is unchanged. We replace the parametric forms in Section 3 with more general functional forms:

- The gross damage function is now $\max\{\lambda(\Lambda, t), 0\}$, where t is the distance from the frontier, $\lambda_1 > 0$, and $\lambda_2 < 0$;
- For regime $i \in \{e, l, r\}$ to invest $M_i \geq 0$, it costs $c(M_i)$, where $c(0) = 0$, $c' > 0$, and $c'' > 0$;
- For a location that is t distance away from G_i, regime i's **military strength** on that location is $\max\{m(M_i, t, \beta), 0\}$.

In addition, we make the following assumptions of the military strength function $m(M_i, t, \beta)$ and the border function $b\left(G_l, G_r, M_l, M_r, \beta\right) \in [G_l, G_r]$:

Assumption 1 (Military strength)**.** *The function $m(M, t, \beta)$ satisfies the following properties:*

1. *(Monotonicity) $m_1 > 0$, $m_2 < 0$, $m_3 \leq 0$ (equality if and only if $t = 0$);*

2. *(Distance effect) $m_{22} < 0$, $m_{12} \leq 0$;*

3. *(Effect of Parameter β) $m_{33} \leq 0$, $m_{13} \leq 0$, $m_{23} \leq 0$.*

Assumption 1 is straightforward. Point 1 states that a regime's military strength is increasing in its military investment, decreasing in distance and also in the difficulty of power projection. Point 2 stipulates that military strength deteriorates over distance at an increasing rate. Furthermore, the marginal effectiveness of military investment is decreasing in distance. Point 3 implies that military strength is decreasing in β at an increasing rate. In addition, a higher β would reduce the marginal effectiveness of military investment and increase the decline in military strength over distance.

When $S = 2$, the border b that separates regimes l and r depends on the regimes' capital city locations (G_l and G_r) and military investments (M_l and M_r).

Assumption 2 (Border formation)**.** *The border $b(G_l, G_r, M_l, M_r, \beta) \in [G_l, G_r]$ satisfies the following:*

1. *(Monotonicity) $\frac{\partial b}{\partial M_l} > 0$, $\frac{\partial b}{\partial M_r} < 0$, $\frac{\partial b}{\partial G_l} > 0$, and $\frac{\partial b}{\partial G_r} > 0$;*

2. *(Concavity) $\frac{\partial^2 b}{\partial M_l^2} \leq 0$, $\frac{\partial^2 b}{\partial M_r^2} \geq 0$, $\frac{\partial^2 b}{\partial G_l^2} \leq 0$, and $\frac{\partial^2 b}{\partial G_r^2} \geq 0$;*

3. *(Symmetry) When $G_l = 1 - G_r$ and $M_l = M_r$,*
   $\frac{\partial b}{\partial \beta} = 0$, $\frac{\partial^2 b}{\partial \beta \partial G_l} = -\frac{\partial^2 b}{\partial \beta \partial G_r} \leq 0$, $\frac{\partial^2 b}{\partial G_l^2} \leq \frac{\partial^2 b}{\partial G_l \partial G_r} \leq \frac{\partial^2 b}{\partial G_r^2}$.

According to Point 1 of Assumption 2, a regime expands territorially when it increases its military investment or moves its capital city toward its rival. Point 2 ensures that the SOCs are satisfied. Point 3 states that when the two regimes are symmetric, the effect of an increase in β on the border is zero (since the effects on both sides cancel out each other); a higher β decreases the marginal gain of moving the capital closer to the competitor; the diminishing returns to moving the capital toward the competitor are strong enough so that the cross derivatives are always smaller in magnitude than the second derivatives.

A.3.2 Proof of Proposition 1

Proof. Consider $x^*(\Lambda) \in [0, 1]$ such that $\lambda(\Lambda, x^*(\Lambda)) = 0$. In other words, $x^*(\Lambda)$ is the leftmost location where the gross damage caused by the threat emanating from the left is zero. If such $x^*(\Lambda)$ does not exist, let $x^*(\Lambda) \equiv 1$.

First, consider a two-sided threat.

1. If $\Lambda \leq \Lambda_I$, the revolution constraint is not violated even if military investment is zero. Since regime e's payoff is decreasing in its military investment, it is optimal to invest zero regardless of the location of the capital city.

2–3. Given that the two-sided threat is symmetric, there is no loss of generality if we assume $G_e \leq 1/2$. Since regime e always receives gross tax revenue y, maximizing its net tax revenue is equivalent to minimizing its military expenditure:

$$\min_{G_e, M_e} c\left(M_e\right)$$

subject to

$$\begin{aligned} \lambda\left(\Lambda, a\right) &\leq m\left(M_e, G_e - a, \beta\right), \text{and} \\ \lambda\left(\Lambda, 1-a-\delta\right) &\leq m\left(M_e, a+\delta-G_e, \beta\right), \end{aligned}$$

where $a \geq 0$ is the leftmost location in the empire that suffers zero net damage when the military investment is M_e. If $M_e = 0$ satisfies both inequalities above, then the optimal solution is a corner one in which regime e invests zero in the military. This is the case when $\Lambda \leq \Lambda_I$ (as discussed above).

Now consider the case in which the solution is interior with a positive military investment, i.e., $M_e > 0$. Note that the first inequality must bind in equilibrium. Otherwise, regime e can increase its net tax revenue by reducing M_e and increasing G_e. Hence, we have

$$\lambda\left(\Lambda, a\right) = m\left(M_e, G_e - a, \beta\right).$$

Next consider (i) the case in which the second inequality does not bind, and (ii) the case in which it binds.

Case (i): if $\lambda\left(\Lambda, 1-z\right) < m\left(M_e, z-G_e, \beta\right)$, it follows that $G_e^* = a$. Otherwise, regime e can increase its net tax revenue by reducing M_e and G_e simultaneously. The same argument also implies that $G_e^* = \max\{0, 1-x^*\left(\Lambda\right)-\delta\}$. Therefore, the optimal military spending M_e^* must satisfy:

$$\lambda\left(\Lambda, \max\{0, 1-x^*\left(\Lambda\right)-\delta\}\right) = m\left(M_e^*, 0, \beta\right).$$

Case (ii): $\lambda\left(\Lambda, 1-z\right) = m\left(M_e, z-G_e, \beta\right)$ because Λ exceeds some threshold (since $\partial\lambda/\partial\Lambda > 0$).

Let Λ_{II} denote this threshold (which we will define later). Since $\Lambda \geq \Lambda_{II}$,

$$\lambda(\Lambda, a) = m(M_e^*, G_e^* - a, \beta), \text{and}$$
$$\lambda(\Lambda, 1-\delta-a) = m(M_e^*, \delta + a - G_e^*, \beta).$$

If $G_e \leq 1/2$, then $a \leq 1-\delta-a$ and $\delta+a-G_e \geq G_e-a$ because military strength is symmetric about the capital city and the threats are symmetric from the two frontiers. Applying total differentiation,

$$\begin{bmatrix} \lambda_2(\Lambda, a) + m_2(M_e^*, G_e^* - a, \beta) & -m_1(M_e^*, G_e^* - a, \beta) \\ -\lambda_2(\Lambda, 1-\delta-a) - m_2(M_e^*, \delta + a - G_e^*, \beta) & -m_1(M_e^*, \delta + a - G_e^*, \beta) \end{bmatrix} \begin{bmatrix} da \\ dM_e \end{bmatrix}$$
$$= \begin{bmatrix} -\lambda_1(\Lambda, a) \\ -\lambda_1(\Lambda, 1-\delta-a) \end{bmatrix} d\Lambda + \begin{bmatrix} m_2(M_e^*, G_e^* - a, \beta) \\ -m_2(M_e^*, \delta + a - G_e^*, \beta) \end{bmatrix} dG_e.$$

It is easy to show that

$$\Delta = \begin{vmatrix} \lambda_2(\Lambda, a) + m_2(M_e^*, G_e^* - a, \beta) & -m_1(M_e^*, G_e^* - a, \beta) \\ -\lambda_2(\Lambda, 1-\delta-a) - m_2(M_e^*, \delta + a - G_e^*, \beta) & -m_1(M_e^*, \delta + a - G_e^*, \beta) \end{vmatrix}$$
$$> 0 \text{ because } m_1 > 0,\ \lambda_2 < 0 \text{ and } m_2 < 0.$$

Hence

$$\frac{dM_e}{dG_e} = \frac{\begin{vmatrix} \lambda_2(\Lambda, a) + m_2(M_e^*, G_e^* - a, \beta) & m_2(M_e^*, G_e^* - a, \beta) \\ -\lambda_2(\Lambda, 1-\delta-a) - m_2(M_e^*, \delta + a - G_e^*, \beta) & -m_2(M_e^*, \delta + a - G_e^*, \beta) \end{vmatrix}}{\Delta}$$
$$< 0 \text{ because } \lambda_2 < 0, \lambda_{22} \geq 0, m_2 < 0 \text{ and } m_{22} < 0.$$

Since $\partial c / \partial M_e > 0$ and therefore $\partial V_e / \partial M_e < 0$, we have

$$\frac{dV_e}{dG_e} = \frac{\partial V_e}{\partial M_e} \frac{dM_e}{dG_e} > 0.$$

Similarly, if $G_e > 1/2$, then $\frac{dM_e}{dG_e} > 0$ and $\frac{dV_e}{dG_e} < 0$. Hence, to maximize its net tax revenue, regime e should locate its capital at $1/2$ and $a = \frac{1}{2}(1-\delta)$. The optimal military spending M_e^* satisfies

$$\lambda\left(\Lambda, \frac{1}{2}(1-\delta)\right) = m(M_e^*, \delta/2, \beta).$$

(Λ_{II}, M_e) is the solution to the following system:

$$\lambda\left(\Lambda_{II}, \frac{1}{2}(1-\delta)\right) = m\left(M_e, \delta/2, \beta\right), \text{ and}$$
$$\lambda\left(\Lambda_{II}, \max\left\{0, 1 - x^*\left(\Lambda_{II}\right) - \delta\right\}\right) = m\left(M_e, 0, \beta\right).$$

Finally, consider a one-sided threat.

4. If $\Lambda \leq \bar{\Lambda}_I$, then $x^*(\Lambda) \leq 1 - \delta$ so that the fraction of protected area is no less than δ even if there is no military investment. Since regime e's payoff is decreasing in its military investment, the optimal military investment is zero and the capital city is located between 0 and 1.

5. If $x^*(\Lambda) > 1 - \delta$, regime e has to make a strictly positive military investment. Since military strength decreases over distance ($m_2 < 0$), it should locate its capital city at the point at which the revolution constraint just binds. This implies that $G_e^* = 1 - \delta$ and M_e^* solves $\lambda(\Lambda, 1-\delta) = m(M_e^*, 0, \beta)$. □

A.3.3 Proof of Proposition 2

Before proving the proposition, it is useful to characterize the outcome of interstate competition in the absence of external threats ($\Lambda = 0$).

Lemma 1. *When there is no external threat, given some locations of capital cities G_l and G_r, the equilibrium military investments M_l^* and M_r^* satisfy*

$$\frac{\partial b}{\partial M_l} y - \frac{\partial c}{\partial M_l} = 0, \textit{ and}$$
$$-\frac{\partial b}{\partial M_r} y - \frac{\partial c}{\partial M_r} = 0.$$

For any symmetric equilibrium capital locations G_l^* and G_r^* that satisfy $G_l^* \neq 0$ and $G_r^* \neq 1$ (i.e., interior solutions), we have

$$\left(\frac{\partial b}{\partial G_l} + \frac{\partial b}{\partial M_l}\frac{\partial M_l}{\partial G_l} + \frac{\partial b}{\partial M_r}\frac{\partial M_r}{\partial G_l}\right) y - \frac{\partial c}{\partial M_l}\frac{\partial M_l}{\partial G_l} = 0, \textit{ and}$$
$$-\left(\frac{\partial b}{\partial G_r} + \frac{\partial b}{\partial M_l}\frac{\partial M_l}{\partial G_r} + \frac{\partial b}{\partial M_r}\frac{\partial M_r}{\partial G_r}\right) y - \frac{\partial c}{\partial M_r}\frac{\partial M_r}{\partial G_r} = 0.$$

Proof. First consider the second stage of interstate competition. Given G_l and G_r, the optimization problems for regimes l and r are

$$\max_{M_l} V_l = b\left(G_l, G_r, M_l, M_r, \beta\right) y - c\left(M_l\right), \text{ and}$$
$$\max_{M_r} V_r = (1 - b\left(G_l, G_r, M_l, M_r, \beta\right))y - c\left(M_r\right).$$

The respective FOCs are

$$\frac{\partial b}{\partial M_l} y - \frac{\partial c}{\partial M_l} = 0, \text{ and}$$
$$-\frac{\partial b}{\partial M_r} y - \frac{\partial c}{\partial M_r} = 0,$$

Given the setup, it cannot be an equilibrium for any regime under interstate competition to invest zero in the military, i.e., it must be the case that $M_l^* > 0$ and $M_r^* > 0$. Since $\partial^2 c/\partial^2 M_l < 0$ and $\partial^2 c/\partial^2 M_r < 0$, the SOCs are guaranteed if $\partial^2 b/\partial M_l^2 \leq 0$ and $\partial^2 b/\partial M_r^2 \geq 0$.

The second-stage equilibrium military investments by regime l and regime r are $M_l^*\left(G_l, G_r\right)$ and $M_r^*\left(G_l, G_r\right)$, respectively. Let $b^*\left(G_l, G_r, \beta\right) \equiv b\left(G_l, G_r, M_l^*\left(G_l, G_r\right), M_r^*\left(G_l, G_r\right), \beta\right)$, $c_l^*\left(G_l, G_r\right) \equiv c\left(M_l^*\left(G_l, G_r\right)\right)$, and $c_r^*\left(G_l, G_r\right) \equiv c\left(M_r^*\left(G_l, G_r\right)\right)$.

Consider the first stage, in which regimes l and r decide their capital city locations:

$$\max_{G_l} V_l = b^*\left(G_l, G_r, \beta\right) y - c_l^*\left(G_l, G_r\right), \text{ and}$$
$$\max_{G_r} V_r = (1 - b^*\left(G_l, G_r, \beta\right))y - c_r^*\left(G_l, G_r\right).$$

The respective FOCs are

$$\frac{\partial b^*}{\partial G_l} y - \frac{\partial c_l}{\partial G_l} = 0, \text{ and}$$
$$-\frac{\partial b^*}{\partial G_r} y - \frac{\partial c_r}{\partial G_r} = 0.$$

Hence, we have

$$\left(\frac{\partial b}{\partial G_l} + \frac{\partial b}{\partial M_l}\frac{\partial M_l}{\partial G_l} + \frac{\partial b}{\partial M_r}\frac{\partial M_r}{\partial G_l}\right) y - \frac{\partial c}{\partial M_l}\frac{\partial M_l}{\partial G_l} = 0, \text{ and}$$
$$-\left(\frac{\partial b}{\partial G_r} + \frac{\partial b}{\partial M_l}\frac{\partial M_l}{\partial G_r} + \frac{\partial b}{\partial M_r}\frac{\partial M_r}{\partial G_r}\right) y - \frac{\partial c}{\partial M_r}\frac{\partial M_r}{\partial G_r} = 0.$$

which, given the FOCs in the second stage, implies

$$\frac{\partial b}{\partial G_l}+\frac{\partial b}{\partial M_r}\frac{\partial M_r}{\partial G_l}=0,\text{ and }\frac{\partial b}{\partial G_r}+\frac{\partial b}{\partial M_l}\frac{\partial M_l}{\partial G_r}=0.$$

□

We now present the proof of Proposition 2:

Proof. First, we denote the equilibrium in the absence of external threats by $M_l^*=M_r^*=M^*$ and $G_l^*=1-G_r^*=G^*$.

Now consider a two-sided threat. Let Λ_{III} solve

$$\lambda\left(\Lambda_{III},\frac{1}{2}(1-\delta)\right)=m\left(M^*,G^*-\frac{1}{2}(1-\delta),\beta\right).$$

It is clear that if $\Lambda\le\Lambda_{III}$, the revolution constraint of the two regimes does not bind, and vice versa. It should also be clear that if the threat is sufficiently large, the revolution constraint must be binding.

The proof for a one-sided threat is similar. The only change required is to replace Λ_{III} with $\bar{\Lambda}_{III}$, which solves

$$\lambda\left(\bar{\Lambda}_{III},1-\delta\right)=m\left(M^*,G^*-(1-\delta),\beta\right).$$

□

A.3.4 Technical Note on Implications 4 and 5

We provide the proof of the following claim, which embodies Implications 4 and 5 and characterizes the fiscal viability of empires and competitive states under different threat scenarios:

Claim 3. ***(Viability)***

1. *Under a one-sided threat,* $V_e^*>V_l^*+V_r^*$*;*
2. *Under a two-sided threat, when* $\Lambda\ge\Lambda_{II}$ *and* β *is sufficiently large,* $V_e^*<V_l^*+V_r^*$.

Proof. First, consider the case of a one-sided threat. Suppose that, contrary to Claim 3,

$$V_e^*<V_l^*+V_r^*\ .$$

Then, regime e can mimic the choices of regime l, set $G_e=G_l^*$ and $M_e=M_l^*$, and obtain a payoff that is weakly greater than the sum of the net tax revenues of regimes l and r, which is a contradiction. Hence, it must be the case that

$$V_e^*\ge V_l^*+V_r^*\ .$$

In fact, the inequality has to be strict since regime r makes a non-zero military investment. Hence,

$$V_e^* > V_l^* + V_r^* \ .$$

Next, consider the case of a two-sided threat. For a centralized regime, when $\Lambda > \Lambda_{II}$,

$$\lambda\left(\Lambda, \frac{1}{2}(1-\delta)\right) = m\left(M_e^*, \delta/2, \beta\right),$$

so that

$$\frac{dM_e}{d\beta} = -\frac{m_3}{m_1} > 0.$$

Therefore,

$$\frac{dV_e}{d\beta} = -\frac{\partial c}{\partial M_e}\frac{dM_e}{d\beta} < 0.$$

Since $\frac{d^2 M_e}{d\beta^2} = -\frac{m_1 m_{33} - m_3 m_{13}}{(m_1)^2} \geq 0$,

$$\frac{d^2 V_e}{d\beta^2} = -\frac{\partial^2 c}{\partial M_e^2}\left(\frac{dM_e}{d\beta}\right)^2 - \frac{\partial c}{\partial M_e}\frac{d^2 M_e}{d\beta^2} < 0.$$

Moving on to interstate competition, consider a symmetric interior equilibrium ($G_l^* \neq 0$ and $G_r^* \neq 1$). Given the symmetry, it suffices for us to focus on regime l alone. Recall that when the revolution constraint is binding,

$$\frac{\partial M_l}{\partial \beta} = \frac{\partial M_r}{\partial \beta} \geq 0 \text{ and } \frac{\partial G_l}{\partial \beta} = -\frac{\partial G_r}{\partial \beta} \leq 0,$$

and equality holds when $G_l^* = \frac{1}{2}(1-\delta)$ and $G_r^* = \frac{1}{2}(1+\delta)$.

When β is large enough such that $G_l^* = \frac{1}{2}(1-\delta)$ and $G_r^* = \frac{1}{2}(1+\delta)$, $m\left(M_l, G_l^* - \frac{1}{2}(1-\delta), \beta\right) = \lambda\left(\Lambda, \frac{1}{2}(1-\delta)\right)$ and $m\left(M_r, (1-G_r^*) - \frac{1}{2}(1-\delta), \beta\right) = \lambda\left(\Lambda, \frac{1}{2}(1-\delta)\right)$. It follows that $\frac{\partial V_l}{\partial \beta} = 0$ because $\frac{\partial M_l}{\partial \beta} = \frac{\partial M_r}{\partial \beta} = 0$ and $\frac{\partial G_l}{\partial \beta} = -\frac{\partial G_r}{\partial \beta} = 0$. Since $\frac{dV_e}{d\beta} < 0$ and $\frac{d^2 V_e}{d\beta^2} < 0$, we conclude that when β is sufficiently large, $V_e^* < V_l^* + V_r^*$. □

In the proof above, we impose minimal assumptions on the shapes of $m(\cdot)$ and $b(\cdot)$. If we are willing to assume that $m(\cdot)$ and $b(\cdot)$ adopt certain common specific functional forms, it can be shown that the "sufficiently large" value of β that characterizes Claim 3 can be much lower than the level defined above (details available upon request).

A.3.5 Proof of Positive Tax Reimbursement when $S = 1$

Here, we show that, if the cost function of military investment is sufficiently convex, the empire will provide some tax reimbursement instead of relying solely on building the military to satisfy the revolution constraint.

Claim 4. *If $\frac{\partial c(M)}{\partial M} > \frac{\partial m(M,0,\beta)}{\partial M}$ for all M, regime e provides a strictly positive tax reimbursement ($R_e > 0$).*

Proof. Suppose regime e faces a one-sided threat. (The proof for a two-sided threat is similar.) The optimization problem is given by

$$\max_{G_e, M_e, R_e} V_e = y - c\left(M_e\right) - R_e$$

subject to

$$\begin{aligned}
&R_e \geq 0, \\
&m\left(M_e, G_e - (1-\delta), \beta\right) + R_e = \lambda\left(\Lambda, 1-\delta\right), \text{ and} \\
&m\left(M_e, \left|x - G_e\right|, \beta\right) \geq \lambda\left(\Lambda, x\right) \text{ for some } x \in \left[0, x^*\left(\Lambda\right)\right].
\end{aligned}$$

Let $\bar{x} \in \left[0, x^*\left(\Lambda\right)\right]$ such that $m(M_e, \left|\bar{x} - G_e\right|, \beta) = \lambda\left(\Lambda, \bar{x}\right)$. Since $\lambda_2 < 0$, $\bar{x} \geq G_e$ (otherwise, the empire can increase its net tax revenue by increasing G_e and decreasing M_e or R_e). If $R_e = 0$, then $G_e = 1 - \delta$ since $m_2 < 0$. If $G_e^* > 1 - \delta$, it must be the case that $R_e^* > 0$ (otherwise, if $R_e^* = 0$, the empire can increase its net tax revenue by decreasing G_e and M_e simultaneously). Therefore, it suffices to compare $R_e^* = 0$ and $R_e^* > 0$ when $G_e = 1 - \delta$. When $G_e = 1 - \delta$, the Lagrangian optimization problem is

$$\max_{M_e, R_e} \mathcal{L}_e = y - c\left(M_e\right) - R_e + \phi R_e + \gamma\left(m\left(M_e, 0, \beta\right) + R_e - \lambda\left(\Lambda, 1-\delta\right)\right)$$

where ϕ and γ are the Lagrangian multipliers. The first order conditions are given by

$$\begin{aligned}
M_e :& \quad -c_M + \gamma m_1\left(M_e, 0, \beta\right) = 0, \\
R_e :& \quad -1 + \phi + \gamma = 0, \\
\phi :& \quad \phi R_e \geq 0 \text{ and either } \phi = 0 \text{ or } R_e = 0, \\
\gamma :& \quad \lambda\left(\Lambda, 1-\delta\right) = m\left(M_e, 0, \beta\right) + R_e.
\end{aligned}$$

Suppose $R_e^* = 0$. Then,

$$\begin{aligned} c_M\left(M_e^*\right) &= \left(1-\phi^*\right) m_1\left(M_e^*, 0, \beta\right), \text{ and} \\ \lambda\left(\Lambda, 1-\delta\right) &= m\left(M_e^*, 0, \beta\right). \end{aligned}$$

Note that $dM_e^*/d\beta = 0$ since $dm\left(M_e^*, 0, \beta\right)/d\beta = 0$.

Alternatively, suppose $R_e^{**} > 0$. This implies $\phi^{**} = 0$, and

$$\begin{aligned} c_M\left(M_e^{**}\right) &= m_1\left(M_e^{**}, 0, \beta\right), \text{ and} \\ \lambda\left(\Lambda, 1-\delta\right) &= m\left(M_e^{**}, 0, \beta\right) + R_e^{**}. \end{aligned}$$

Note that $dM_e^{**}/d\beta = 0$ and $dR_e^{**}/d\beta = 0$ since $dm_1\left(M_e^{**}, 0, \beta\right)/d\beta = 0$. Furthermore,

$$\begin{aligned} m\left(M_e^*, 0, \beta\right) &= m\left(M_e^{**}, 0, \beta\right) + R_e^{**} \\ &> m\left(M_e^{**}, 0, \beta\right). \end{aligned}$$

This implies that $M_e^* > M_e^{**}$. Now, let

$$\begin{aligned} \Psi &\equiv c\left(M_e^*\right) - c\left(M_e^{**}\right) - R_e^{**} \\ &= c\left(M_e^*\right) - c\left(M_e^{**}\right) - \left(m\left(M_e^*, 0, \beta\right) - m\left(M_e^{**}, 0, \beta\right)\right) \\ &= \left(c\left(M_e^*\right) - m\left(M_e^*, 0, \beta\right)\right) - \left(c\left(M_e^{**}\right) - m\left(M_e^{**}, 0, \beta\right)\right) \end{aligned}$$

The empire should set $R_e > 0$ only if $\Psi > 0$. Since $M_e^* > M_e^{**}$, $\Psi > 0$ if for all $M \geq 0$,

$$\frac{\partial c\left(M\right)}{\partial M} > \frac{\partial m\left(M, 0, \beta\right)}{\partial M}.$$

$\square$

A.4 Multiple Military Bases

In Section 3, we assume that every regime can only set up one military base, which we refer to as the capital city. While this assumption clearly does not apply to a rich present-day democracy, it aptly reflects the nature of power in an age of "natural states," when political power was inherently personal and determined by one's access to the use of organized violence (North et al., 2009). Because *political power grows out of the barrel of a gun*, historically, a spatially extensive empire that maintained two or more comparable armies commanded by different individuals to safeguard different parts of the empire either behaved like multiple states (e.g., the Holy Roman empire) or would fragment into multiple states (e.g., the Carolingian empire, the Mongol empire). Political decentralization was inevitable because the alternative centers of power enjoyed vast autonomy that was guaranteed by the long distance between them and the national capital city.

Below, we discuss historical examples drawn from China and Rome to illustrate this point further (A.4.2). Before that, we first relax the assumption and allow each regime to set up multiple auxiliary armies under the constraint that the central army has to be measuredly stronger than the regional armies to deter insubordination or outright usurpation (A.4.1). We show that relaxing the "one military per regime" assumption in this manner would not affect our results and conclusions.

A.4.1 Alternative Assumption

Historically, one important constraint that empires faced in setting up auxiliary armies was the risk of military usurpation. Once empowered, commanders of the auxiliary armies often sought to break away or to replace the incumbent ruler (A.4.2).

Suppose that regime $i \in \{e, l, r\}$ may set up any number of auxiliary military bases outside its capital city. To model agency cost in the form of military usurpation, we assume that to prevent a military usurpation, the strength of regime i's central army (based in the capital city) must be no less than the strength of the individual auxiliary armies.

Formally, preventing military usurpation requires $m\left(M_i, \left|G_i - G_k\right|, \beta\right) \geq m\left(M_k, 0, \beta\right)$, where G_k and M_k are the location and military investment of auxiliary military base k. If this constraint is violated, a military usurpation occurs, and the regime receives a negative payoff.

For simplicity, assume that it costs $c(M_i)$, where $c(0) = 0$, $c' > 0$, and $c'' > 0$, for regime i to set up its central army as before, but after that the auxiliary armies can be set up at zero cost.[38]

As Figure 18 illustrates, allowing regime i to set up multiple auxiliary armies has the effect of increasing its ability to project military power from $m(M_i, t, \beta)$ to $\hat{m}(M_i, t, \beta)$, where $\hat{m}(M_i, t, \beta)$ is

[38]For instance, suppose armies outside the capital city are organized to farm so as to feed themselves. This practice is known as *tuntian* in Chinese history.

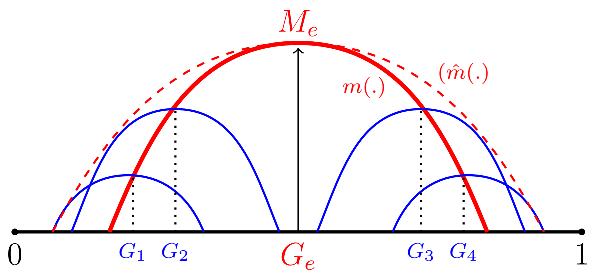


Figure 18: The empire sets up a central army and four auxiliary armies. Its ability to project military power is depicted by the dashed red line, as opposed to the thick red line, which depicts the military power of an empire with no auxiliary armies.

a function that envelops the $m(M_i, t, \beta)$ functions of regime i's central and auxiliary armies, i.e., $\hat{m}(M_i, t, \beta) = \max_{G_k \in [0,t]} m(m(M_i, |G_i - G_k|, \beta), |G_k - t|, \beta)$.

Hence, if we relax the "one military per regime" assumption in the model with this alternative setup, all that is required is to replace $m(M_i, t, \beta)$ [i.e., $M_i - \beta t^2$ in Section 3] with $\hat{m}(M_i, t, \beta)$. The analysis and all results remain qualitatively unchanged.

A.4.2 Historical Discussion: China and Rome

Interestingly, the lessons of Chinese and Roman history also provide ample evidence to support our argument that it is inherently unstable for a state to maintain two or more comparable political-military centers. During the mid-Tang dynasty, the Xuanzong emperor (r. 712–756) implemented a polycentric political-military system and devolved much of the central government's political authority to frontier military governors with the goal of improving military responsiveness and effectiveness. However, Xuanzong's favorite and most powerful frontier governor, An Lushan, infamously revolted in 755 as the military might of An's army fed popular suspicion of his political ambitions, which ironically compelled An to revolt. A similar development took place during the early Ming dynasty with the implementation of a de facto twin-capital system in which the emperor resided in the populous south and his uncle, the Prince of Yan, coordinated border defense in the strategic north. Mutual suspicion again led to the outbreak of a bloody civil war in 1399 with the Prince of Yan emerging as the eventual victor. To prevent history from repeating itself, the usurper moved the capital city from Nanjing ('southern capital') to Beijing ('northern capital') in order to maintain direct control of the large army along the northern border.

Similarly, the history of the Roman Empire was beset with the problems posed by maintaining multiple large armies based far away from the capital. The turmoil of the late Republic was partly due to the ability of powerful senators such as Marius, Sulla, Caesar, Mark Antony, and Octavian to build up powerful field armies that they could then use to turn on their enemies and in some cases on the Republic itself.

During the Principate (27 BC to AD 180), Roman emperors were usually sufficiently secure to entrust independent commands to generals in charge of large field armies based far from the capital. However, even in this period this was a source of political tension as indicated by the careers of Germanicus (15 BC–AD 19), who died under suspicious circumstances, and Corbulo (7–67), who was ordered to commit suicide. More penitently, it was the commanders of the largest field armies who rebelled in 69 and 193 that led to devastating civil wars in both cases. After the death of Severus Alexander (r. 222–235), the empire nearly fell apart due to a series of rebellions, civil wars, and external invasions.

As a result, in the later Roman Empire, emperors rarely trusted subordinates with large field armies. From Gallenius (r. 260–268) onwards, mobile armies known as *comitatensis* were set up to replace the large frontier armies and accompany the emperor to war. This innovation reduced the problem of rebellion by successful field commanders. However, a single emperor could not campaign on more than one front at a time. Diocletian, therefore, inaugurated a fourfold division of the empire known as the Tetrarchy, under which four emperors would jointly rule the empire from four separate capitals. This system, however, proved unstable; it did not last long beyond his retirement in 306 and its collapse led to a series of civil wars that only finally ended with Constantine's reunification of the empire in 324. Civil wars reemerged during periods of imperial division in 337–350, 360–361, 383–388, and 392–395. By the end of the fourth century, the centrifugal forces affecting the empire had led to the *de facto* permanent division of the empire into East and West, and from this point on, the two empires coexisted as separate political entities until the fall of the Western Empire in 476.

A.5 Historical Discussion: The Byzantine Empire

In this subsection we briefly review Byzantine history in light of our theory. After the collapse of Rome, the Eastern Roman empire survived. At least until the Arab invasions of the mid-seventh century, it remained a large empire on the scale of Han China or the Ottoman Empire. Following the historical scholarship, we refer to the East Roman empire as the Byzantine empire for the period after AD 600.

As noted in the main text, the Byzantine empire faced a two-sided threat: in the North it faced attacks from the Huns, Avars, Bulgars, Slavs, Magyars, Pechenegs, and Cumans; from the East it was threatened by Persians and later the Arabs. The existence of such a two-sided threat might seem to imply that Byzantine as a large empire should have collapsed. Does the fact that it endured for so long provide a counterexample to our analysis?

To address this question, it is important to note that it was only during the sixth and seventh centuries that the Eastern Roman empire was a truly large-scale empire. Furthermore, even during this period, the Eastern Roman empire were rarely successful in maintaining both their northern and their eastern borders for long periods of time. One reason for its ability to survive intact during the fifth century is the fact that its eastern frontier remained largely peaceful, as the Persian empire also faced invasion threats from the steppe. In Whittow's words:

> During the fifth century both empires had been preoccupied by other enemies to their west and east respectively [...] the Romans were struggling with the direct and indirect consequences of the Hun invasions; the Persians waging war with first the Kidarites and then the Hepthalites (Whittow, 1996, 41).

The precondition for the reconquest of western territories in the mid-sixth century, therefore, was peace in the east.[39] Similarly, it was a two-front war in the early seventh century that nearly saw the demise of the entire empire, culminating in the siege of Constantinople by Avars and Persians in 626. After the seventh century, the remaining rump of the empire that survived the Arab invasions was a considerably smaller state than the empire of Justinian had been. It still faced a two-sided threat. It attempted to defend against these threats in two ways. First, there was the institutional development of *themes* that, like the feudal system in western Europe, represented a decentralization of political and military authority, and which had the advantage of providing a measure of defense in depth (Treadgold, 1997). Second, as Luttwak (2009) documents, the Byzantines became exceptionally adept at using bribes, diplomacy, and espionage to play their enemies off one another. Together these allowed the Byzantine empire to survive for centuries, but as a medium-sized state, instead of as a hegemonic empire.

[39] Whittow (1996, 48) argues that 'Justinian's western offensive had only been possible during periods of peace with Persia, and once war restarted the Balkans, Italy and Africa had to make do with the limited forces that were left.'

A.6 Further Empirical Evidence

Table 8: Correlation Matrices of the Variables Employed in Section 4.1

	y_t	y_t^{alt}	x_t	z_{1t}	z_{2t}	w_{1t}	w_{2t}	w_{3t}	w_{4t}	w_{5t}	w_{6t}	w_{7t}	w_{8t}
y_t	1.00												
y_t^{alt}	-0.87***	1.00											
x_t	-0.08	0.04	1.00										
z_{1t}	-0.03	-0.01	0.15*	1.00									
z_{2t}	0.07	-0.03	0.02	0.47***	1.00								
w_{1t}	-0.23***	0.16*	0.04	0.37***	0.10	1.00							
w_{2t}	-0.02	-0.05	-0.07	0.56***	0.24***	0.40***	1.00						
w_{3t}	0.06	0.02	-0.01	-0.20**	-0.37***	0.02	-0.14*	1.00					
w_{4t}	-0.46***	0.45***	-0.12	-0.07	-0.23**	0.02	-0.01	-0.09	1.00				
w_{5t}	-0.22**	0.12	0.16*	0.29***	0.08	0.40***	0.23**	-0.03	-0.14*	1.00			
w_{6t}	0.22**	-0.20**	-0.19**	0.49***	0.22**	0.18*	0.50***	-0.26***	-0.13	-0.12	1.00		
w_{7t}	0.05	-0.08	0.14*	0.72***	0.63***	0.29***	0.47***	-0.20**	-0.23***	0.27***	0.51***	1.00	
w_{8t}	0.10	-0.12	0.04	-0.13	-0.22**	-0.01	0.00	0.06	0.29***	0.06	-0.03	-0.28***	1.00

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Below, we test the robustness of the results presented in Section 4 by replacing *Fragmentation* with *#Regime* (the *log* number of regimes in China Proper in decade t. See Table 1) as the dependent variable. The results obtained are consistent with those discussed in in Section 4.

Tables 9 and 10 are analogous to Tables 2 and 3 in Section 4. For example, Column (a) of Table 9 suggests that, while nomadic attacks appear to have negligible effect on the number of regimes in China proper, the lagged effect is significant: every additional nomadic attack is associated with a 1.37 percent decrease in the number of regimes in China proper one decade later. In the VAR specification, the corresponding estimate is -1.30 percent.

Table 9: ADL Model

	(a)	(b)
Dependent variable:	#Regimes	#Regimes
#Regimes: Lag 1	1.061***	1.019***
	(0.0664)	(0.0678)
#Regimes: Lag 2	-0.337***	-0.325***
	(0.0901)	(0.0901)
#Regimes: Lag 3	0.146**	0.113*
	(0.0635)	(0.0646)
Nomadic attacks	0.00379	0.00409
	(0.00614)	(0.00642)
Nomadic attacks: Lag 1	-0.0137**	-0.0126**
	(0.00614)	(0.00634)
Additional controls	No	Yes
Observations	203	203
R-squared	0.781	0.798
AIC	0.158	0.184

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 10: VAR Model

	(a)	(b)
Dependent variable:	#Regimes	Nomadic attacks
#Regimes: Lag 1	1.026***	1.904***
	(0.0660)	(0.721)
#Regimes: Lag 2	-0.343***	-1.330
	(0.0882)	(0.963)
#Regimes: Lag 3	0.131**	-0.254
	(0.0638)	(0.696)
Nomadic attacks: Lag 1	-0.0130**	0.333***
	(0.00626)	(0.0684)
Nomadic attacks: Lag 2	0.000943	0.236***
	(0.00654)	(0.0715)
Nomadic attacks: Lag 3	-0.00269	0.000803
	(0.00632)	(0.0690)
Additional controls	Yes	Yes
Observations	203	203

A.7 Population Fluctuations

Pop. ('000)	t	t^2	t^3	t^4	t^5	t^6	N	Adj. R^2
China	69.4***	-	-	-	-	-	14	0.58
Europe	54.2***	-	-	-	-	-	14	0.67
China	-77.8	0.081***	-	-	-	-	14	0.78
Europe	-68.8***	0.068***	-	-	-	-	14	0.94
China	44.2	-0.092	$6.4 \cdot 10^{-5}$	-	-	-	14	0.79
Europe	-0.62	-0.029	$3.6 \cdot 10^{-5}$*	-	-	-	14	0.95
China	-144.9	0.44	$-4.1 \cdot 10^{-4}$	$1.3 \cdot 10^{-7}$	-	-	14	0.80
Europe	-12.0	0.026	$7.1 \cdot 10^{-6}$	$8.1 \cdot 10^{-9}$	-	-	14	0.95
China	300.5	-1.61	0.0027	$-1.9 \cdot 10^{-6}$*	$4.6 \cdot 10^{-10}$*	-	14	0.86
Europe	144.7	-0.72	0.0011	$-7.1 \cdot 10^{-7}$	$1.6 \cdot 10^{-10}$*	-	14	0.96
China	242.3	-1.21	0.0018	$-8.9 \cdot 10^{-7}$	$-5.0 \cdot 10^{-11}$	$9.6 \cdot 10^{-14}$	14	0.84
Europe	34.4	0.036	$-6.7 \cdot 10^{-4}$	$1.2 \cdot 10^{-6}$	$-8.1 \cdot 10^{-10}$	$1.8 \cdot 10^{-13}$	14	0.96

Constant terms are not reported. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 11: Fitting Year Polynomials to Chinese and European Population Data. Adjusted R^2 is higher for Europe than for China in each case. See discussion in Section 5.3.