A Theory of Power Structure and Institutional Compatibility: China vs. Europe Revisited*

Ruixue Jia† Gérard Roland‡ Yang Xie§

August 27, 2021

Abstract

The literature on institutions and development has often contrasted the more inclusive, open-access, and equal societies with other less inclusive ones, but the scholarship comparing Imperial China and Premodern Europe has emphasized different inclusive features of institutions in both societies, such as the non-hereditary access to elite status and the less oppressed peasantry in China, and the stronger rule of law and property rights in Europe. To reconcile these views, we provide a framework based on observed institutional differences along two dimensions of the power structure of society: the degree of the Ruler's absolute power, which was lower in Europe, and the level of asymmetry between the Elites and the People in terms of their power and rights, which was lower in China. Our model shows that a more balanced Elite–People relationship can be compatible with a more absolutist Ruler. This is because the stronger the absolute power, the more effectively such balance will stabilize an autocratic rule, and the greater the Ruler's incentive to promote such balance. Discussion and stylized facts support the theory's relevance. Our theory helps understand the relationship between major components of inclusive institutions and the logic behind autocratic stability.

Keywords: absolute power, Elite–People relationship, autocratic stability, inclusive institution, rule of law, access to elite status

*We are grateful to Avner Greif and Joel Mokyr for their thorough and constructive feedback. We have also benefited from thoughtful comments from Chris Bidner, Gary Cox, Wei Cui, Georgy Egorov, Patrick Francois, Murat Iyigun, Mark Koyama, Ling Li, Weijia Li, Zhao Liu, Zhaotian Luo, Pietro Ortoleva, Albert Park, Jean-Laurent Rosenthal, Jared Rubin, Larry Samuelson, Tuan-Hwee Sng, Konstantin Sonin, David Stasavage, Guido Tabellini, Chenggang Xu, and Xueguang Zhou. We also thank participants in online seminars hosted by Fudan, Harvard, Tsinghua, UBC, UCSD, and USC, the economic history reading group at Monash, and the 2021 ASSA, CESI, and SIOE meetings for their valuable comments. We thank Ming Zhang for his excellent research assistance. We appreciate the support from the EDI grant "Endogenous Political Fragmentation: The Role of Property Rights in Historical Perspective" and the valuable feedback from EDI. An earlier version of the paper has been circulated as NBER Working Paper 28403 and CEPR Discussion Paper 15700.

†UCSD and LSE, CEPR, and NBER; rxjia@ucsd.edu.

‡UC Berkeley, CEPR, and NBER; groland@econ.berkeley.edu.

§UC Riverside; yang.xie@ucr.edu.

1 Introduction

The very influential literature on institutions and development has often cast societies in a light that contrasts the more inclusive, open-access, and equal ones with the more extractive, limited-access, and unequal ones (e.g., North, 1989; North and Weingast, 1989; Acemoğlu, Johnson and Robinson, 2001, 2005a,b; North, Wallis and Weingast, 2009; Acemoğlu and Robinson, 2012; Cox, North and Weingast, 2019). At the same time, the scholarship comparing Imperial China and Premodern Europe seems to contradict this dichotomic classification. It has been emphasized that, although rule of law and protection of property rights against the Ruler were weaker in China, some other institutional features of China were more inclusive compared to Europe: for example, the access to elite status was primarily non-hereditary and governed by the civil service exam, peasants enjoyed a greater degree of freedom, and land ownership was less concentrated (e.g., Finer, 1997a,b; Fukuyama, 2011; Tackett, 2014; Zhang, 2017; Acemoğlu and Robinson, 2019; Stasavage, 2020; Greif, Mokyr and Tabellini, Forthcoming; the survey by Qian and Sng, 2021). Are these views of institutional differences contradictory to each other, and can they be reconciled?

In this paper, we reconcile these views by providing a framework that analyzes institutional differences along two, instead of one, dimensions of the *power structure* of society, which is about how power and rights were allocated across the Ruler, the Elites, which included primarily the lords in Europe and bureaucrats in China, and the common People. First, in Europe, the power and rights of the Elites and People were less conditional on the Ruler's will, i.e., the *absolute power of the Ruler* was weaker. This was reflected in, for example, the different strengths of rule of law, property rights, and whether the king or emperor had ultimate ownership and control over land and population. Second, in China, the *Elite–People relationship in terms of their power and rights* was less asymmetric. This was indicated, for example, by how the access to elite status was governed, how much freedom the peasantry enjoyed, and how unequal land ownership was. The characterized power structure differences were the most prominent during the 9–14th centuries, with persistence beyond, between the society in the historical core of Imperial China and the western–central European society where feudalism was once influential.

After examining rich comparative historical narratives based on our power structure framework, we formalize this framework in a simple game-theoretical model of the relationship between the two dimensions of the power structure. We start with a Ruler, who prefers to maintain a particular status quo of autocratic rule, and a Challenger, who could try to alter it. Since the Challenger can be either a foreign threat not under the Ruler's rule, a conspiring elite, or a rebellious population under the Ruler's rule, since the Challenger's goal

does not necessarily involve dethroning the Ruler, and since the challenge can be armed or nonviolent, our model is sufficiently general to cover a wide range of threats that could destabilize an autocratic rule. In the model, we assume that the success of a potential challenge in altering the status quo depends on whether the Elites and People choose to side with the Ruler. In the model, more symmetric power and rights between Elites and People is represented by less unequal payoffs, if they have not defied the Ruler; we model a stronger absolute power of the Ruler as a greater proportional reduction in the payoffs of the ruled, i.e., a heavier punishment, after they unsuccessfully defy the Ruler.

Analysis of the model leads to a comparative institutional theory on the compatibility between the two dimensions of the power structure. In the analysis, we first take the level of the Ruler's absolute power as exogenous and analyze how it affects the stability of autocratic rule and the Ruler's perspective about the Elite–People relationship. Everything starts from reading the absolute power of the Ruler as the *conditionality* of power and rights of the ruled on the Ruler's will. Given any non-zero level of such conditionality, the more power and rights the People enjoy when they have not defied the Ruler, the more they will lose if they defy the Ruler, and, therefore, the more they will be willing to side with the Ruler during a challenge. We call this the *punishment* effect of more power and rights of the People. Knowing that a stronger alliance between the Ruler and People has worsened the prospect of a challenge to the Ruler, the Elites will be more willing to side with the Ruler, too. We call this the *political alliance* effect.[1] The Challenger would then be deterred from challenging the status quo, stabilizing the autocratic rule and thus creating an incentive for the Ruler to promote a more symmetric Elite–People relationship. Since a stronger absolute power of the Ruler implies a greater aforementioned conditionality, it will make the initial punishment effect and, therefore, the total stabilizing effect stronger. The Ruler's incentive to promote a more symmetric Elite–People relationship will thus be greater when the Ruler has a stronger absolute power. Having a more absolutist Ruler can thus be compatible with a more symmetric Elite–People relationship, reconciling the seemingly contradictory views in the literature on the institutional differences between Imperial China and Premodern Europe.

As we show in Online Appendix E, the insights and results from the theory are robust in a Markov game in which the ruled covet the Ruler's throne and all players take continuation values into consideration. We also show in Online Appendix C that the compatibility result also holds in the other direction, i.e., the stabilizing effect of a stronger absolute power of

[1] As remarked by Orwell (1947, p. 17), this idea of the Ruler and the People "being in a sort of alliance against the upper classes" is "almost as old as history" in Europe; in China the same idea can be traced to not later than *Han Feizi* (Watson, 1964, p. 87) from the 3rd century BC, which has been the most representative text in the Chinese Legalist tradition since then.

the Ruler is increasing in the level of symmetry between the Elites and People.

A few additional implications arise about the power structure in the long run when we extend the model by endogenizing the level of the Ruler's absolute power. For example, because a more absolutist Ruler can take away more of the power and rights of the People once they unsuccessfully defy him, he is more willing to grant more power and rights to them in the first place. Therefore, the People may prefer the Ruler to enjoy a stronger absolute power, defying less often and enjoying their granted power and rights under a more stable autocratic rule. This makes it possible for the power structure of a strong absolute power and a relatively symmetric Elite–People relationship to be incentive-compatible for both the Ruler and the People, and therefore to persist. In another extension, we allow the current political stability to influence the future power structure. We show that two societies that differ slightly in their power structure or autocratic stability could diverge into different steady states – one with a stronger absolute power of the Ruler, a more symmetric Elite–People relationship, and a higher stability of autocratic rule than the other.

Given these results, we further explore the historical relevance of our theory. We first discuss how our theory can help understand specific institutions. For example, we can interpret the use of the civil service exam to govern the access to the bureaucracy in China and the rise of cities in Europe both as resulting from the Ruler's efforts to reduce the Elite–People asymmetry.[2] Second, we examine the auxiliary predictions from our model about the impact of the power structure on the stability of autocratic rule. We systematically compare Imperial China and Premodern Europe in the frequency of wars, the risk of deposition to a Ruler in a given year, and the resilience of unified autocratic rule. Consistent with the predictions of our model, the data show that autocratic rule was more stable in China than in Europe over the 9–14th centuries, when the differences in the power structure were the most prominent, with persistence in later centuries.

Our paper contributes to the political economy literature on institutions and development by investigating the relationship between major components of inclusive institutions (e.g., the aforementioned; Besley and Persson, 2011, 2014; Mokyr, 2016; Acemoğlu and Robinson, 2019). The literature often analyzes society by categorizing it into two estates (e.g., state vs. society, elites vs. mass, those with vs. those without access to political and economic resources and decisions), and it has taught a general lesson that a strong interdependence and synergy exists between institutional arrangements that are conducive to sustainable political, economic, and social development, such as rule of law and property rights on the one hand and a more open access to elite status on the other hand (e.g.,

[2]To be sure, the development of cities in Europe was to a large extent based on autonomous factors and exogenous shocks, but as we show below, various kings acted to help it.

North, Wallis and Weingast, 2009; Besley and Persson, 2011, 2014; Acemoğlu and Robinson, 2012). We extend the two-estate framework into a three-estate one. By that, we show that the more repressive an institution is in the dimension of the Ruler's absolute power, the more inclusive it may be in the dimension of the power and rights equality between the Elites and People, and this pattern may well persist. This seemingly paradoxical result is, to our knowledge, new to the literature.

Our paper also contributes to the literature on the strategies that a ruling class can use to fend off challenges to their rule. For example, democratization and enfranchisement as in Acemoğlu and Robinson (2000, 2001) can serve as a credible commitment to redistribution by *shifting the decision power* to the median voter, avoiding destructive revolution; Acemoğlu, Verdier and Robinson (2004) propose that an *ad-hoc* divide-and-rule *policy* given weak institutional constraints can intensify the collective action problem among the ruled, making the ruler safer; Padró i Miquel (2007) argues that the fear by the ruled of falling under an even worse ruler *external to* the incumbent ruler can help discipline the ruled, achieving stability for the incumbent ruler. Our analysis suggests that an absolutist Ruler can co-opt the People and thus secure his autocratic rule by making the People's power and rights more comparable to the Elites', for example, by promoting meritocratization. Compared with the literature, strategies of this type are unique in the sense that they *do not change* where the absolute power lies, provide an *ex-ante committed* payoff schedule through the *institutional design* of the power structure, and the Ruler's incentive to engage in such strategies depends on the absolute power of the Ruler in *the same* power structure.

Conceptually the closest to us in this thread of literature, Persico (2021) shows in a general model for political regimes that as long as civil liberties are imperfectly protected, which is similar to our notion that the power and rights of the ruled are conditional on the Ruler's will, a politician will always have an incentive to promise equal treatment across citizens, trying to win their coordinated support, which is similar to the more symmetric Elite–People relationship in our context with the political alliance effect involved. Concurrently and independently developed, our paper and Persico (2021)'s paper complement each other: Persico (2021) focuses on *policy* treatment and *provision of public goods,* whereas we focus on the compatibility within the *power structure* and its implications for *political stability.*

It has been well documented in the literature that the unified autocratic rule of a dominant state could hardly be maintained in Europe since the fall of the Roman Empire, while in Imperial China it was relatively resilient (e.g., Finer, 1997a,b; Scheidel, 2019). Studies have argued that the political divergence has been highly consequential for economic development (e.g., Rosenthal and Wong, 2011; Mokyr, 2016). A few explanations of the political divergence have been proposed from the environmental, geographical, or geopolitical perspective

(e.g., Wittfogel, 1957; Jones, 1981; Turchin, 2009; Dincecco and Wang, 2018; Scheidel, 2019; the survey by Qian and Sng, 2021). Among recent contributions, Ko, Koyama and Sng (2018) and Fernández-Villaverde, Koyama, Lin and Sng (2020) show that the spatial distribution of external threats and the existence or not of a high-productivity, traversable core geographical region could have played important roles in facilitating the unification of China and fragmentation of Europe.

We contribute to the literature on the same topic but from the institutional front. First, we introduce the power structure approach to the literature, and we show that the characterized power structure differences between the two societies can explain the differences in their autocratic stability. The same approach can also shed light on some important variations and changes within China and Europe and on other parts of the world, as we briefly discuss in Sections 2.1 and 5.

Second, Acemoğlu and Robinson (2019) have emphasized the role of the state–society struggle in understanding the political divergence. We complement their view by a richer strategic dynamics: a political alliance could exist between the head of the state, i.e., the Ruler, and the lower classes of society, i.e., the People, together against the state apparatus members or higher classes of society, i.e., the Elites, especially when the Ruler's absolute power is strong. We can thus provide a more general understanding of a few specific institutional arrangements, on which the literature has focused to explain the political divergence, such as the wage of the tax-collecting agents, the tax on the masses, and the development of fiscal capacity (Gennaioli and Voth, 2015; Ma and Rubin, 2019), the availability and capacity of bureaucracy (Stasavage, 2020), and meritocracy as an informational solution to the institution–power friction (Huang and Yang, 2021). That is to say, we can read the low wage–low tax equilibrium and bureaucracy with meritocratic recruitment in Imperial China as a more symmetric Elite–People relationship and thus a stronger Ruler–People alliance; these were compatible with the stronger absolute power of the Ruler and higher autocratic stability in China.

Finally, some studies have explored, along the technological, geographical, and economic lines, the more exogenous factors behind the initial differences in the power structure (e.g., McNeill, 1982; Roland, 2020; Stasavage, 2020). Although not focusing on the origin of these initial differences, we show that such differences can be incentive-compatible and exhibit self-reinforcing dynamics over time, together with the persistent difference in the autocratic stability. This dual divergence of the power structure and autocratic stability complements the divergence of culture and its co-evolution with political institutions (Greif and Tabellini, 2010, 2017; Greif, Mokyr and Tabellini, Forthcoming).

The paper is organized as follows. Section 2 briefly presents historical narratives in

the power structure framework on the institutional differences between Imperial China and Premodern Europe. Section 3 presents the settings, analysis, and extensions of the model. Section 4 explores the historical relevance of the theory with further discussion and stylized facts. Section 5 concludes the paper.

2 Power Structure in Historical Narratives

2.1 Scope and Focus of Historical Narratives

In this section, we discuss historical narratives on the institutional differences between Imperial China and Premodern Europe along the two dimensions of the power structure of society. In the narratives, by "China," we consider the society in "the historical core of Imperial China," i.e., "the traditionally agrarian part of China south of the Great Wall and east of the Tibetan Plateau" (Fernández-Villaverde, Koyama, Lin and Sng, 2020, p. 8, 12). By "Europe," unless clarified otherwise, we follow the focus of Bloch (1962a,b), Finer (1997b, Book III, Part III), and Blaydes and Chaney (2013), i.e., the Romano–Germanic influenced or assimilated society in western and central Europe where feudalism was once influential. This society was "[h]emmed in by these three blocs, Mohammedan, Byzantine, and Slav" (Bloch, 1962a, p. xxvi) and "comprised principally the British Isles, the Scandinavian countries, France, Germany, Italy, and northern Spain" (Finer, 1997b, p. 855).[3]

The most relevant period of the characterized power structure differences was the 9–14th centuries, with persistence beyond. This period covered the rise and decline of feudalism in Europe (e.g., Bloch, 1962a,b; Ganshof, 1952), with the Black Death taking place in the middle of the 14th century; in Imperial China, it was since the Tang dynasty (618–907) that political institutions had largely been stable, after the swings during the 800 preceding years (e.g., Yan, 2009).

Admittedly, important variations and changes in the power structure existed across polities and over time within China and within Europe. At the same time, "over and above" these variations and changes, historians have emphasized "the predominant quality of a common civilization" in Europe and the "evolving axis" or "theme" of the institutional and cultural characteristics of Chinese society during our focused period (e.g., Bloch, 1962a, p. xxvi; Yan, 2009, p. 11–12). We follow this insight in our narratives: we try to identify the "ideal type" of the differences in the power structures between Imperial China and Premod-

[3]Following Blaydes and Chaney (2013), "Europe" by this definition covered all countries under the section "The Barbarian West" and the subsections "The British Isles," "France," "The Low Countries," "Italy," "The Iberian Peninsula," "The German-speaking States," "Scandinavia," and "Crusader States" under the section "Europe" in Morby (1989).

ern Europe, sometimes discussing specific polities or periods as examples; we also adopt the *longue durée* approach by focusing on significant, persistent features of the power structures.

By no means denying the aforementioned important variations and changes within China and Europe, our narratives can provide a benchmark to help interpret them in the power structure framework. For example, within Europe, consistent with Proposition 3 below, Bloch (1962a, p. 180) observes the co-existence of a stronger absolute power of the king over his vassals and less significant class distinctions between the lords and the peasantry in Germany from the 10th century to the end of the Middle Ages, compared with the power structure in France at that time; within China, we discuss in Online Appendix G how our framework can help interpret the dynastic cycles in Chinese history, each of which was marked by a co-decline of the absolute power of the Ruler, the Elite–People symmetry, and the stability and effectiveness of autocratic rule (e.g., Skinner, 1985; Usher, 1989; Dillon, 1998).

We summarize the historical narratives in Table 1, and we elaborate on them below.

2.2 Absolute Power of the Ruler

The first difference we emphasize in the power structure between Imperial China and Premodern Europe is that Chinese Rulers enjoyed a stronger absolute power than their European counterparts, by which we mean that the power and rights of the ruled were more dependent on the Ruler's will in China compared to Europe. This difference is first reflected in the strength of rule of law, and then in the ultimate ownership and control over the most important assets in historical societies: land and population. Next, we summarize the narratives and explain how we model the degree of absolute power based on these narratives.

Strength of rule of law. As noted by many scholars, Chinese emperors were less constrained by rule of law (Finer, 1997a,b; Stasavage, 2016; Acemoğlu and Robinson, 2019; Ma and Rubin, 2019, p. 227; Greif, Mokyr and Tabellini, Forthcoming).[4] As put by Finer (1997a,b, p. 455, 836), all the ruled, including the top bureaucrats, were "subjects not citizens" and had only "duties not rights"; as observed by Fukuyama (2011, p. 290) and

[4]The Chinese Ruler had the obligation to act benevolently towards the ruled and to follow the "Mandate of Heaven" (e.g., Zhao, 2009), but as noted by Stasavage (2016, p. 148), "the concept of a Mandate of Heaven never extended to obtaining consent, nor did it involve assembling representatives to achieve this goal." Finer (1997a, p. 462) also notes: "[i]deally, government must be of the people, for the people: but, emphatically, Mencius never for a moment hints that it can ever be by the people. Very much the reverse. ... Nor did a dissatisfied populace have the right to rebel." Perry (2008) further contrasts the right to rebel in the Anglo-American tradition, which is against tyranny and for liberty, i.e., about the Ruler's absolute power and γ in our model, with the People's rights in the Chinese tradition, which is for livelihood and against poverty and socioeconomic injustice, i.e., mainly about the Elite–People balance and β in our model.

Table 1: Power Structure in Imperial China and Premodern Europe

	China	Europe	Examples of references
Absolute power of the Ruler			
Strength of rule of law	Ruler less constrained by law	Ruler constrained by Church and law	Bloch (1962b), Unger (1977), Mann (1986) Finer (1997a,b), Tamanaha (2004) Fukuyama (2011), Acemoğlu and Robinson (2019) Greif, Mokyr and Tabellini (Forthcoming)
Ultimate ownership of land	Reserved for Ruler; confiscation legitimate when Ruler deemed it necessary	Confiscation highly constrained; Ruler expected to "live of his own"	Chao and Chen (1982), Levi (1988) Finer (1997b), Wang (2000)
Ruler's control over population	Ruled considered Ruler's subjects; harsh penalty for disloyalty	Limited control; much less harsh punishment for disloyalty	Bloch (1962a), Lander (1961), Levenson (1965) Mann (1986), Finer (1997a,b)
Asymmetry in power and rights between Elites and People			
General comparison	Much less unbalanced	Elites a supreme class; oppressive to the poor	Bloch (1962b), Lü (1944), Finer (1997b)
Hereditary vs. non-hereditary access to elite status	Non-hereditary, elite status governed through the civil service exam	Hereditary nobility	Kemp (1970), Finer (1997b), Wickham (2009) Yan (2009), Parish (2010), Tackett (2014)
Inequality in land ownership	Mostly free and landowning peasantry; land ownership less concentrated	Serfdom common in Middle Ages; land ownership much more concentrated	Esherick (1981), Chao and Chen (1982) Beckett (1984), Finer (1997a), Wickham (2009) Tackett (2014), von Glahn (2016), Zhang (2017)
Inheritance rule	Partible inheritance	Primogeniture increasingly more common	Goody, Thirsk and Thompson (1976) Goldstone (1991), Bertocchi (2006) von Glahn (2016)

Unger (1977, p. 104), "law was only the positive law that [the emperor] himself made" and it "could be as general or as particular as the policy objectives of the rulers might require."[5]

In contrast, European Rulers faced strong constraints from the Christian church (Mann, 1986; Fukuyama, 2011; Johnson and Koyama, 2019; Scheidel, 2019; Greif, Mokyr and Tabellini, Forthcoming). Given the Pope's threat to delegitimize and excommunicate them, "[k]ings ... could not defy the Pope for very long," as shown in many examples (Southern, 1970, p. 130).[6] The king also faced much tighter legal constraints. In the famous words of Bracton (1968, vol. 2, p. 33), "[t]he king must ... be under the law, because law makes the king." Having emerged from the 9th-century customary law, a man's right to judge and resist when his king had acted unlawfully had been repeatedly recognized by significant legal documents through the Middle Ages (Bloch, 1962b, p. 172–173).[7] Importantly, this right was "not subject to the king's justice" and "not upon the desires of the king" (Tamanaha, 2004, p. 26).[8]

Ultimate ownership of land. While land could be owned by individuals in normal times in China, the ultimate legitimacy of land ownership was always reserved for the Ruler, so the emperor could re-centralize the ownership when he deemed it necessary (Chao and Chen, 1982; Wang, 2000). Since even before the Qin dynasty unified China in 221 BC, land confiscation from the noble families and landed gentry had been a common practice of the Chinese Ruler to raise revenue for military projects (Ebrey and Walthall, 2013).[9] Depending on the emperor's will, systematic persecutions against Buddhism, Manichaeism, and other religions also repeatedly happened, regularly entailing large-scale confiscation of temple properties (de Groot, 1903, ch. 2).

In contrast, when European Rulers needed revenues, they could usually not confiscate

[5]For example, the founding emperor of the Ming dynasty created "law beyond the law" when he was frustrated by the Great Ming code of his own, while insisting that only he could use the newly created law (Brook, 2010, p. 87). Unger (1977, ch. 2) discusses the characteristics of law in Imperial China in detail.

[6]Famous examples include the dramatic scenes of Henry IV of Germany at Canossa, Henry II of England at Canterbury, and King John of England at Dover, and the destruction of Holy Roman Emperor Frederick II's family.

[7]Bloch (1962b, p. 173) raises examples of "the English Great Charter of 1215; the Hungarian 'Golden Bull' of 1222; the Assizes of Jerusalem; the Privilege of the Brandenburg nobles; the Aragonese Act of Union of 1287; the Brabantine charter of Cortenberg; the statute of Dauphiné of 1341; the declaration of the communes of Languedoc (1356)."

[8]For more extensive discussion on the rule of law, see Finer (1997b), Tamanaha (2004), Fukuyama (2011), Vincent (2012), Fernández-Villaverde (2016), Acemoğlu and Robinson (2019), and Greif, Mokyr and Tabellini (Forthcoming).

[9]Among famous early examples, Duke Xiao of the Qin state confiscated land from the feudal nobles in the 340s BC, sharing it among the peasants; in 114 BC, Emperor Wu of Han confiscated land from nobles and merchants to raise additional revenue to fund the Han–Xiongnu War.

land from the Elites or the Church, at least during the 9–14th centuries.[10] Instead, they had to exchange rights or resources with revenues. Levi (1988, p. 99) states it clearly: "[d]uring the medieval period, a monarch was expected to 'live of his own' (*vivre du sien*). That is, funds for the monarch were to come from royal lands and customary dues. ... Should monarchs need more, even if it was to fund a campaign on behalf of the country as a whole, they had to obtain assent to some form of 'extraordinary' taxation. They could neither expropriate property at will nor rely on a regular levy."[11]

Ruler's control over population. As the population were subjects of the Ruler in China, he could reward or punish anyone arbitrarily, which precisely reflected his absolute power (Levenson, 1965, p. 39; Finer, 1997a, p. 455). Consistent with the emphasis of Confucianism on the loyalty of the ruled to the Ruler (Greif, Mokyr and Tabellini, Forthcoming), one person's rebellion, treason, or even slight disobedience, regardless of her social status, would be punished extremely harshly, usually leading to eradication of the whole family line (Finer, 1997b, p. 778).[12] Sometimes mere suspicion from the Ruler could guarantee the calamity, as shown in the fall of Princess Taiping in 713.[13] Following the Legalist tradition in Chinese political philosophy, the absolute right to impose harsh punishment on the ruled could effectively help the Ruler to grip control of both the Elites and the People, despite sometimes significant administrative constraints (e.g., Watson, 1964; Sng, 2014).

In contrast, in feudal Europe, the king, in practice, did not have direct control over peasants because the latter were controlled by their overlords; the peasants could, as a rule, be punished by local courts controlled by their overlords, and the king did not have control over these local courts (Bloch, 1962a). Although loyalty was also emphasized in Europe and enforced through mechanisms like oaths, treason was punished much less harshly than in

[10]European Rulers could become more capable of expropriating the Church as their absolute power grew, but mainly in a later period, i.e., the 16–18th centuries, and especially during the Reformation. One may also note that the famous dissolution of the English monasteries under Henry VIII (1536–1541) was accompanied by the rise of the gentry, who were commoners but their power and rights grew relative to the peerage (Heldring, Robinson and Vollmer, 2021). This was consistent with Proposition 3 below about the compatibility between a greater absolute power of the Ruler and a more balanced Elite–People relationship.

[11]See also Finer (1997b, p. 887) for a similar observation. Besides, when Louis XIV managed to tax the nobility for the first time, the taxes happened only at the end of his reign and were insignificant in size and subject to numerous exemptions (McCollim, 2012). Expropriations did happen but mostly under Eminent Domain (Reynolds, 2010); in case of serious crimes like treason, the nature of the crime had to be determined by law, not merely the Ruler's will (Lander, 1961).

[12]In a famous case, when Fang Xiaoru, a prominent minister, refused to write an inaugural address for Emperor Yongle of Ming, the emperor sentenced 873 people to death, including Fang's family, kinfolk, friends, and students, before having Fang himself executed.

[13]In 713, Emperor Xuan of Tang, merely suspecting that his aunt Princess Taiping had been planning a coup, forced her to commit suicide and executed several dozens of her extended family and allies. Literary inquisitions for merely *potentially* subversive attitudes to the Ruler were also conducted at a frequency and scale much more significant than in Europe (e.g., Xue, 2021).

China. First, although execution of the traitor and attainder could apply, killing the family seldom applied, and the attainder would often later be reversed (Lander, 1961).[14] Second, it was common in the feudal system for a vassal to have two or more overlords (Bloch, 1962a) and when in conflict, he could simply choose which to follow (e.g., Cantor, 1964, p. 202; Tuchman, 1978; Mann, 1986). Eventually, as Finer (1997b, p. 881) observes, the Ruler's control over the population was "abysmal" and he "could not always count on the fidelity of the vassal," precisely because his lack of ability to punish them: "after all, [they were] in possession of his lands and what could he do if defeated?"

Formalization in our model. Motivated by these narratives, we assume that the Ruler, Elites, and People are sharing a surplus of size π; when the Ruler had survived a challenge to the status quo of his rule in which the ruled did not side with him, he could punish the defiers by having them enjoy only γ of their share of the surplus. Given the initial distribution of the surplus, a lower γ indicates that more of the power and rights of the ruled are conditional on the Ruler's will, i.e., the Ruler has a stronger absolute power and is more capable of exerting punishment on the ruled for defiance.

2.3 Asymmetry in Power and Rights between Elites and People

The differences in the power structure between Imperial China and Premodern Europe lied also in the relationship between the Elites and the People. In Bloch's words, the disparity between "[a] subject peasantry" and "the supremacy of a class of specialized warriors" was one of "the fundamental features of European feudalism" (Bloch, 1962b, p. 167), and his final verdict on the system concerns only its bindings on the Rulers and its oppressiveness to the poor (Bloch, 1962b, p. 173). In contrast, prominent Chinese historian Lü Simian summarizes the scenario in Imperial China elegantly: "once the father or elder brother takes the throne, the sons and younger brothers," who are princelings themselves, "will become mere commoners" in terms of their power and rights (Lü, 1944, p. 347). Finer (1997b, p. 836) also comments that "[i]n practice, China was a two-class society" where the ruled, including "even the higher mandarins" and commoners, were having relatively equivalent power and rights in front of the Ruler. The difference in the Elites–People relationship was reflected by differences in, for example, the dominance of the hereditary versus non-hereditary access to elite status, inequality in land ownership, and the inheritance rule. As above, we summarize the narratives and then explain how we formalize them in our model.

[14]For example, during the reigns from Henry VI to Henry VII of England, 64% of the attainders were eventually reversed (Lander, 1961, p. 149).

Hereditary vs. non-hereditary access to elite status. In Medieval Europe, elite status was governed primarily by hereditary nobility. As Finer (1997b, p. 879–880) explains, "lineage [was] much more important than initiation," while "the very right to be a vassal (i.e. to hold a fief) [was] confined to those already noble!" Government positions, especially in courts and the army, were reserved for aristocrats. Although ordinary peasants routinely performed military service as a privilege in the early Middle Ages, this stopped to be the case later and was reserved for knights and higher titled nobles (for more discussion, see, e.g., Wickham, 2009). Access to priesthood and religious orders was not forbidden to commoners, but even after the Gregorian reform in the 11th century, "the abolition of . . . the hereditary ecclesiastical benefice" had remained a "formidable task" in western Christendom until as late as the 13th century (Kemp, 1970, p. 1; Parish, 2010, p. 88–92).

In contrast, as early as during the 5–4th century BC, accompanied by reforms that strengthened the absolute power of the Ruler, the Warring States in China had started to abolish hereditary titles and make elite status dependent solely on military merit and open to the common People (Yan, 2009, p. 23–24). To facilitate the fluid exchange between the Elites and the People, the Sui dynasty (581–619) established the civil service exam to regulate elite status, and the exam system was greatly developed during the Tang dynasty (618–907). Notably, the exam was in principle open to almost all adult males, and elite status gained via success in the exam could not be inherited. Following the destruction of the aristocratic clans during the fall of the Tang dynasty, elite status in China had been governed mainly by the exam system. Sustained by "a culture of merit," the resulting Chinese Elites were "more diffuse [and] justified . . . on the basis of talent and education" instead of hereditary titles, which "would constitute one of the most striking distinctions between Chinese and Western societies over the course of the subsequent millennium" (Tackett, 2014, p. 3–5).

It may be worthwhile to comment here on the difference between the *de jure* and *de facto* access to elite status. First, given the lack of comparative historical evidence on the *de facto* difference in the access to elite status between Imperial China and Premodern Europe for the 9–14th centuries, when our characterization of the power structure was the most relevant, we do not take a strong stand on this subject.[15]

[15]In Medieval Europe, with the system of hereditary aristocracy, the most visible way of social preferment for commoners was the Church, which could have been comparable to the civil service exam in China. According to Herlihy (1973) and Barrow (2015), however, it was still mainly the few landowners, patricians, or clerics themselves, if allowed, who sent their children to the clergy, since only they needed to cut down the numbers of heirs *and* could afford losing precious family labor. In practice the relevance of ecclesiastical careers to commoners was thus limited. Herlihy (1973) also identifies three main patterns of social mobility in Medieval Europe, and ecclesiastical careers were not among them. For later periods only scattered evidence for China and England is available. Ho (1959) documents that during 1752–1938, 78%–88% of Cambridge students came from elite families, whereas during the 13–19th centuries, only 50%–65% of the highest degree holders (Jinshi) in the Chinese civil service exam system came from elite families; Clark (2014, p. 86) shows

Second, we emphasize in our framework the *de jure* difference in the access to elite status. It is important to note that the *de jure* access to elite status alone can shape the belief in society about the *de facto* access, affecting the stability of the autocratic rule. For example, Bai and Jia (2016) show empirically that China's abolition of the civil service exam in 1905 caused an increase in revolutionary activities against the Qing court, contributing to the end in 1912 of not only the Qing dynasty but also the imperial era. One interpretation for such evidence is that the People's belief in the alliance with the Ruler was temporarily broken when the abolition of the civil service exam shut down the main *de jure* access of the commoners to elite status.

Inequality in land ownership. Circumstances on land ownership inequality are also suggestive. In Imperial China, peasants "were mostly free" (Finer, 1997a, p. 205), "land-owning peasantry had been the main agent and form of agricultural production," and they "had mostly enjoyed the freedom of choice" (Chao and Chen, 1982, p. 192–193).[16] In contrast, in early-Medieval Europe, mostly between the 8th and 10th centuries, small peasants became gradually expropriated by rich aristocrats as well as by the Church, making peasants gradually fall entirely under the control of landlords. This happened in many ways, as documented by Wickham (2009): First, in the aftermath of the Viking incursions, some landlords became richer and acquired more land, usually from poor peasants, either through payment or expropriation. Tenant peasants faced higher rents and greater control over their labor. They became gradually submitted to the judicial control of landlords and completely lost their freedoms to become feudal serfs. The only escape route for encaged peasants was to flee to the cities, a process that accelerated with the Black Death, but those living in the countryside remained heavily under the control of landlords until much later on.[17] In the 17th century in England, around 70% of the land was still owned by landlords and gentry (Beckett, 1984). Almost all scholars on China would agree that the corresponding number remained below 45% from the 6th century to modern China (e.g., Esherick, 1981; Chao and Chen, 1982).[18] Even during the Tang dynasty when the aristocratic families still had considerable political influence, they "did not maintain large landed estates over multiple generations" (Tackett,

that the surname-approach estimate of the intergenerational correlation of elite status for England during 1380–1858 is about 0.81–0.85, whereas Hao and Clark (2012) show that the estimate from the same approach for Zhejiang and Jiangsu in China during 1645–1810 is about 0.81–0.89. These results suggest that during the studied periods, the *de facto* social mobility in China was comparable to that in England, if not significantly higher.

[16]See von Glahn (2016, ch. 6, 8) for a similar observation from the mid-late Tang dynasty on.

[17]It is important to note that the stronger property rights of land in Europe documented by historians in reality concern mainly whether the rights of landlords were independent of the arbitrary will of the Ruler, not whether small peasants enjoyed certain rights in their normal, everyday life.

[18]For extensive discussion on the many works on England and China, see Zhang (2017).

2014, p. 12).

Inheritance rule. The differences in land ownership concentration are partly related to differences in inheritance rules. China gradually switched from primogeniture to partible inheritance in the Qin and Han dynasties (221 BC–220), while primogeniture became more common in Europe during the Middle Ages (Goody, Thirsk and Thompson, 1976; Bertocchi, 2006; von Glahn, 2016, ch. 2, 8). The consequence of these rules on elite privilege is intuitive: partible inheritance makes it more difficult for elite families to accumulate assets over generations. As Goldstone (1991, p. 380) observed, in China, "land was generally divided among heirs, and over a few generations such division could easily diminish the land holdings of gentry families. At the same time, peasants, who could purchase clear and full title to their lands, might expand their holdings through good luck or hard work. Thus the difference between the gentry and the peasantry was not landholding per se, but rather the cultivation, prestige, and influence that came from success in the imperial exams."

Formalization in our model. Motivated by these narratives, we capture the relative power of the Elites and the People by a simple parameter β. With the surplus of size π mentioned above, the Elites will get a and the People will get βa, if they have been loyal to the Ruler, where $0 \leq \beta \leq 1$ and a higher β indicates a more symmetric Elites–People relationship. This approach allows us to avoid modeling the exact mechanism of each specific institution, for example, the hereditary versus non-hereditary access to elite status, land ownership, freedom of peasants, and the inheritance rule; we can instead focus on their general implications on the Elite–People relationship in the power structure.

Remarks. To be sure, both China and Europe experienced changes and challenges of the power structure over the centuries. It should not be surprising that multiple Rulers in Europe, especially during the early modern period, attempted to make the Elite–People relationship more balanced. Nevertheless, the weaker *de facto* power of the Ruler and the multiple checks on executive power by the Elites in Europe generally made it less possible for the Ruler to consistently succeed in these kinds of endeavors.[19] In Online Appendix D, we show that our main model can be extended to accommodate this interpretation, where we allow the current political stability to affect the future power structure. In Section 4, helped by our theory, we discuss further the rise of cities in Medieval Europe, another phenomenon

[19]For example, Louis XIV insisted on depriving the nobility of actual power after the rebellions of the Fronde, attempted to choose ministers and officials on merit, and used commoners to replace aristocrats. Even though he succeeded temporarily, access to nobility through a judiciary and administrative office became practically barred in 18th-century France.

related to the Ruler's hope to enlist the People as allies against the Elites by granting more power and rights to urban commoners.

3 Comparative Institutional Analysis

Now we introduce the setting of our model. We assume that there is a Ruler (R), who prefers a certain status quo of autocratic rule. The nature of the status quo is open to interpretation: for example, it can be a peaceful, unified autocratic rule across the territory. There is also a Challenger (C), who is unhappy about and can challenge the status quo. She could be one or a group of nobles, lords, or bureaucrats, or some common people who are under R's rule, or a foreign threat who is not under R's rule; her challenge may or may not seek to dethrone R or be violent. With such flexibility in interpretation, the model is sufficiently general to accommodate different types of threat to autocratic rule, such as external conflict, elite revolt, coup, or secession, popular uprising, independence war, and other apparently non-violent attempt to alter the status quo, with or without a competing claim over the ruling position.

Besides R and C, we assume that there are also the Elites (E), which represents the nobles, lords, and bureaucrats, and the People (P), which includes peasants and urban commoners in the model. When interpreting E and P, depending on the identity of C, we exclude the initial challenger from E and P. For example, if C were a group of elites, then E would be the other elites; if C were a group of members of the people, then P would be the other members of the people. We interpret E and P's actions as whether all significant members of each estate actively side with and fully support R to preserve the status quo or not, focusing on the alliance across R, C, E, and P. Naturally, unanimous actions were rare in reality both within E and within P; the model can be easily extended to analyze the collective action problem within each estate.

As we have included both E and P in the model, there can be little doubt that, given their significant political, economic, and military resources, the Elites' position had a great impact on autocratic stability in history, whereas concerns may arise about whether the common People were relevant, especially in Medieval Europe. In response to the concerns, we first provide in Online Appendix F historical examples where the People's position were critical in determining the outcome of a conflict, an important type of threats to the stability of the Ruler's rule in both Europe and China. Second, our model will be able to explain as an equilibrium outcome that in Europe the autocratic stability looked largely reliant on the Elites but not the People: later analysis in Section 3.1.2 suggests that if the Elite–People relationship is extremely asymmetric, as in Medieval Europe, then the People will almost

never actively support the Ruler when called upon, making their action seemingly irrelevant and the Elites' position apparently decisive to the status quo. Finally, one may also note that even if we did not observe any significant move of the People in reality, it does not suggest that the People were irrelevant; on the contrary, they may have been influential on the off-equilibrium path, which we could not observe but may have been instrumental in supporting the observed outcome as an equilibrium.

The four players in our model play a game of two stages. Stage 2 is about the stability of the status quo of autocratic rule, where C, E, and P play a subgame while taking as given the power structure. Stage 1 is about R's design of the power structure, where R chooses the degree of asymmetry between E and P in terms of their power and rights, while foreseeing Stage 2 and taking as given the level of his absolute power. Across the two stages, we assume that all players maximize their own expected payoff. Given the two-stage structure, we now introduce in detail and analyze Stage 2, and then move back to Stage 1.

3.1 Stage 2: Stability of Autocratic Rule

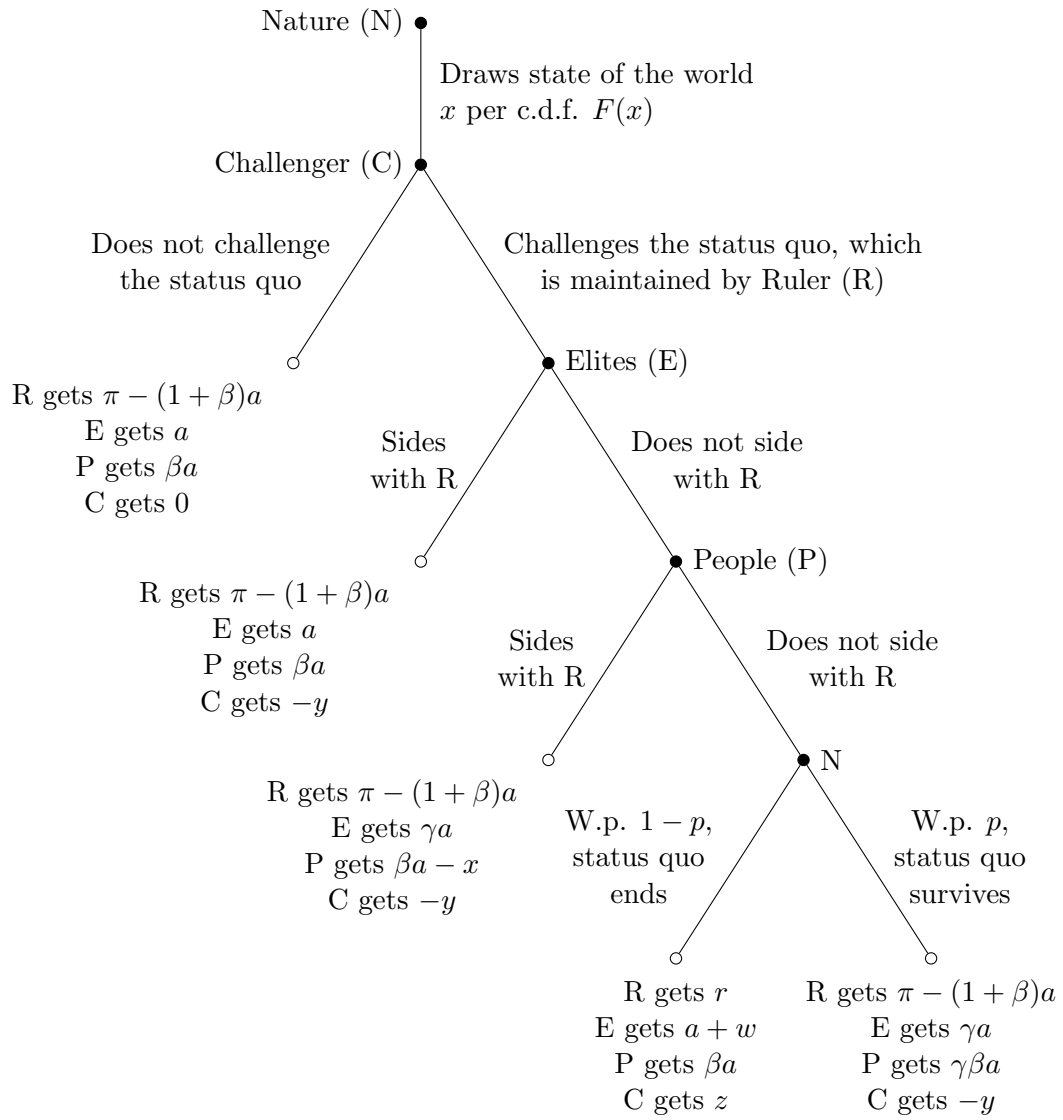
3.1.1 Setting

Figure 1 presents the setting of Stage 2. Nature (N) first randomly draws a state of the world $x \geq 0$, following the exogenous cumulative distribution function $F(x)$. The state of the world x will appear later in the game as the cost born by P if she sides with R.

Given x, C will decide whether to challenge the status quo, which is maintained by the rule of R. If C does not challenge, then C will get her default payoff 0; E will get her status quo payoff $a > 0$, which is exogenous; P will get βa, where $\beta \in [0, 1]$ measures the power symmetry between E and P in the status quo and is exogenous at this stage; R will get the exogenous total surplus π net of the sum of E and P's status quo payoffs $(1 + \beta)a$, which is $\pi - (1 + \beta)a$ in total. Stage 2 then ends there.

If C instead does challenge, then E will decide whether to side with R. If E sides with R, then the status quo will survive. Stage 2 will end there with R, E, and P all getting their status quo payoffs, respectively, while the failed challenge will incur an exogenous loss $y > 0$ to C, leaving her the payoff $-y$.

If E instead does not side with R, then it will be P's turn to decide whether to side with R. If P decides to side with R, then the state of the world x comes in as the cost incurring to P for the choice, while the status quo will survive. In this scenario, C will still get $-y$ for the failed challenge; R will still get his status quo payoff $\pi - (1 + \beta)a$; P will get her status quo payoff βa but net of the cost x, which is $\beta a - x$ in total; E will now suffer a punishment because she has not sided with R, getting only γa instead of her status quo payoff a, where



$x \geq 0$, $a > 0$, $\pi - 2a > r$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$, $0 < p < 1$, $w > 0$, $y > 0$, $z > 0$

Figure 1: Stage 2: Stability of Autocratic Rule

$\gamma \in [0, 1]$ is exogenous. A lower γ measures a stronger absolute power of R to punish its subjects who have defied him. For simplicity, we assume that the destroyed part of E's status quo payoff, $(1-\gamma)a$, evaporates and is not going to R; assuming otherwise would complicate Stage 1 with few additional insights. Stage 2 then ends there.

If P does not side with R either, then R will be left on his own. N will then determine randomly whether the status quo will survive. With exogenous probability $p \in (0, 1)$, the status quo will survive, so C will still get $-y$ for the failed challenge; R will still get his status quo payoff $\pi - (1 + \beta)a$; E will be punished, getting γa; P will be punished, too, getting $\gamma\beta a$; as above, we still assume that the destroyed parts $(1 - \gamma)a$ and $(1 - \gamma)\beta a$ evaporate

and are not going to R. Stage 2 then ends there.

With probability $1-p$, the status quo will end, leaving C with an exogenous prize $z>0$ and R an exogenous reservation payoff r, where we assume, intuitively, $\pi-2a>r$ so that, given any $\beta\in[0,1]$, R would prefer the status quo to survive. P will still get her status quo payoff βa, while E will now get an exogenous incentive $w>0$ for having not sided with R, in addition to her status quo payoff a, so her total payoff will be $a+w$. Stage 2 then ends there.

About the random elements, we assume that N's draws of x and whether the status quo will survive on R's own are mutually independent. About the informational environment, we assume that in Stage 2 there is complete and perfect information. We will thus use backward induction to solve for subgame perfect equilibria.

For simplicity, we assume that E and P will side with R if indifferent, respectively, and C will not challenge if indifferent, ruling out mixed strategies. Online Appendix A shows that the insights from our results would remain robust if mixed strategies were allowed.

Before analyzing Stage 2, we make a few remarks on the conceptual and technical issues around the current setting.

Remarks. First, as discussed in Section 2, we interpret Imperial China as having a high β and a low γ, while Premodern Europe has a low β and a high γ. The β–γ characterization of the power structure captures the idea that power and rights are specific to estates and scenarios, as β measures the E–P asymmetry and γ measures how much the power and rights of the ruled depend on whether they have defied R.

Second, an alternative approach to model the E–P relationship is to assume that E and P's status quo payoffs are $(1-\beta')\,a'$ and $\beta' a'$, respectively, where $\beta'\in[0,1/2]$ measures the E–P symmetry and $a'>0$ measures the sum of their status quo payoffs, instead of a and βa, respectively, as in our current approach. Comparing the two approaches, first, as shown in Proposition 1 below, C and E will follow P's strategy in equilibrium in Stage 2, and all further results depend only on how γ and β or β' would affect P's best strategy in the equilibrium. Since P's status quo payoffs have the same form in the two approaches, i.e., either βa or $\beta' a'$, the two approaches will thus derive the same theoretical results. That said, as shown in Section 3.2 below, the current approach will create a political–economic trade-off for R in Stage 1, making R's problem non-trivial. This is achieved without the help of any additional modeling device that would be necessary if the alternative approach were adopted. In light of these considerations, we opt for our current approach.[20]

[20]On the empirical side, there is little historical evidence comparing the Elites' power and rights between Imperial China and Premodern Europe. This has already made it difficult for us to generate empirical implications related to a or a'. The consequence of the different empirical implications from the two approaches

Third, as mentioned, C can be an outsider or an elite member or part of the people; the incentive for E not to side with R also depends on the specific context.[21] Thus, for generality and simplicity, we model any incentives of C and E that are additional to the power structure via the exogenous variables w, y, and z that are added to C and E's payoffs. On the robustness of our analysis to this approach, first, modeling these incentives as multiplicative terms would not affect our results, since Proposition 1 below will suggest that in the focal equilibrium, these additional incentives are irrelevant at the margin. Second, in Online Appendix E, we endogenize these additional incentive by collapsing C and E into a single player E from the inside of the autocratic rule, making her look forward infinitely in a Markov game, and allowing her to replace R if her challenge succeeds. We show parallel results in Online Appendix E to all results in the main text.

Fourth, P's incentive not to side with R depends also on the specific context, for example, P's level and prospect of income, R's level of legitimacy, whether and how severely R is in a crisis, and whether P has an opportunity to revolt, all of which can be affected in turn by many random factors. We thus model this random component of her incentive that is additional to the power structure as a single, exogenously drawn, state-of-the-world variable, i.e., the random cost x added to P's payoff when he sides with R. Modeling it alternatively as a reward for not siding with R would not affect our analysis.

Fifth, in the current setting, we have assumed that C, E, and P move sequentially. As we will show, this has the advantage of simplicity when we highlight the political alliance channel through which γ and β affect E and C's equilibrium strategies by affecting P's equilibrium strategy. The political alliance channel always exists, unless P has moved strictly earlier than both C and E, which is unrealistic; any other sequences of moves, for example, C, E, and P moving simultaneously or E and P moving simultaneously after C, would not affect the insights of our analysis.

Finally, we have chosen not to focus on the possibility of contracting among R, C, E, and P at Stage 2. It is not too unreasonable in reality, since any threat R or C can exert upon E and P depends on the status quo's own survival or the success of C's challenge, respectively, and any reward R or C can promise to E and P is not too credible, since the need for cooperation will disappear once the status quo survives or C's challenge succeeds, respectively. That said, when considering R's preference about β at Stage 1, one can interpret a higher β as an implicit contract between R and P where R grants more rights to P in exchange for support in scenarios where P would not support R with a lower β. At the same time, the severity

is thus limited, too.

[21] For example, E could hope to replace R in the challenge, or simply to get more power, rights, or other economic interests, or even to secede from the Ruler, without necessarily taking the ruling position; similarly, C could hope to replace R, or to secede from R, or simply to loot a great fortune in the challenge.

of the credibility problem may be endogenous to the β–γ power structure. A more explicit exploration on the contracting across R, C, E, and P could be an interesting direction for future research.

3.1.2 Equilibrium Characterization

We start the backward induction from P's strategy. In any subgame perfect equilibrium, P will side with R if and only if

$$\beta a - x \geq (1-p) \cdot \beta a + p \cdot \gamma \beta a, \tag{1}$$

i.e., the cost of siding with R is not greater than the probability-adjusted punishment for not siding with R in case that C's challenge fails:

$$x \leq p \cdot (1-\gamma)\beta a \equiv \hat{x}. \tag{2}$$

As mentioned when introducing the players of the model, one may note here that if the power structure has an extremely asymmetric relationship between E and P, i.e., β is close to zero, then the critical threshold $\hat{x}$ will be extremely low, i.e., in equilibrium P will almost never actively help R out when called upon, making R largely reliant on E. P may thus look irrelevant to the fate of the status quo, but E must still consider P's strategy when solving for his own best strategy.

Now consider E's best strategy while expecting P's strategy in equilibrium, i.e., to side with R if and only if $x \leq \hat{x}$. When $x \leq \hat{x}$, P would side with R, so E will side with R; when $x > \hat{x}$, P would not side with R, so E will not side with R if and only if

$$a < (1-p) \cdot (a+w) + p \cdot \gamma a, \tag{3}$$

i.e., the incentive for not siding with R is greater than the probability-adjusted punishment in case that C's challenge fails:

$$w > \frac{p}{1-p} \cdot (1-\gamma)a. \tag{4}$$

This analysis implies that if this condition does not hold, then in any subgame perfect equilibrium, E will always side with R so that it will be impossible for the status quo to end. Such equilibria are empirically irrelevant, as in reality the chance for the status quo to end was always strictly positive; such equilibria are also theoretically trivial, in the sense that E and P will always side with R regardless of the state of the world. Therefore, to narrow

our focus onto empirically more relevant and theoretically less trivial scenarios, we assume $w > a \cdot p/(1-p)$ so that for any $\gamma \in [0,1]$, in any subgame perfect equilibrium, E will not side with R if and only if $x > \hat{x}$.

Under this assumption, now consider C's strategy while expecting these strategies of E and P in equilibrium. When $x \leq \hat{x}$, E would side with R, so C will not challenge the status quo; when $x > \hat{x}$, E and P would not side with R, so C will challenge the status quo if and only if

$$0 < (1-p)z - py, \tag{5}$$

i.e., the prize from a successful challenge is greater than the probability-adjusted loss from a failed challenge:

$$z > \frac{p}{1-p} \cdot y. \tag{6}$$

This analysis implies that if this condition does not hold, then in any subgame perfect equilibrium, C will never challenge the status quo. Similar to the discussion above, such equilibria are empirically irrelevant and theoretically trivial. Therefore, to further narrow our focus onto empirically more relevant and theoretically less trivial scenarios, we further assume $z > y \cdot p/(1-p)$ so that in any subgame perfect equilibrium, C will challenge the status quo if and only if $x > \hat{x}$.

Note that under the two assumptions we have introduced, we have found the unique strategy of each player in any subgame perfect equilibrium, so these strategies constitute a unique subgame perfect equilibrium. To summarize:

Proposition 1. *If $w > a \cdot p/(1-p)$ and $z > y \cdot p/(1-p)$, then for any $\beta \in [0,1]$ and $\gamma \in [0,1]$, there exists a unique subgame perfect equilibrium at Stage 2, in which C will challenge the status quo if and only if $x > \hat{x}$, E will not side with R if and only if $x > \hat{x}$, and P will not side with R if and only if $x > \hat{x}$, where $\hat{x} \equiv p \cdot (1-\gamma)\beta a$.*

This equilibrium is indeed theoretically non-trivial, since in the equilibrium, whether C will challenge the status quo and start a challenge and whether E and P will side with R all depend on the state of the world; this equilibrium is also empirical relevant, since in the equilibrium, a challenge of the status quo can happen and E and P may not side with R, i.e., the probability of challenge $1 - F(\hat{x})$ can be strictly positive and the survival probability of the status quo

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p) \tag{7}$$

can be strictly lower than one. Therefore, to focus on this equilibrium, from now on we assume that the condition in Proposition 1 holds, i.e., $w > a \cdot p/(1-p)$ and $z > y \cdot p/(1-p)$.

3.1.3 Impact of Power Structure on Stability of Autocratic Rule

How does the β–γ power structure shape the probability of challenge and the survival probability of the status quo in equilibrium?

Proposition 2. *At Stage 2, a higher β and a lower γ decrease the probability of challenge and increase the survival probability of the status quo of autocratic rule in equilibrium.*

Proof. By Proposition 1, the probability of challenge is $1-F(\hat{x})$ and the survival probability of the status quo is $S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p)$, so a higher $\hat{x}$ lowers $1-F(\hat{x})$ and raises S. Since a higher β and a lower γ increase $\hat{x} \equiv p \cdot (1-\gamma)\beta a$, the proposition then follows. □

The intuition of Proposition 2 deserves more discussion. In the model, β and γ influence the stability of the status quo in equilibrium by their impacts on P, E, and C's equilibrium strategies. We discuss each of these impacts. First, the impacts of β and γ on P's strategy in equilibrium are straightforward: by Equation (2), P's strategy hinges on the comparison between her cost x for siding with R and the probability-adjusted punishment $\hat{x} \equiv p(1-\gamma)\beta a$ for not siding with R in case C's challenge fails; both a higher β and a lower γ impose a heavier punishment $(1-\gamma)\beta a$, making P more willing to side with R in equilibrium. We can say that these impacts work through a generic, *punishment* channel.

Second, the impact of γ on E's strategy in equilibrium generally has two channels. The first is again the punishment channel: a lower γ imposes a heavier punishment $(1-\gamma)a$ on E in case C's challenge fails, making E more willing to side with R *given any strategy of P*, including the one in equilibrium. The second, which is new, is a strategic, *political alliance* channel: a lower γ makes P more willing to side with R in equilibrium, lowering the chance for C's challenge to succeed and, therefore, making E more willing to side with R in the first place.[22] Therefore, through both channels, a lower γ makes E more willing to side with R in equilibrium.

In the specific case of Proposition 2, under the condition $w > a \cdot p/(1-p)$, E always prefers "both herself and P not siding with R" to "herself siding with R", and further to "herself not siding with R while P siding with R." Meanwhile, P will always either side with or not side with R, and her decision solely depends on x, so E does not face strategic uncertainty about P. Therefore, a heavier punishment upon E brought by a lower γ would not change the fact that E's best response to P's strategy in equilibrium is to "follow" P's strategy, i.e.,

[22] To see the point, observe that when deciding whether to side with R, E compares the payoff of doing so, i.e., a, versus the payoff of not doing so, i.e., $\mathbf{P}[\text{P sides with R}|x,\gamma] \cdot \gamma a + \big(1 - \mathbf{P}[\text{P sides with R}|x,\gamma]\big) \cdot \big((1-p)\cdot(a+w) + p\cdot\gamma a\big)$, where P's strategy is represented by $\mathbf{P}[\text{P sides with R}|x,\gamma]$. There are two channels via which γ can influence this comparison: first, γ can affect γa in the payoff of siding with R, which is the punishment channel; second, γ can affect $\mathbf{P}[\text{P sides with R}|x,\gamma]$, which is the political alliance channel.

to switch between to side or not to side with R at $x = \hat{x}$. Therefore, the punishment channel is muted and we observe only the political alliance channel.[23]

Finally, the impact of β on E's strategy and the impacts of β and γ on C's strategy in equilibrium have only the political alliance channel: β does not affect E's payoffs at any of the five ending nodes of the game, and β and γ do not affect C's payoffs at these nodes, either, but a higher β makes E more willing to side with R by making P more willing to side with R in equilibrium, whereas a higher β and a lower γ make C more reluctant to challenge by making P and E more willing to side with R in equilibrium.

To summarize, Proposition 2 reveals that both a higher β and a lower γ will make P more willing to side with R, thus E more willing to side with R, and, therefore, C more reluctant to challenge the status quo in the first place. The probability of challenge is then lowered and the status quo becomes more stable. In our specific setting, a generic punishment channel appears in β and γ's impacts on P's strategy; it exists in γ's impact on E's strategy but is muted, with only a strategic political alliance channel visible; in β's impact on E's strategy and β and γ's impacts on C's strategy, only the political alliance channel exists. All these make the impacts of β and γ on political stability come from only their impacts on P's switching threshold $\hat{x}$, providing much simplicity for the result.

Proposition 2 thus highlights that how well R can form an alliance with P is critical in shaping the stability of autocratic rule.[24] This proves crucial in R's design of the power structure at Stage 1, which comes below. Also, by Proposition 2, compared with Europe, both a higher β and a lower γ make an autocratic rule more stable in China. We will come back to this implication in Section 4.

3.2 Stage 1: Design of Power Structure

3.2.1 Setting

This stage characterizes how R's incentive to promote the symmetry between E and P can depend on the level of his absolute power. We assume that R at this stage simply chooses

[23] If E faced strategic uncertainty about P, the punishment channel would not be muted. For example, suppose E did not observe x when deciding whether to side with R. She would then compare a versus $\int_0^{\hat{x}} \gamma a \cdot dF(x) + \int_{\hat{x}}^{\infty} \big((1-p)\cdot(a+w) + p\cdot\gamma a\big)\cdot dF(x)$. As a lower γ will strictly lower the latter sum by lowering γa, its impact on E's decision via the punishment channel would be visible.

[24] Chapter 17 in *Han Feizi* argues that "too much compulsory labor service" upon the People (low β) would make it easy for the Elites to shelter the People in exchange for their financial and political support against the Ruler (low $\hat{x}$), damaging the Ruler's "long lasting benefit" (low S, Watson, 1964, p. 87). This argument follows exactly the modeled impact of β on the stability of autocratic rule via the political alliance channel in this analysis and Online Appendix E.

β, while foreseeing the equilibrium at Stage 2 and taking γ as given. R's program is thus

$$\max_{\beta} \quad V^R \equiv (\pi - (1+\beta)a) \cdot S + r \cdot (1 - S), \text{ subject to} \tag{8}$$

$$0 \leq \beta \leq 1, \quad S = 1 - \big(1 - F(\hat{x})\big) \cdot (1 - p), \quad \hat{x} = p \cdot (1 - \gamma)\beta a, \tag{9}$$

where V^R is R's expected payoff from Stage 2. Without losing generality, we also assume that the state of the world x's probability density function is always strictly positive while finite in the relevant range, i.e., satisfies $f(x) \in [\underline{f}, \bar{f}] \subset (0, \infty)$ over $x \in [0, pa]$.

3.2.2 Institutional Compatibility

How will R choose β by the program? There is a political–economic trade-off: by Proposition 2, a higher β will increase the survival probability S of the status quo at Stage 2, which is political; at the same time, it will decrease the status quo payoff $\pi - (1+\beta)a$ at Stage 2, which is economic.

The economic side of the trade-off is straightforward: a higher β will decrease the status quo payoff at a marginal rate of a. The political side is less so, as it depends on the impact of β on the survival probability, i.e., $dS/d\beta$. Intuitively, this impact is largely governed by γ: a higher γ suggests that P will not lose much of her status quo payoff after she has not sided with R and C's challenge has failed, so any additional status quo payoff would not make her much more loyal to R and, therefore, it will not make E much more loyal toward R, and neither would C be much more reluctant to challenge.

The key assumption that leads to this intuition is that the punishment upon P, i.e., $(1 - \gamma)\beta a$, is multiplicative between $1 - \gamma$ and β. We find this assumption uncontroversial, since in reality, given the punishing institution against defying behaviors, the ones who own more would often be more concerned about losing it.

We can formalize this intuition by showing that the impact of β on the survival probability of the status quo can be approximated by two positive and increasing functions of $1 - \gamma$:

Lemma 1 (Impact of β on stability governed by γ)**.** *There exist* $\underline{c} \equiv (1-p)p\underline{f} > 0$ *and* $\bar{c} \equiv (1-p)p\bar{f} > \underline{c}$ *such that*

$$\underline{c}a \cdot (1 - \gamma) \leq \frac{dS}{d\beta} \leq \bar{c}a \cdot (1 - \gamma). \tag{10}$$

Proof. By Proposition 1, the marginal impact of β on S is

$$\frac{dS}{d\beta} = (1 - p) \cdot \frac{dF(\hat{x})}{d\beta} = (1 - p)pf\left(\hat{x}\right) \cdot a \cdot (1 - \gamma), \tag{11}$$

where $\hat{x} \equiv (1-\gamma)\beta p \cdot a \in [0, pa]$. By $f(x) \in [\underline{f}, \bar{f}]$ over $x \in [0, pa]$, the lemma follows. $\square$

The proof of Lemma 1 also suggests that the approximation would be exact if and only if the state of the world x followed a uniform distribution, i.e., $f(\hat{x})$ is a constant. Such an assumption could be arbitrary. Therefore, our approximating result captures the most robust part of the intuition.

Lemma 1 suggests that R's trade-off around β is largely governed by γ, too:

Proposition 3 (Institutional compatibility)**.** *At Stage 1, if $\gamma < \underline{\gamma} \equiv 1 - 1/(\pi - 2a - r)\underline{c}$, then R will prefer β to be as high as possible, i.e., $\beta^* = 1$; if $\gamma > \bar{\gamma} \equiv 1 - p/(\pi - a - r)\bar{c}$, then R will prefer β to be as low as possible, i.e., $\beta^* = 0$, where $\underline{\gamma} < \bar{\gamma} < 1$. Further, if $\pi > 2a + r + 1/\underline{c}$, then $\underline{\gamma} > 0$.*

Proof. The marginal impact of β on R's expected payoff in equilibrium at Stage 2 is

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS. \tag{12}$$

By Lemma 1, $\beta \in [0,1]$, and $S \in [p,1]$, we have

$$\begin{aligned} \frac{dV^R}{d\beta} &\geq \big(\pi - (1+\beta)a - r\big) \cdot \underline{c}a \cdot (1-\gamma) - aS \\ &\geq \big((\pi - 2a - r) \cdot \underline{c} \cdot (1-\gamma) - 1\big) \cdot a, \end{aligned} \tag{13}$$

so if

$$(\pi - 2a - r) \cdot \underline{c} \cdot (1-\gamma) - 1 > 0, \tag{14}$$

i.e.,

$$\gamma < 1 - \frac{1}{(\pi - 2a - r) \cdot \underline{c}} \equiv \underline{\gamma}, \tag{15}$$

then $dV^R/d\beta > 0$. At the same time, we have

$$\begin{aligned} \frac{dV^R}{d\beta} &\leq \big(\pi - (1+\beta)a - r\big) \cdot \bar{c}a \cdot (1-\gamma) - aS \\ &\leq \big((\pi - a - r) \cdot \bar{c} \cdot (1-\gamma) - p\big) \cdot a, \end{aligned} \tag{16}$$

so if

$$(\pi - a - r) \cdot \bar{c} \cdot (1-\gamma) - p < 0, \tag{17}$$

i.e.,

$$\gamma > 1 - \frac{p}{(\pi - a - r) \cdot \bar{c}} \equiv \bar{\gamma}, \tag{18}$$

then $dV^R/d\beta < 0$. Finally, note $\underline{\gamma} < \bar{\gamma} < 1$, and $\underline{\gamma} > 0$ is equivalent to $\pi > 2a + r + 1/\underline{c}$. The proposition is then proven. □

The intuition of Proposition 3 is as follows. When γ is sufficiently low, a higher β will increase the punishment P will face in case C's challenge fails, so the increase in the stability at Stage 2 will be significant; therefore, the political side of R's trade-off at Stage 1 will always be dominant; R then prefers the highest possible β. If γ is sufficiently high, the opposite will happen, and R will prefer the lowest possible β.

Proposition 3 explains the institutional compatibility within the power structure across Imperial China and Premodern Europe: as in European history, a high γ, which represents a weak absolute power of the Ruler, and a low β, which represents a highly unbalanced relationship between the Elites and People's rights and power, are compatible, while as in Chinese history, a low γ and a high β are compatible.

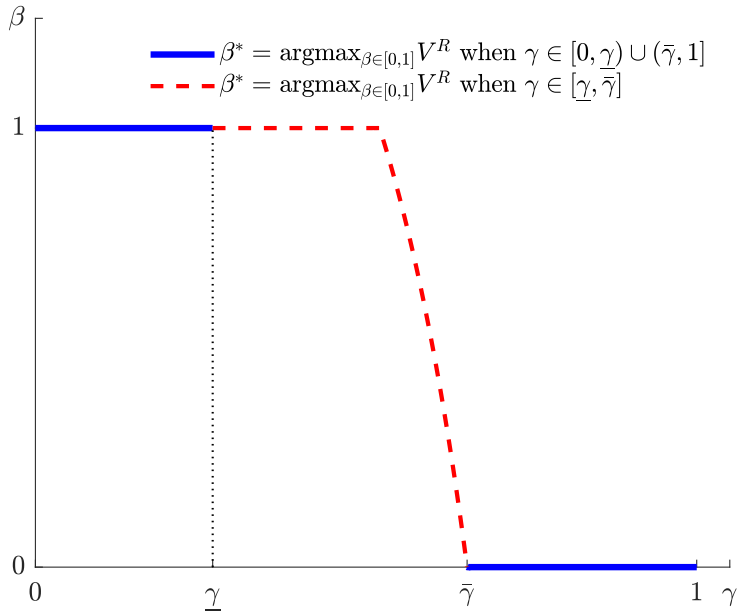
One may wonder why we did not show a result for $\gamma \in [\underline{\gamma}, \bar{\gamma}]$. It is not straightforward to derive such a result without further restrictions on the distribution of x. To see this point, observe that

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS \quad \text{and} \quad \frac{dS}{d\beta} = (1-p)pf\left(\hat{x}\right) \cdot a \cdot (1-\gamma). \tag{19}$$

A lower γ increases S, $1-\gamma$, and $\hat{x}$, but its impact on $f(\hat{x})$ depends on properties of $f(\cdot)$. Therefore, any unambiguous result about the impact of $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ on R's preference over β would rely on further restrictions on the distribution of x, which would have to be more or less arbitrary. As an example, Online Appendix B derives a result that R will generally prefer a higher β given a lower γ with an additional restriction on the distribution of x. For theoretical robustness, Proposition 3 only touches upon the extreme cases and, therefore, the first-order implications of γ. In light of all these, we provide a numerical example in Figure 2, where we plot R's choice β^* against γ: consistent with Proposition 3, $\beta^* = 1$ if $\gamma < \underline{\gamma}$, while $\beta^* = 0$ if $\gamma > \bar{\gamma}$; silent in Proposition 3, given the specification of the example, β^* weakly decreases with γ over $\gamma \in [\underline{\gamma}, \bar{\gamma}]$.

3.3 Extensions: Endogenizing the Absolute Power of the Ruler

In the analysis above we have taken the level of the absolute power of the Ruler γ as exogenous. Here we introduce two examples of extensions in which we endogenize γ and derive additional implications.



Specification: $F(x) = 1 - e^{-x}$, $p = 0.8$, $\pi = 20$, $a = 0.6$, $r = 5$. Under this specification, $\pi - 2a > r$. The Ruler's expected payoff in equilibrium at Stage 2 is denoted as V^R. The blue, solid line plots β^* when $\gamma \in [0, \underline{\gamma}) \cup (\bar{\gamma}, 1]$, which is consistent with Proposition 3. The red, dashed line plots β^* when $\gamma \in [\underline{\gamma}, \bar{\gamma}]$, about which Proposition 3 is silent.

Figure 2: Ruler's choice β^* a function of $\gamma \in [0, 1]$

People's perspective on the Ruler's absolute power. One may argue that γ would eventually depend on the legitimacy that P has granted to R in the first place. Along this argument, if before Stage 1 P has an opportunity to choose γ, how would her preference of γ look like?

Corollary 1. *If P could choose γ before Stage 1, then P would prefer any $\gamma < \underline{\gamma}$ over any $\gamma > \bar{\gamma}$.*

Proof. Given the β–γ power structure, P's expected payoff at Stage 2 is

$$\begin{aligned} V^P &= \gamma\beta a \cdot \big(1 - F(\hat{x})\big) \cdot p + \beta a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot p\Big) \\ &= \Big(1 - \big(1 - F(\hat{x})\big) \cdot p \cdot (1 - \gamma)\Big) \cdot \beta a. \end{aligned} \tag{20}$$

By Proposition 3, if $\gamma > \bar{\gamma}$, R will choose $\beta = 0$; if $\gamma < \underline{\gamma}$, R will choose $\beta = 1$. Therefore,

$$V^P\big|_{\gamma < \underline{\gamma}, \beta = 1} > 0 = V^P\big|_{\gamma > \bar{\gamma}, \beta = 0}. \tag{21}$$

The corollary is then proven. □

The intuition is as follows. On the equilibrium path at Stage 2, P will never side with

R when called upon. Therefore, she will receive either her status quo payoff βa or her post-punishment payoff $\gamma\beta a$. Given a sufficiently high $\gamma > \bar{\gamma}$, R will prefer the lowest possible $\beta = 0$ at Stage 1, so P will receive exactly a zero payoff eventually; any sufficiently low $\gamma < \underline{\gamma}$ will induce R to choose $\beta = 1$, granting P a strictly positive payoff eventually. P will then prefer any sufficiently low $\gamma < \underline{\gamma}$ over the sufficiently high $\gamma > \bar{\gamma}$ before Stage 1.

To clarify, we focus on the extreme case to highlight that it is not always the case that P will prefer a high to a low γ; instead, P may tolerate a quite absolutist R. We will come back to this insight in Section 4 when discussing the bureaucracy with the civil service exam in China.

Allowing current stability to shape future power structure. One may also argue that the European Rulers might have wanted to raise β but were not able to do so. Online Appendix D provides a response to this argument in several steps. First, it is easy to see that before Stage 2, R will prefer γ to be as low as possible, since a lower γ stabilizes the autocratic rule without sacrificing the status quo payoff, as seen in Equation (8).

Second, Proposition 2 implies that, if the total surplus π is sufficiently big, then the political side of R's trade-off with respect to β will be dominant, as long as the conditionality of the power and rights of the ruled exists, i.e., $\gamma < 1$. In that case, any R would prefer β to be as high as possible, as in Corollary D.1 in Online Appendix D.

Third, the last two results suggest that when the total surplus is sufficiently big, any R would like to invest in a lower γ and a higher β at the same time. Given this preference, we can consider an alternative setting in which Stage 2 gets played repeatedly over different periods and, instead of letting R choose β only once, we can justify a mechanical link from the current stability of autocratic rule in equilibrium to the future power structure, thereby endogenizing the absolute power of the Ruler in the future: the more stable R's autocratic rule is today, the more successful he would be in investing in the power structure toward the direction that he would favor, so the higher the degree of R's absolute power and the more symmetric the Elite–People relationship tomorrow. This effect on the future power structure, by Proposition 2, would eventually lead to a higher future stability, creating a dynamic complementarity.

Finally, given this dynamic complementarity, multiple stable steady states of (β, γ, S), i.e., the power structure and stability of autocratic rule, may exist, and, among these steady states, the stronger the absolute power of R and the more symmetric the Elite–People relationship, the higher the stability of autocratic rule, as derived in Proposition D.2 in Online Appendix D. A dual divergence of the power structure and stability of autocratic rule from slightly different initial conditions can thus appear, as shown in Proposition D.3 in Online

Appendix D. Here we summarize the implication as follows:

Corollary 2. *Compared with Premodern Europe, Imperial China could have been given a slightly lower γ, a slightly higher β, or a slightly higher S at very early times. This slight difference could have led the two societies to diverge into different stable steady states, where compared to Europe, China had a lower γ, a higher β, and a higher S.*

4 Further Discussion and Stylized Facts

4.1 Understanding Specific Institutions

Bureaucracy and civil service exam in China. Our model can help us understand specific institutions without explicitly modeling them in detail. One such example is the Chinese bureaucracy with the civil service exam, the hallmark of the Chinese imperial institutions (e.g., Finer, 1997a,b). Following our model, we can read it primarily as the Ruler raising β by generalizing the access to elite status between the Elites and People. By Proposition 3, Chinese Rulers had a great incentive to do so because they enjoyed a low γ, i.e., a strong absolute power. This is consistent with the fact that the civil service exam was first introduced during the Sui dynasty (581–619), when the absolute power of the Ruler had recovered from the low level during the Six Dynasties (220–589) (Yan, 2009). Given the bureaucratic system, the Elites became mainly bureaucrats who were appointed by the Ruler, so they became further reliant on the Ruler for legitimacy, making their power and rights more conditional on the Ruler's will, i.e., further lowering γ. Not only did the Ruler favor the stability of autocratic rule, i.e., a high S, as the result of the combination of a consolidated generalized access to elite status and a strong absolute power, i.e., a higher β and a low γ, but also by Corollary 1, the People might have been satisfied with the power structure and the resulting stability, without too much appetite for stronger rule of law or property rights.

Cities in Medieval Europe. In a similar vein, we can also read part of the rise of cities in Medieval Europe in relation to the Ruler's effort to raise β, by issuing charters that granted certain rights to the People in cities against other local Elites. This effort could eventually help stabilize the Ruler's autocratic rule. For example, "Philip [II of France] knew that in recognizing a commune, he was binding the citizens of that town to him. At critical moments in the reign the communes ... proved staunch military supporters. ... From the point of view of the communes ... the king was their natural ally, a counter to the main opponents of their

independence, the Church or the magnates" (Bradbury, 1998, p. 236).[25]

By Lemma 1 and Proposition 3, however, this stabilizing effect was not guaranteed when the Ruler's absolute power was as weak as it was in Medieval Europe. As a European Ruler was generally constrained by his own charters, he would find it difficult to punish the cities by retracting the granted rights. Because of this, granting more power and rights to cities might not help the Ruler much in creating a political alliance with urban commoners and securing his position. In this sense, when a Ruler in Europe freed a city from its feudal lords, he ran the risk of having freed it also from himself. Notable examples of this risk can be found during the rise of cities and boroughs in England and the free imperial cities in the Holy Roman Empire.[26]

The greater danger in the stabilizing effect under a less absolutist Ruler implies that this kind of Rulers could be more reluctant to promote the development of cities. This is consistent with the observation that within Premodern Europe, the urbanization rate increased on average in the states that were governed with a weak or no assembly, which indicated a stronger absolute power of the Ruler, whereas decreased in the long run in the states governed with a strong assembly, which indicated a weaker absolute power of the Ruler (Stasavage, 2020, p. 191–192). Given the uncertainty of the stabilizing effect under the generally weak absolute power of European Rulers, together with the dual divergence of the power structure and stability as in Corollary 2 and Online Appendix D, the European population that enjoyed cities' privileges was eventually relatively small at the eve of the modern times (Cantor, 1964; de Vries, 1984, p. 76). This reflected the limited success of the Ruler's effort to raise β in Premodern Europe.

4.2 Comparing Stability of Autocratic Rule

Proposition 2 states that a stronger absolute power of the Ruler and a more symmetric relationship between the Elites and People, as in Imperial China compared to Premodern

[25]Philip II's practice followed Louis VII, who "gave encouragement to the commune movement and received reciprocal support from the communities, at the expense of local lords" (Bradbury, 1998, p. 32). Relatedly, on the economic consequences of cities freeing peasants from local lords, see Cox and Figueroa (Forthcoming).

[26]In England, in May 1215, facing rebelling barons, John of England chartered the right of Londoners to elect their own mayor, together with other rights, "[i]n a last attempt to win the city" (Williams, 1963, p. 6). This proved futile: in June, still, "discontent citizens joined the barons in enforcing the signing of Magna Carta; the Mayor [of London] was the only commoner whose name appeared among the signatories" (Porter, 1994, p. 25–26). Magna Carta eventually extended the city rights by confirming in Article 13 that "the city of London [and] all other cities, boroughs, towns, and ports shall have all their liberties and free customs" (McKechnie, 1914, p. 241). Similarly, in the Holy Roman Empire, "the free towns had been winning valuable privileges in addition to those which they already possessed, and the wealthier among them, like Lübeck and Augsburg, were practically *imperia in imperio,* waging war and making peace, and ruling their people without any outside interference" from the Emperor (Holland, 1911, p. 342).

Europe, imply a higher stability of autocratic rule. As the nature of the challenge and the status quo of autocratic rule in our model are open to flexible interpretation, Proposition 2 can generate several auxiliary predictions that we could bring to data from historical China and Europe.

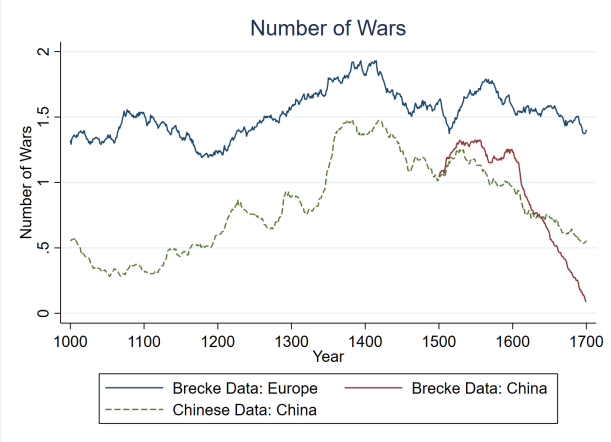
Number of wars. First, if we interpret the challenge in our model as an armed conflict, Proposition 2 then predicts that anyone in Europe who preferred an alternative to the status quo would be more willing to start a war than her counterpart in China. Note that this prediction does not depend on the challenger's identity and her status in the respective status quo: she could be either a foreign power, a rebellious local lord or regional governor, or a group of commoners. We also find it difficult to argue for a systematic difference in the number of all these possibly relevant entities between China and Europe in either way. Therefore, we should compare the total number of wars that challenged a status quo in the Chinese society with the number for the European society, regardless of the identity of the challengers. As defined in Section 2.1, here the "Chinese society" is the society in the historical core of Imperial China, whereas the "European society" is the Romano–Germanic influenced or assimilated society in western and central Europe where feudalism was once influential.

We are not aware of systematic evidence on this subject that covers the period of our interest. That said, Brecke (1999) provides comprehensive information on wars in Europe from 900 onwards and in China from only 1400 onwards. We complement the data with information from the Chinese Military History (2003) project from 900.[27] We further identify for each war whether it was fought to challenge a status quo in the Chinese or European society, respectively.

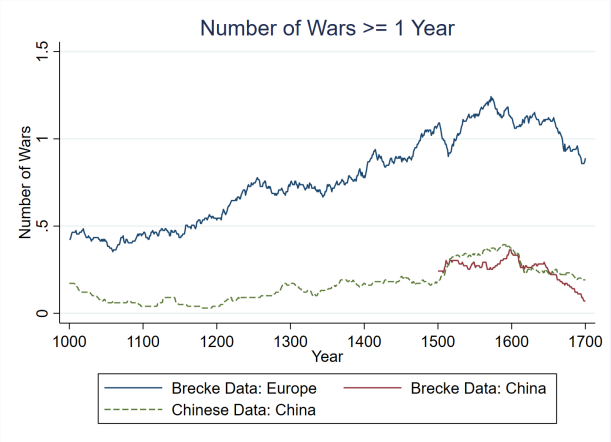
Figure 3 reports the number of wars breaking out in each given year that challenged a status quo in the Chinese or European society. For robustness, we plot the retrospective 100-year moving-averages; when doing so, for each retrospective 100-year window, we calculate the Olympic average in the window, i.e., taking the average in the window after removing one of the highest and one of the lowest values in the window. Besides reporting the result for wars of all lengths in Figure 3a, we also restrict our attention to more significant wars that lasted longer than one year, three years, and five years, respectively, in Figures 3b–3d.

Across Figures 3a–3d, we see the same pattern. First, Brecke (1999)'s data and our data

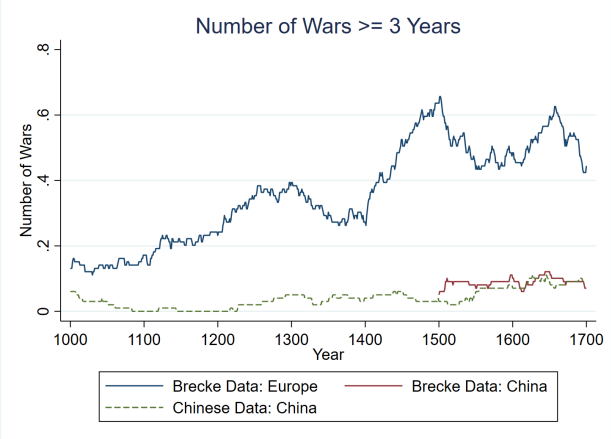
[27] The original data in the Chinese Military History (2003) project are at the level of individual battles. We first compare the battle-level data from the Chinese Military History (2003) project with the war-level data from Brecke (1999) to understand Brecke (1999)'s criteria of categorization. Complementing the criteria with information from Wu (2016) and Tian (2019), we finally manually categorize the battles recorded in the Chinese Military History (2003) project into wars.



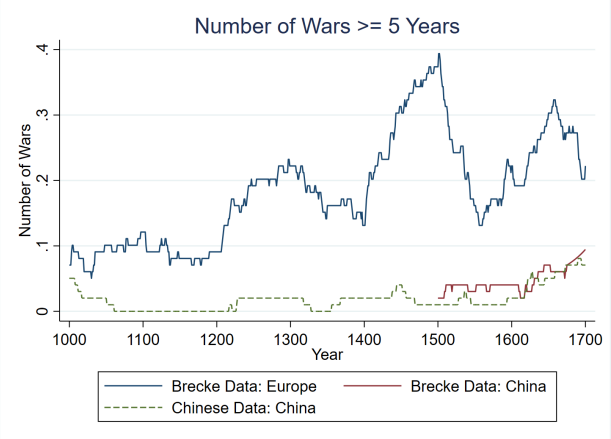
(a) All wars



(b) One-year or longer wars



(c) Three-year or longer wars



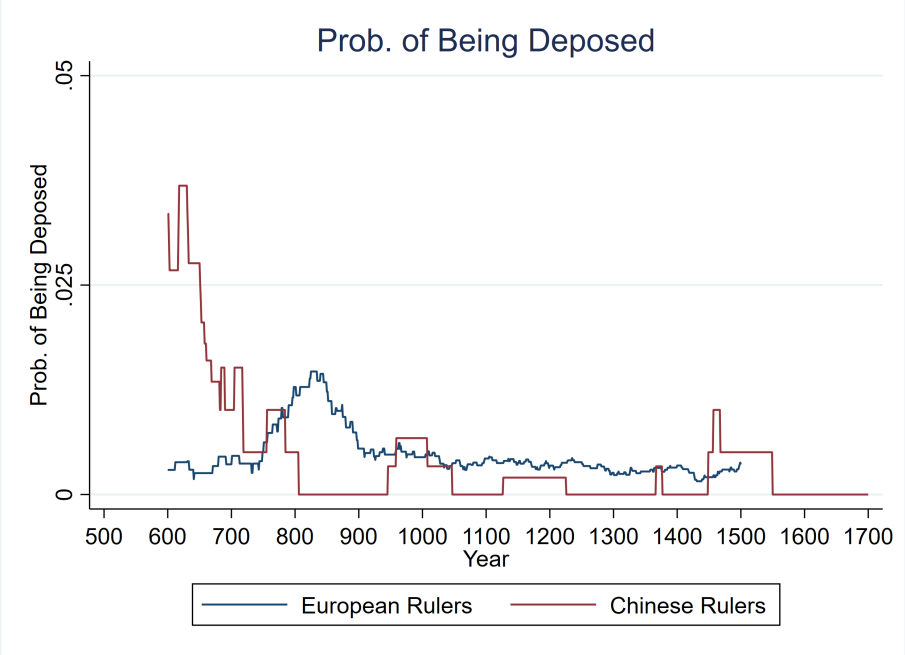
(d) Five-year or longer wars

Olympic average within each retrospective 100-year window. "Brecke Data: Europe" include wars in Brecke (1999) that challenged a status quo in the European society, i.e., the Romano–Germanic influenced or assimilated society in western and central Europe where feudalism was once influential; "Brecke Data: China" include wars in Brecke (1999) that challenged a status quo in the Chinese society, i.e., the society in the historical core of Imperial China; "Chinese Data: China" include wars in our Chinese war data that challenged a status quo in the Chinese society. For more details of our Chinese war data, see Footnote 27.

Figure 3: Number of wars starting in a given year that challenged a status quo in the Chinese or European society

give comparable numbers of wars that challenged a status quo in the Chinese society during 1400–1700, strengthening our confidence about our data. Second, the figures show that the number of wars for Europe was consistently higher than that for China from 900 to 1700. We thus conclude that during 900–1700, there were significantly more wars challenging a status quo in Europe than those challenging a status quo in China, consistent with Proposition 2.

Risk of deposition. Second, if we interpret the challenge in our model as to remove the Ruler from the ruling position, Proposition 2 then predicts that a Chinese Ruler should have faced a lower risk of deposition in each given year than a European Ruler. On the data, the historical information of all monarchies in the world has been compiled by Morby (1989) and some of it has been used in a few studies (e.g., Blaydes and Chaney, 2013; Kokkonen and Sundell, 2014). Using the same data, to compare the risk of deposition between China and Europe, we first calculate for each given year a measure of the risk of deposition in that year, i.e., the share of the Rulers who were deposed in that year among all the Rulers who had been in power in that year; we then visualize in Figure 4 the comparison between China and Europe by plotting the retrospective 100-year moving-averages of the measure. For robustness, again, the Olympic average is used.



Olympic average within each retrospective 100-year window. Following Blaydes and Chaney (2013), "European Rulers" include all the ones who assumed power before 1500 and are under the section "The Barbarian West" or the subsections "The British Isles," "France," "The Low Countries," "Italy," "The Iberian Peninsula," "The German-speaking States," "Scandinavia," and "Crusader States" under the section "Europe" in Morby (1989). "Chinese Rulers" include all the ones under the subsection "China" under the section "The Far East" in Morby (1989).

Figure 4: Risk of deposition to a Ruler in a given year, China vs. Europe

As shown in Figure 4, the risk of deposition to a Ruler in a given year was generally lower in China than in Europe during the 9–14th centuries, when the two differences we emphasize in the power structure between the two societies were the most prominent. That said, there existed a short period around 970 where the risk in China appeared to be higher, when

China entered the Five Dynasties and Ten Kingdoms period. In light of this, we conduct a Kolmogorov–Smirnov test to check whether the differences in the risks between China and Europe are systematic. The test reports that at a significance level of 0.1%, we can accept the claim that the risk of deposition to a Ruler in a given year was generally lower in China than in Europe during 800–1400, whereas the opposite claim must be rejected. These results are consistent with Proposition 2.[28]

Resilience of unified autocratic rule. Finally, if we interpret the status quo of autocratic rule in our model as a unified one across the territory, Proposition 2 then predicts that a unified autocratic rule should have been more resilient in China than in Europe. As discussed, the literature has well documented that China had been more unified than Europe in history. Among many other measures, here we present in Figure 5 only a replication of the comparison by Scheidel (2019, fig. 1.11) as an example, plotting the share of the population in the continent Europe that was controlled by the largest polity in the continent, together with the same measure for East Asia, where China is located.

As shown in the figure, since 800, in East Asia, the population share in the largest polity, which was the dominant empire in China, had usually been above 75%, except for short subperiods of turbulence. In contrast, the number for Europe had been below 20%, consistent with a more fragmented pattern. This comparison is consistent with Proposition 2. Therefore, our model provides a power-structure approach to the unification–fragmentation cleavage between Imperial China and Premodern Europe.

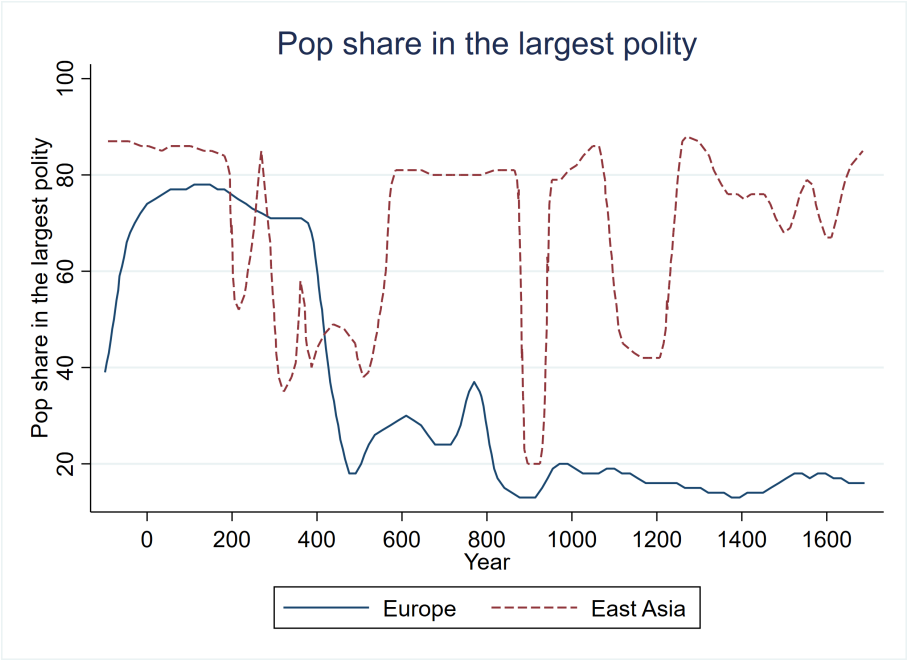
In sum, consistent with our model, we find a persistent difference in the stability of autocratic rule, whether measured by the number of wars, the Ruler's risk of deposition, or the resilience or vulnerability of unified autocratic rule.[29]

5 Conclusion

In this paper, we provide a power structure framework to reconcile a series of views on the institutional differences between Imperial China and Europe that are seemingly contradictory

[28] One may recall that Hoffman (2015) shows that during the 16–18th centuries, major European sovereigns were seldom deposed after losing a war. On the 7–14th centuries, when our characterization of the power structure was more relevant, however, Eisner (2011) shows that the risk of regicide, which would surely lead to but was not the only way to deposition, had remained high in Europe. Eisner (2011) also shows that at that time battle death was a major risk to European rulers and being murdered with an external power involved was also not rare.

[29] To clarify, the dual divergence of the power structure and autocratic stability depicted in Corollary 2 and Online Appendix D could have already taken place around 800. If so, we would not predict an increase in stability differences but only the persistence of such differences.



Replicated from Scheidel (2019, fig. 1.11). According to Scheidel (2019, fig. 1.1), "Europe" is defined primarily by the common geographical term, i.e., the continent Europe, whereas "East Asia" includes nowadays China, Mongolia, North Korea, South Korea, and Japan.

Figure 5: Percentage of population claimed by the largest polity, Europe vs. East Asia

in the light of the literature on institutions and development. In this framework, we read the institutional differences along two dimensions of the power structure: Chinese Rulers had a stronger absolute power, while the relationship between the Elites and People in terms of their power and rights was more asymmetric in Europe.

By building a model and analyzing how the power structure can shape the stability of an autocratic rule, we show that, once we recognize that the Ruler's absolute power is about the conditionality of the power and rights of the ruled on the Ruler's will, a more symmetric Elite–People relationship will strengthen the political alliance between the Ruler and the People, thus creating more loyalty to the Ruler, deterring potential challenges, and stabilizing the autocratic rule. Importantly, this effect and, therefore, the Ruler's incentive to promote a more symmetric Elite–People relationship depend on the Ruler's absolute power. This suggests that an absolutist Ruler can be compatible with a more symmetric Elite–People relationship. A society can thus be repressive in one institutional dimension but inclusive in another at the same time, a new result to the literature.

This comparative institutional theory explains the coexistence of the two power structure differences between Imperial China and Premodern Europe. Besides guiding us to understand specific institutions, our model also suggests a higher stability of autocratic rule in

Imperial China. This implication is supported by stylized facts about the number of wars, risk of deposition, and resilience of unified autocratic rule.

Admittedly, our theory is highly stylized as we capture the power structure with only two parameters, and we only examine the stability of autocratic rule as the outcome of the power structure. The benefit of doing so is that we can deliver our key insights in a simple manner. That said, our framework can be applied to understand other political, economic, and social outcomes. For example, the Chinese Legalist tradition had emphasized the absolute power of the Ruler, while the apparent dominance of Confucianism in China had been reflected by the institution and content of the civil service exam, a key component of the more balanced Elite–People relationship. Our Proposition 2 thus explains why the Chinese Rulers had promoted the Confucianism–Legalism confluence as the dominant political culture (e.g., Qin, 1998; Yan, 2004; Zhao, 2015); Proposition 3 sees Legalism as the more fundamental side within the confluence; Corollaries 1 and 2 explain why such a culture may have been accepted by the People and persistent over time, respectively.

There can also be more insights to gain if one applies our power structure framework to other parts of the world. For example, Blaydes and Chaney (2013) show that Christian kings in Western Europe enjoyed a higher political stability than Muslim sultans during the 9–15th centuries. This difference can be explained in our framework: lords in feudal Europe owned land and military forces on a regular basis, suggesting a high status quo payoff a to the Elites, while Mamlukism in the Muslim world was designed to remove elite Mamluks "from the luxuries of settled life" (Blaydes and Chaney, 2013, p. 23), suggesting a low a; one can show in our model that a higher a would increase the stability of autocratic rule. We thus hope that our study opens new avenues for future research.

References

Acemoğlu, Kamer Daron, Simon Johnson, and James Alan Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* **91**: 1369–1401.

Acemoğlu, Kamer Daron, Simon Johnson, and James Alan Robinson. 2005a. Institutions as a fundamental cause of long-run growth. *Handbook of economic growth* **1**: 385–472.

Acemoğlu, Kamer Daron, Simon Johnson, and James Alan Robinson. 2005b. The rise of Europe: Atlantic trade, institutional change, and economic growth. *American Economic Review* **95**: 546–579.

Acemoğlu, Kamer Daron, and James Alan Robinson. 2000. Why did the west extend the

franchise? democracy, inequality, and growth in historical perspective. *Quarterly Journal of Economics* **115**: 1167–1199.

Acemoğlu, Kamer Daron, and James Alan Robinson. 2001. A theory of political transitions. *American Economic Review* **91**: 938–963.

Acemoğlu, Kamer Daron, and James Alan Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. New York: Crown Publishing Group.

Acemoğlu, Kamer Daron, and James Alan Robinson. 2019. *The Narrow Corridor: States, Societies, and the Fate of Liberty*. New York: Penguin Books.

Acemoğlu, Kamer Daron, Thierry Verdier, and James Alan Robinson. 2004. Kleptocracy and divide-and-rule: A model of personal rule. *Journal of the European Economic Association* **2**: 162–192.

Bai, Ying, and Ruixue Jia. 2016. Elite recruitment and political stability: The impact of the abolition of China's civil service exam. *Econometrica* **84**: 677–733.

Barrow, Julia Steuart. 2015. *The Clergy in the Medieval World: Secular Clerics, Their Families and Careers in North-Western Europe c.800–c.1200*. Cambrdige: Cambridge University Press.

Beckett, John Vincent. 1984. The pattern of landownership in England and Wales, 1660–1880. *Economic History Review* **37**: 1–22.

Bennett, Michael John. 1999. *Richard II and the Revolution of 1399*. Stroud: Sutton.

Bertocchi, Graziella. 2006. The law of primogeniture and the transition from landed aristocracy to industrial democracy. *Journal of Economic Growth* **11**: 43–70.

Besley, Timothy John, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton: Princeton University Press.

Besley, Timothy John, and Torsten Persson. 2014. The causes and consequences of development clusters: State capacity, peace, and income. *Annual Review of Economics* **6**: 927–949.

Blaydes, Lisa, and Eric Chaney. 2013. The feudal revolution and Europe's rise: Political divergence of the Christian West and the Muslim world before 1500 CE. *American Political Science Review* **107**: 16–34.

Bloch, Marc. 1962a. *Feudal Society, Volume I: The Growth of Ties of Dependence*. London: Routledge, 2nd edition. Translated by L. A. Manyon from *La Société Féodale, Volume I: La Formation des Liens de Dépendance,* 1939, Paris: Éditions Albin Michel. Reprint, 2004.

Bloch, Marc. 1962b. *Feudal Society, Volume II: Social Classes and Political Organization.* London: Routledge, 2nd edition. Translated by L. A. Manyon from *La Société Féodale, Volume II: Les Classes et le Gouvernement des Hommes,* 1940, Paris: Éditions Albin Michel. Reprint, 2005.

Bracton, Henry de. 1968. *Bracton on the Laws and Customs of England.* Cambridge: Belknap Press. Translated with annotations by Samuel E. Thorne from *Bracton De Legibus et Consuetudinibus Angliæ*, George E. Woodbine (ed.), 1915, New Haven: Yale University Press.

Bradbury, Jim. 1998. *Philip Augustus: King of France 1180–1223.* London: Routledge. Reprint, 2013.

Brecke, Peter. 1999. Violent conflicts 1400 AD to the present in different regions of the world. Paper prepared for the 33rd Peace Science Society (International) North American Meeting, Ann Arbor, October 8–10, 1999.

Brook, Timothy. 2010. *The Troubled Empire: China in the Yuan and Ming Dynasties.* Cambridge: Belknap Press.

Cantor, Norman Frank. 1964. *The Civilization of the Middle Ages.* New York: Harper Perennial. Reprint, 1994.

Chao, Kang, and Chung-yi Chen. 1982. *A History of China's Land Institutions.* Taipei: Linking Publishing. Reprint, 2006, Beijing: New Star Publisher.

Chinese Military History. 2003. *Chronicle of Wars in Chinese History.* Beijing: People's Liberation Army Publishing House.

Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton: Princeton University Press.

Cox, Gary Walter, and Valentín Figueroa. Forthcoming. Political fragmentation, rural-to-urban migration and urban growth patterns in western Eurasia, 800–1800. *European Review of Economic History* .

Cox, Gary Walter, Douglass Cecil North, and Barry Robert Weingast. 2019. The violence trap: A political-economic approach to the problems of development. *Journal of Public Finance and Public Choice* **34**: 3–19.

de Groot, Jan Jakob Maria. 1903. *Sectarianism and Religious Persecution in China: A Page in the History of Religions, Volume I.* Amsterdam: Johannes Müller.

de Vries, Jan. 1984. *European Urbanization, 1500–1800.* Cambridge: Harvard University Press.

Dillon, Michael O. 1998. Dynastic cycle. In Dillon, Michael O. (ed.) *China: A Historical and Cultural Dictionary.* London: Curzon Press, 87.

Dincecco, Mark, and Yuhua Wang. 2018. Violent conflict and political development over the long run: China versus Europe. *Annual Review of Political Science* **21**: 341–358.

Ebrey, Patricia Buckley, and Anne Walthall. 2013. *East Asia: A Cultural, Social, and Political History*. Boston: Cengage Learning, 3rd edition.

Eisner, Manuel. 2011. Killing kings: Patterns of regicide in Europe, AD 600–1800. *British Journal of Criminology* **51**: 556–577.

Esherick, Joseph W. 1981. A note on land distribution in prerevolutionary China. *Modern China* **7**: 387–411.

Fernández-Villaverde, Jesús. 2016. Magna Carta, the rule of law, and the limits on government. *International Review of Law and Economics* **47**: 22–28.

Fernández-Villaverde, Jesús, Mark Koyama, Youhong Lin, and Tuan-Hwee Sng. 2020. The fractured-land hypothesis. National Bureau of Economic Research Working Paper No. 27774.

Finer, Samuel Edward. 1997a. *The History of Government from the Earliest Times, Volume I: Ancient Monarchies and Empires*. New York: Oxford University Press.

Finer, Samuel Edward. 1997b. *The History of Government from the Earliest Times, Volume II: The Intermediate Ages*. New York: Oxford University Press.

Fukuyama, Francis. 2011. *The Origins of Political Order: From Prehuman Times to the French Revolution*. New York: Farrar, Straus and Giroux.

Ganshof, François-Louis. 1952. *Feudalism*. London: Longmans, Green and Co. Translated by Philip Grierson from *Qu'est-ce que la Féodalité?* 1947, Bruxelles: Office de Publicité, S.C., 2nd edition.

Gennaioli, Nicola, and Hans-Joachim Voth. 2015. State capacity and military conflict. *The Review of Economic Studies* **82**: 1409–1448.

Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World: Population Change and State Breakdown in England, France, Turkey, and China, 1600–1850*. Berkeley: University of California Press.

Goody, Jack, Joan Thirsk, and Edward Palmer Thompson (eds.) . 1976. *Family and Inheritance: Rural Society in Western Europe, 1200–1800*. Cambridge: Cambridge University Press.

Greif, Avner, Joel Mokyr, and Guido Tabellini. Forthcoming. *Two Paths to Prosperity: Culture and Institutions in Europe and China, 1200–2000*.

Greif, Avner, and Guido Tabellini. 2010. Cultural and institutional bifurcation: China and Europe compared. *American Economic Review* **100**: 135–140.

Greif, Avner, and Guido Tabellini. 2017. The clan and the corporation: Sustaining cooperation in China and Europe. *Journal of Comparative Economics* **45**: 1–35.

Grummitt, David. 2014. *A Short History of the Wars of the Roses*. London: Bloomsbury Publishing.

Hao, Yu, and Gregory Clark. 2012. Social mobility in China, 1645–2012: A surname study. Working paper, University of California, Davis.

Heldring, Leander, James Alan Robinson, and Sebastian Vollmer. 2021. The long-run impact of the Dissolution of the English monasteries. *Quarterly Journal of Economics* .

Herlihy, David. 1973. Three patterns of social mobility in medieval history. *Journal of Interdisciplinary History* **3**: 623–647.

Ho, Ping-ti. 1959. Aspects of social mobility in China, 1368–1911. *Comparative Studies in Society and History* **1**: 330–359.

Hoffman, Philip T. 2015. *Why Did Europe Conquer the World?* Princeton: Princeton University Press.

Holland, Arthur William. 1911. Imperial cities or towns. In Chisholm, Hugh (ed.) *Encyclopædia Britannica*, volume 14. Cambridge: Cambridge University Press, 11th edition, 342.

Huang, Yasheng, and Clair Zhuqing Yang. 2021. The great political divergence. Working paper.

Jiang, Yonglin. 2011. *The Mandate of Heaven and the Great Ming Code*. Seattle: University of Washington Press.

Johnson, Noel D., and Mark Koyama. 2019. *Persecution and Toleration: The Long Road to Religious Freedom*. New York: Cambridge University Press.

Jones, Eric. 1981. *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*. Cambridge: Cambridge University Press.

Kemp, Brian Richard. 1970. Hereditary benefices in the medieval English Church: A Herefordshire example. *Historical Research* **43**: 1–15.

Ko, Chiu Yu, Mark Koyama, and Tuan-Hwee Sng. 2018. Unified China and divided Europe. *International Economic Review* **59**: 285–327.

Kokkonen, Andrej, and Anders Sundell. 2014. Delivering stability – Primogeniture and autocratic survival in European monarchies 1000–1800. *American Political Science Review* **108**: 438–453.

Lander, Jack Robert. 1961. Attainder and forfeiture, 1453 to 1509. *Historical Journal* **4**: 119–151.

Levenson, Joseph Richmond. 1965. *Confucian China and Its Modern Fate, Volume 2: The Problem of Monarchial Decay*. Berkeley: University of California Press.

Levi, Margaret. 1988. *Of Rule and Revenue*. Los Angeles: University of California Press.

Lü, Simian. 1944. *A General History of China, Volume 2*. Shanghai: Kai Ming Bookstore. Reprint, 1992, Shanghai: East China Normal University Press.

Ma, Debin, and Jared Rubin. 2019. The paradox of power: Principal-agent problems and administrative capacity in Imperial China (and other absolutist regimes). *Journal of Comparative Economics* **47**: 277–294.

Mann, Michael. 1986. *The Social Sources of Power, Volume 1: A History of Power from the Beginning to AD 1760*. Cambridge: Cambridge University Press.

McCollim, Gary B. 2012. *Louis XIV's Assault on Privilege: Nicolas Desmaretz and the Tax on Wealth*. Rochester: University Rochester Press.

McKechnie, William Sharp. 1914. *Magna Carta: A Commentary on the Great Charter of King John*. Glasgow: James MacLehose and Sons, 2nd edition.

McNeill, William Hardy. 1982. *The Pursuit of Power: Technology, Armed Force, and Society Since AD 1000*. University of Chicago Press.

Mokyr, Joel. 2016. *A Culture of Growth: The Origins of the Modern Economy*. Princeton: Princeton University Press.

Morby, John Edwin. 1989. *Dynasties of the World: A Chronological and Genealogical Handbook*. New York: Oxford University Press.

Morton, Arthur Leslie. 1938. *A People's History of England*. London: Lawrence and Wishart. Reprint, 1999.

North, Douglass Cecil. 1989. Institutions and economic growth: An historical introduction. *World Development* **17**: 1319–1332.

North, Douglass Cecil, John Joseph Wallis, and Barry Robert Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.

North, Douglass Cecil, and Barry Robert Weingast. 1989. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* **49**: 803–832.

Orwell, George. 1947. *The English People*. London: Collins.

Padró i Miquel, Gerard. 2007. The control of politicians in divided societies: The politics of fear. *Review of Economic Studies* **74**: 1259–1274.

Parish, Helen. 2010. *Clerical Celibacy in the West: C.1100–1700*. Farnham: Ashgate Publishing Limited.

Perry, Elizabeth Jean. 2008. Chinese conceptions of "right": From Mencius to Mao – and now. *Perspectives on Politics* **6**: 37–50.

Persico, Nicola. 2021. A theory of non-democratic redistribution and public good provision. Working paper.

Porter, Roy Sydney. 1994. *London: A Social History*. London: Hamish Hamilton.

Qian, Jiwei, and Tuan-Hwee Sng. 2021. The state in Chinese economic history. Working paper, National University of Singapore.

Qin, Hui. 1998. The "large-community standard" and traditional Chinese society: Part I. *Sociological Studies* **13**: 12–21.

Reynolds, Susan. 2010. *Before Eminent Domain: Toward a History of Expropriation of Land for the Common Good*. Chapel Hill: University of North Carolina Press.

Roland, Gérard. 2020. The deep historical roots of modern culture: A comparative perspective. *Journal of Comparative Economics* **48**: 483–508.

Rosenthal, Jean-Laurent, and Roy Bin Wong. 2011. *Before and Beyond Divergence: The Politics of Economic Change in China and Europe*. Cambridge: Harvard University Press.

Scheidel, Walter. 2019. *Escape from Rome: The Failure of Empire and the Road to Prosperity*. Princeton: Princeton University Press.

Seward, Desmond. 1978. *The Hundred Years War: The English in France 1337–1453*. London: Penguin. Reprint, 1999.

Skinner, George William. 1985. Presidential address: The structure of Chinese history. *Journal of Asian Studies* **44**: 271–292.

Sng, Tuan-Hwee. 2014. Size and dynastic decline: The principal-agent problem in late imperial China, 1700–1850. *Explorations in Economic History* **54**: 107–127.

Southern, Richard William. 1970. *Western Society and the Church in the Middle Ages*. Baltimore: Penguin Books.

Stasavage, David. 2016. Representation and consent: Why they arose in Europe and not elsewhere. *Annual Review of Political Science* **19**: 145–162.

Stasavage, David. 2020. *The Decline and Rise of Democracy: A Global History from Antiquity to Today*. Princeton: Princeton University Press.

Tackett, Nicolas. 2014. *The Destruction of the Medieval Chinese Aristocracy*. Cambridge: Harvard University Asia Center.

Tamanaha, Brian Z. 2004. *On the Rule of Law: History, Politics, Theory*. Cambridge: Cambridge University Press.

Tian, Zhaolin. 2019. *History of Warfare in China*. Nanjing: Jiangsu People's Publishing.

Tuchman, Barbara Wertheim. 1978. *A Distant Mirror: The Calamitous 14th Century*. New York: Alfred A. Knopf.

Turchin, Peter. 2009. A theory for formation of large empires. *Journal of Global History* **4**: 191–217.

Unger, Roberto Mangabeira. 1977. *Law in Modern Society: Toward a Criticism of Social Theory*. New York: The Free Press.

Usher, Dan. 1989. The dynastic cycle and the stationary state. *American Economic Review* **79**: 1031–1044.

Vincent, Nicholas. 2012. *Magna Carta: A Very Short Introduction*. Oxford: Oxford University Press.

von Glahn, Richard. 2016. *An Economic History of China: From Antiquity to the Nineteenth Century*. Cambridge: Cambridge University Press.

Wang, Jiafan. 2000. *A General History of China*. Shanghai: East China Normal University Press.

Watson, Burton DeWitt. 1964. *Han Fei Tzu: Basic Writings*. New York: Columbia University Press.

Wickham, Chris. 2009. *The Inheritance of Rome: A History of Europe from 400 to 1000*. London: Penguin.

Williams, Gwyn Alfred. 1963. *Medieval London: From Commune to Capital*. London: Athlone Press.

Wittfogel, Karl August. 1957. *Oriental Despotism: A Comparative Study of Total Power*. New Haven: Yale University Press.

Wood, Charles Tuttle. 1975. The deposition of Edward V. *Traditio* **31**: 247–286.

Wu, Guoqing. 2016. *History of Warfare in China*. Beijing: People's Press.

Xue, Melanie Meng. 2021. Autocratic rule and social capital: Evidence from imperial China. Working paper.

Yan, Buke. 2004. Bureaucratic political institutions during the beginning of the imperial era – Qin and Han dynasties. In Wu, Zongguo (ed.) *Studies on the Bureaucratic Political Institutions in Ancient China*. Beijing: Peking University Press, 19–85.

Yan, Buke. 2009. *Peaks and Troughs: Political Civilization from the Qin Dynasty to the Southern and Northern Dynasties*. Beijing: Peking University Press.

Zhang, Taisu. 2017. *The Laws and Economics of Confucianism: Kinship and Property in Preindustrial China and England*. Cambridge: Cambridge University Press.

Zhao, Dingxin. 2009. The Mandate of Heaven and performance legitimation in historical and contemporary China. *American Behavioral Scientist* **53**: 416–433.

Zhao, Dingxin. 2015. *The Confucian-Legalist State: A New Theory of Chinese History*. New York: Oxford University Press.

Online Appendix

A Allowing for Mixed Strategies

In this section we allow for mixed strategies at Stage 2 by dropping the earlier assumption that E and P will side with R when indifferent and C will not challenge when indifferent. We then characterize all the subgame perfect equilibria that are empirically relevant (possible for the status quo to end) and can involve mixed strategies at a strictly positive share of the states of the world. We then examine whether the main insights from the main text would maintain.

When doing so, we adopt a few additional assumptions without losing much generality. We consider only the empirically relevant, nontrivial case $\gamma < 1$. We also assume that x is a continuous random variable so that its distribution does not have any mass point, and that $F(ap) < 1$ so that $1 - F(\hat{x}) > 0$ always holds.

By backward induction, in any subgame perfect equilibrium at Stage 2, P will side with R when $x < \hat{x}$ and not side with R when $x > \hat{x}$.

Taking this into consideration, in any subgame perfect equilibrium, E will side with R when $x < \hat{x}$; when $x > \hat{x}$, E will side with R if

$$w < \frac{p}{1-p} \cdot (1-\gamma)a; \tag{22}$$

E will not side with R if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a; \tag{23}$$

E will side with R with probability $q_E(x)$ if

$$w = \frac{p}{1-p} \cdot (1-\gamma)a, \tag{24}$$

where $q_E(x)$ is a function and satisfies $q_E(x) \in [0,1]$ for any $x > \hat{x}$.

Taking this into consideration, in any subgame perfect equilibrium, C will not challenge when $x < \hat{x}$. When $x > \hat{x}$, C will not challenge if

$$w < \frac{p}{1-p} \cdot (1-\gamma)a; \tag{25}$$

C will also not challenge if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z < \frac{p}{1-p} \cdot y; \tag{26}$$

C will challenge if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z > \frac{p}{1-p} \cdot y; \tag{27}$$

C will challenge with probability $q_C(x)$ if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z = \frac{p}{1-p} \cdot y, \tag{28}$$

where $q_C(x)$ is a function and satisfies $q_C(x) \in [0,1]$ for any $x > \hat{x}$; if

$$w = \frac{p}{1-p} \cdot (1-\gamma)a, \tag{29}$$

however, C will compare

$$0 \text{ vs. } q_E(x) \cdot (-y) + (1 - q_E(x)) \cdot \left((1-p) \cdot z - p \cdot y\right), \tag{30}$$

i.e.,

$$0 \text{ vs. } (1 - q_E(x))(1-p) \cdot z - \left(1 - (1 - q_E(x))(1-p)\right) \cdot y, \tag{31}$$

so C will challenge with probability $q_C(x)$, where, for any $x > \hat{x}$, $q_C(x) = 1$ if

$$z > \frac{1 - (1 - q_E(x))(1-p)}{(1 - q_E(x))(1-p)} \cdot y, \tag{32}$$

$q_C(x) = 0$ if

$$z < \frac{1 - (1 - q_E(x))(1-p)}{(1 - q_E(x))(1-p)} \cdot y, \tag{33}$$

and $q_C(x) \in [0,1]$ if

$$z = \frac{1 - (1 - q_E(x))(1-p)}{(1 - q_E(x))(1-p)} \cdot y. \tag{34}$$

We have then specified all equilibrium strategies at any $x \neq \hat{x}$. Therefore, at Stage 2, the only families of subgame perfect equilibria that are empirically relevant and can involve mixed strategies at a strictly positive share of the states of the world are:

- When $w > \frac{p}{1-p} \cdot (1-\gamma)a$ and $z = \frac{p}{1-p} \cdot y$, in any subgame perfect equilibrium, if $x < \hat{x}$, then C will not challenge, E would side with R, and P would side with R; if $x > \hat{x}$, then C will challenge with probability $q_C(x) \in [0,1]$, E will not side with R, and P will not side with R.

  In any equilibrium of this family, the probability of challenge is $\int_{\hat{x}}^{\infty} q_C(x) dF(x)$, while

the survival probability of the status quo is

$$S = 1 - \int_{\hat{x}}^{\infty} q_C(x) dF(x) \cdot (1-p). \tag{35}$$

All impacts of γ and β on political stability still come from their impacts on $\hat{x}$. All main insights from the main text would then remain.

- When $w = \frac{p}{1-p} \cdot (1-\gamma)a$, in any subgame perfect equilibrium, if $x < \hat{x}$, then C will not challenge, E would side with R, and P would side with R; if $x > \hat{x}$, then C will challenge with probability $q_C(x)$, where $q_C(x)$ depends on

$$z \text{ vs. } \frac{1-(1-q_E(x))(1-p)}{(1-q_E(x))(1-p)} \cdot y, \tag{36}$$

E will side with R with probability $q_E(x)$, and P will not side with R.

In any equilibrium of this family, the probability of challenge is $\int_{\hat{x}}^{\infty} q_C(x) dF(x)$, while the survival probability of the status quo is

$$S = 1 - \int_{\hat{x}}^{\infty} q_C(x)(1-q_E(x))(1-p) dF(x). \tag{37}$$

Still, all impacts of γ and β on political stability come from their impacts on $\hat{x}$. All main insights from the main text would then remain.

B Institutional Compatibility under Additional Restriction on the Distribution of the State of the World

Proposition B.1. *If the distribution of x satisfies*

$$\epsilon \equiv -\frac{x \cdot f'(x)}{f(x)} \leq \bar{\epsilon} \equiv 1 - \frac{a}{\pi - 2a - r} \tag{38}$$

over $x \in [0, pa]$, then a lower $\gamma \in [0,1]$ would make R prefer a higher $\beta \in [0,1]$.

Proof. Observe that

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS, \quad \frac{dS}{d\beta} = (1-p)pf(\hat{x}) \cdot a \cdot (1-\gamma), \tag{39}$$

and

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p). \tag{40}$$

Therefore,

$$\begin{aligned}
\frac{\partial^2 V^R}{\partial\gamma\partial\beta} &= \big(\pi-(1+\beta)a-r\big)\cdot\frac{\partial S}{\partial\gamma\partial\beta}-a\cdot\frac{dS}{d\gamma}\\
&= -\big(\pi-(1+\beta)a-r\big)\cdot(1-p)pa\cdot\big((1-\gamma)f'(\hat{x})\cdot pa\beta+f(\hat{x})\big)-a\cdot\frac{dS}{d\gamma}\\
&= -\big(\pi-(1+\beta)a-r\big)\cdot(1-p)pa\cdot\big(f'(\hat{x})\cdot\hat{x}+f(\hat{x})\big)+a\cdot(1-p)f(\hat{x})p\beta a\\
&= -(1-p)pa\cdot\Big(\big(\pi-(1+\beta)a-r\big)\cdot\big(f'(\hat{x})\cdot\hat{x}+f(\hat{x})\big)-f(\hat{x})\beta a\Big)\\
&= -(1-p)pa\cdot\Big(\big(\pi-(1+\beta)a-r\big)\cdot f'(\hat{x})\cdot\hat{x}+\big(\pi-(1+2\beta)a-r\big)\cdot f(\hat{x})\Big). \quad (41)
\end{aligned}$$

Therefore, $\partial^2 V^R/\partial\gamma\partial\beta \le 0$ if and only if

$$\big(\pi-(1+\beta)a-r\big)\cdot f'(\hat{x})\cdot\hat{x}+\big(\pi-(1+2\beta)a-r\big)\cdot f(\hat{x})\ge 0, \quad (42)$$

i.e.,

$$\epsilon\equiv-\frac{f'(\hat{x})\cdot\hat{x}}{f(\hat{x})}\le\frac{\pi-(1+2\beta)a-r}{\pi-(1+\beta)a-r}=1-\frac{\beta a}{\pi-(1+\beta)a-r}. \quad (43)$$

Since

$$\frac{\beta a}{\pi-(1+\beta)a-r}\in\left[0,\frac{a}{\pi-2a-r}\right], \quad (44)$$

we have

$$1-\frac{\beta a}{\pi-(1+\beta)a-r}\in\left[1-\frac{a}{\pi-2a-r},1\right]. \quad (45)$$

Therefore, $\partial^2 V^R/\partial\gamma\partial\beta \le 0$ can be guaranteed by

$$\epsilon\le 1-\frac{a}{\pi-2a-r}\equiv\bar{\epsilon},\quad\text{where}\quad\bar{\epsilon}<1. \quad (46)$$

The proposition then follows. □

C Dependence of the Stabilizing Effect of the Ruler's Absolute Power on the Elite–People Relationship

Lemma 1 shows that the impact of the Elite–People symmetry on the stability of the status quo is generally increasing in the absolute power of the Ruler. We can show a parallel result when analyzing the impact of the Ruler's absolute power on the stability of the status quo:

Lemma C.1. *The stabilizing effect of the Ruler's absolute power is increasing in the level of the Elite–People symmetry, i.e., $dS/d(1-\gamma)>0$ is increasing in β.*

Proof. By $\hat{x} = p(1-\gamma)\beta a$, $S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p)$, and $f(x) \in [\underline{f}, \bar{f}] \subset (0, \infty)$ over $x \in [0, pa]$, we have

$$\frac{dS}{d(1-\gamma)} = f(\hat{x}) \cdot (1-p) \cdot \frac{d\hat{x}}{d(1-\gamma)} = f(\hat{x}) \cdot (1-p) \cdot p\beta a > 0, \tag{47}$$

which is increasing in β. □

D Allowing Current Stability to Shape Future Power Structure

Based on the equilibrium at Stage 2, R's preference over γ is straightforward: a lower γ stabilizes the status quo (higher S) without any impact on R's status quo payoff; therefore, R will prefer the lowest possible γ.

Proposition 3 also implies:

Corollary D.1 (Higher β almost always preferred by R)**.** *As* $\pi - r \to \infty$, $\underline{\gamma} \to 1^-$.

This result suggests that when the surplus R would enjoy is sufficiently large, given any $\gamma < 1$, R will prefer β to be as high as possible. This result and R's preference over γ allow us to consider the following setting:

- At t:
  - The ruling position's historical strength S_{t-1} is given.
  - $\gamma_t = \gamma(S_{t-1})$ and $\beta_t = \beta(S_{t-1})$ are realized, where $\gamma(S)$ and $\beta(S)$ satisfy $\gamma_S(S) < 0$ and $\beta_S(S) > 0$, respectively.
  - The modeled Stage 2 plays out $S_t = 1 - (1 - F(\hat{x})) \cdot (1-p) \equiv S(\beta_t, \gamma_t, \theta)$ as in the unique subgame perfect equilibrium; θ include all factors that conditional on S_{t-1}, affect S_t but do so not through γ_t or β_t.
- At $t+1$: The same happens.

The dynamics then follows

$$\beta_t = \beta(S_{t-1}), \quad \gamma_t = \gamma(S_{t-1}), \quad S_t = S(\beta_t, \gamma_t, \theta), \tag{48}$$

or just

$$S_t = S\big(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\big). \tag{49}$$

Steady states are then defined by

$$S^* = S\left(\beta^*, \gamma^*, \theta\right), \quad \beta^* = \beta\left(S^*\right), \quad \gamma^* = \gamma\left(S^*\right), \tag{50}$$

or just

$$S^* = S\left(\beta\left(S^*\right), \gamma\left(S^*\right), \theta\right). \tag{51}$$

Existence and stability of steady states. The defining equation of steady states can help establish a few technical results. The first result is about the possible range of S_t in the dynamics:

Lemma D.1. *Any S_t in the dynamics must satisfy $\underline{S} \leq S_t \leq \bar{S}$, where $\underline{S} = p$ and $\bar{S} = 1 - (1-p) \cdot (1 - F(pa)) < 1$.*

Proof. Note that $S_\beta \geq 0$ and $S_\gamma \leq 0$. Therefore, the minimum $\underline{S}$ is reached when $\beta_t = 0$ and $\gamma_t = 1$ and the maximum $\bar{S}$ is reached when $\beta_t = 1$ and $\gamma_t = 0$. The lemma then follows. □

The first result helps establish the second result, which is about the existence of a steady state given a reasonable assumption about $\beta(\cdot)$ and $\gamma(\cdot)$:

Lemma D.2. *If $\beta(\underline{S})$, $\gamma(\underline{S})$, $\beta(\bar{S})$, and $\gamma(\bar{S})$ are all within the range $(0, 1)$, then there exists at least one steady state S^*, at which $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ crosses $S_t = S_{t-1}$ from $S_t > S_{t-1}$ to $S_t < S_{t-1}$, and $0 \leq S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1$.*

Proof. Note that $S_\beta > 0$ and $S_\gamma > 0$ for any $\beta > 0$ and $\gamma < 1$. Therefore, by $\beta(\underline{S}) > 0$ and $\gamma(\underline{S}) < 1$, we have $S\left(\beta\left(\underline{S}\right), \gamma\left(\underline{S}\right), \theta\right) > \underline{S}$; by $0 < \beta(\bar{S}) < 1$ and $0 < \gamma(\bar{S}) < 1$, we have $S\left(\beta\left(\bar{S}\right), \gamma\left(\bar{S}\right), \theta\right) < \bar{S}$. Since $S\left(\beta(s), \gamma(s), \theta\right)$ is continuous in s, the defining equation $S^* = S\left(\beta\left(S^*\right), \gamma\left(S^*\right), \theta\right)$ must have a solution $S^* \in \left[\underline{S}, \bar{S}\right]$, i.e., a steady state exists, at which $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ crosses $S_t = S_{t-1}$ from $S_t > S_{t-1}$ to $S_t < S_{t-1}$. Moreover, note that

$$\frac{dS\left(\beta\left(s\right), \gamma\left(s\right), \theta\right)}{ds} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \geq 0, \tag{52}$$

so $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ is increasing in S_{t-1}. Therefore, at S^*,

$$0 \leq \frac{dS\left(\beta\left(s\right), \gamma\left(s\right), \theta\right)}{ds} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{53}$$

□

The third result is the condition for a steady state to be stable:

Lemma D.3. *A steady state S^* is stable if and only if at S^*, $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ crosses $S_t = S_{t-1}$ from $S_t > S_{t-1}$ to $S_t < S_{t-1}$ and $0 \leq S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1$.*

Proof. First, suppose a steady state S^* is stable, then at S^*, $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ crosses $S_t = S_{t-1}$ and

$$-1 < \frac{dS\left(\beta\left(S^*\right), \gamma\left(S^*\right), \theta\right)}{dS^*} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{54}$$

Note that

$$S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \geq 0, \tag{55}$$

so

$$0 \leq S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{56}$$

Therefore, the crossing must be from $S_t > S_{t-1}$ to $S_t < S_{t-1}$.

The other direction of the lemma is straightforward. The lemma is then proven. □

The last two results establish the existence of stable steady states:

Proposition D.1. *If $\beta(\underline{S})$, $\gamma(\underline{S})$, $\beta(\bar{S})$, and $\gamma(\bar{S})$ are all within the range $(0, 1)$, then there exists at least one stable steady state, and at all the stable steady states, $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ crosses $S_t = S_{t-1}$ from $S_t > S_{t-1}$ to $S_t < S_{t-1}$ and $0 \leq S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1$.*

Institutional compatibility under multiple steady states. Assuming $\beta(\underline{S})$, $\gamma(\underline{S})$, $\beta(\bar{S})$, and $\gamma(\bar{S})$ are all within the range $(0, 1)$, we can have the following result: when multiple steady states exist given θ, any two different steady states must be different in a certain way, i.e., follows institutional compatibility:

Proposition D.2. *Given θ, if there are two steady states $\{S^*, \beta^*, \gamma^*\}$ and $\{S^{*\prime}, \beta^{*\prime}, \gamma^{*\prime}\}$, then any one among the following three statements will imply the other two: 1) $S^* \geq S^{*\prime}$; 2) $\beta^* \geq \beta^{*\prime}$; 3) $\gamma^* \leq \gamma^{*\prime}$.*

Proof. The result follows the three defining equations of steady states and their monotonicity. □

Given multiple steady states, the second result is about the divergence of compatible institutions:

Proposition D.3. *If there are $N \geq 2$ different stable steady states $S_1^* < \cdots < S_N^*$, then there are $N-1$ different unstable steady states $\tilde{S}_1 < \cdots < \tilde{S}_{N-1}$, they satisfy $\underline{S} < S_1^* < \tilde{S}_1 < S_2^* < \tilde{S}_2 < \cdots < S_{N-1}^* < \tilde{S}_{N-1} < S_N^* < \bar{S}$, and the institutional dynamics is determined by the initial strength of the ruling position S_0:*

- *if* $\tilde{S}_n < S_0 < \tilde{S}_{n+1}$*, where* $n = 1, \ldots, N-1$*, then* $S_t \to S^*_{n+1}$ *as* $t \to \infty$*;*
- *if* $\underline{S} \leq S_0 < \tilde{S}_1$*, then* $S_t \to S^*_1$ *as* $t \to \infty$*;*
- *if* $\tilde{S}_{N-1} < S^*_N < \bar{S}$*, then* $S_t \to S^*_N$ *as* $t \to \infty$*.*

Proof. As eventually $S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right)$ has to cross $S_t = S_{t-1}$ from $S_t > S_{t-1}$ to $S_t < S_{t-1}$, we can rank the stable and unstable steady states as proposed. Neighboring unstable steady states then divide the possible range of S into sub-ranges, starting from each of which S_t will converge to the stable steady state in it. □

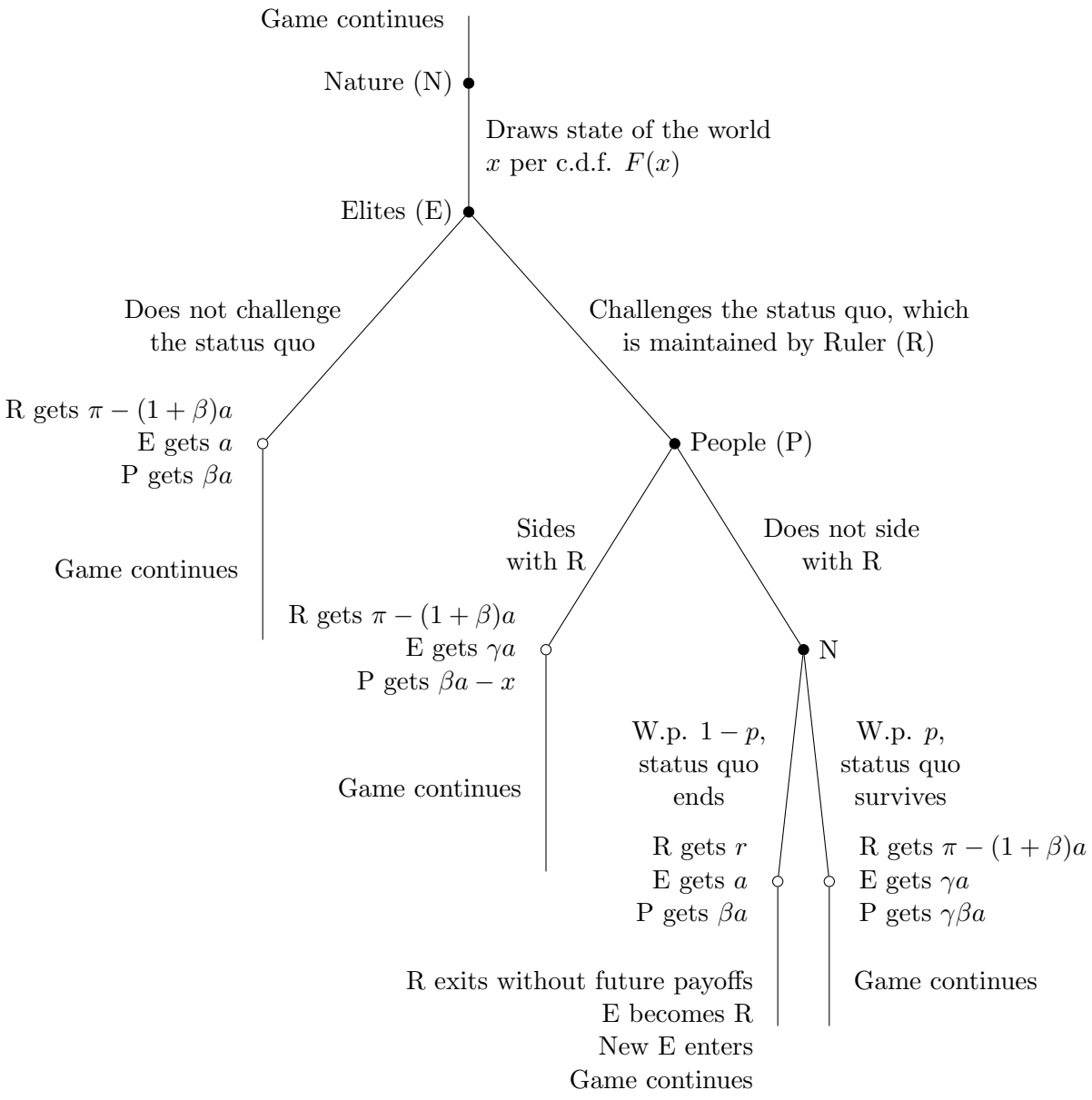
This result implies that the institutional difference between China and Europe can be thought as different stable steady states given the same primitives but different initial strengths S of the ruling position in history, which is compatible with different β and γ at very early times.

E Endogenizing the Challenger and Elites' Incentives in a Markov Game

In this extension of Stage 2 we collapse C and E into a single player E, make her look forward in a Markov game with an infinite number of discrete periods, and allow her to replace R. Figure 6 shows each period of the Markov game.

Compared with Figure 1, Stage 2 will now continue after each period; the prize z for C to challenge and the incentive w for E not to side with R are replaced by the aspiration of E to replace R at the end of this period; the loss y for C if her challenges fails is replaced by the punishment that would reduce E's payoff from the status quo level a to γa. About the stochastic elements of the game, we assume that N's draws of x and whether R will survive the challenge on his own within each period and across periods are mutually independent. About the dynamic elements of the game, we assume that all the players have an infinite horizon with an exogenous intertemporal discount factor $\delta \in (0, 1)$. All other assumptions in the main text remain here.

We will adopt the Markov perfect equilibrium as the solution concept in our analysis. For simplicity, we still assume that E will not challenge and P will side with R if they are indifferent in their decision, respectively, ruling out mixed strategies. Online Appendix E.3 shows that allowing for mixed strategies would accommodate a mixed-strategy equilibrium when and only when pure-strategy equilibria do not exist, while the key insights would remain robust.



$x \geq 0$, $a > 0$, $\pi - 2a > r$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$, $0 < p < 1$

Figure 6: Extended Stage 2: Each period in the Markov game

E.1 Equilibrium Characterization

Now we analyze the extended Stage 2 by first characterizing all possible Markov perfect equilibria and finding the conditions under which they exist. We denote the net present values that the players enjoy at the beginning of each period as V^R, V^E, and V^P, respectively. We have a first result to partially characterize all Markov perfect equilibria:

Lemma E.1. *In any Markov perfect equilibrium, P will side with R if and only if* $x \leq \hat{x} \equiv$

$(1-\gamma)\beta p \cdot a$, *where* $\hat{x} \in [0, pa]$*; when* $x \leq \hat{x}$*, E will not challenge the status quo, and when* $x > \hat{x}$*, E will challenge if and only if the aspiration to replace R in equilibrium dominates the probability-adjusted punishment in case of a failed challenge:*

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a. \tag{57}$$

Proof. In any Markov perfect equilibrium, P will side with R if and only if

$$\beta a - x + \delta V^P \geq (\beta a + \delta V^P) \cdot (1-p) + (\gamma \beta a + \delta V^P) \cdot p, \tag{58}$$

i.e.,

$$x \leq (1-\gamma)\beta p \cdot a \equiv \hat{x}. \tag{59}$$

Given this strategy of P and the continuation strategy of E in the equilibrium, E will not challenge if $x \leq \hat{x}$, since

$$a + \delta V^E \geq \gamma a + \delta V^E \tag{60}$$

holds for any $\gamma \in [0, 1]$ and V^E; when $x > \hat{x}$, E will challenge if and only if

$$a + \delta V^E < (a + \delta V^R) \cdot (1-p) + (\gamma a + \delta V^E) \cdot p, \tag{61}$$

i.e.,

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a. \tag{62}$$

The lemma is then proven. □

Note that the analysis is parallel to Section 3.1.2, the definition of $\hat{x}$ is the same as in Section 3.1.2, and Condition (57) is parallel to Conditions (4) and (6).

By Lemma E.1, only two Markov perfect equilibria are possible. The first one is a secured-R equilibrium:

Proposition E.1 (Secured-R equilibrium in the Markov game)**.** *If*

$$h(\beta, \gamma) \equiv \frac{\pi - (2+\beta)a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \leq 0, \tag{63}$$

then "E never challenges the status quo; P would not side with R if and only if $x > \hat{x}$*" is a Markov perfect equilibrium; in this equilibrium, the survival probability of the status quo is* $S = 1$*.*

Proof. For "E never challenges the status quo; P would not side with R if and only if $x > \hat{x}$"

to be a Markov perfect equilibrium, the condition

$$V^R - V^E \leq \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \tag{64}$$

must hold, where, given E and P's strategies in this equilibrium,

$$V^R = \frac{\pi - (1+\beta)a}{1-\delta} \quad \text{and} \quad V^E = \frac{a}{1-\delta}. \tag{65}$$

The condition is then equivalent to

$$\frac{\pi - (1+\beta)a}{1-\delta} - \frac{a}{1-\delta} \leq \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{66}$$

i.e.,

$$h(\beta, \gamma) \equiv \frac{\pi - (2+\beta)a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \leq 0. \tag{67}$$

The proposition is then proven. □

The intuition of the result is as follows: the function $h(\beta, \gamma)$ measures E's aspiration $V^R - V^E = (\pi - (2+\beta)a)/(1-\delta)$ to replace R given the specified strategies, net of the probability-adjusted punishment $\big(p/(1-p)\delta\big) \cdot (1-\gamma)a$ on E in case the challenge fails. The condition $h(\beta, \gamma) \leq 0$ then suggests that the aspiration cannot dominate the punishment. Lemma E.1 then implies that we have the secured-R equilibrium.

Note that this equilibrium is parallel to the scenario in Section 3.1.2 when Conditions (4) and (6) do not hold. Following the same argument as in Section 3.1.2, this equilibrium is empirically not much relevant, as in reality the chance for R to be ousted was always strictly positive; it is also trivial, in the sense that no challenge will happen in equilibrium.

The second equilibrium is an unsecured-R equilibrium:

Proposition E.2 (Unsecured-R equilibrium in the Markov game)**.** *If*

$$g(\beta, \gamma) \equiv \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a > 0, \tag{68}$$

where

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p) \in [p, 1] \quad \textit{and} \quad \hat{x} \equiv (1-\gamma)\beta p \cdot a, \tag{69}$$

then "E will challenge the status quo if and only if $x > \hat{x}$; P would not side with R if and only if $x > \hat{x}$" is a Markov perfect equilibrium; in this equilibrium, R's stability is $S \leq 1$.

Proof. For "E will challenge the status quo if and only if $x > \hat{x}$; P would not side with R if

and only if $x > \hat{x}$" to be a Markov perfect equilibrium, the condition

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \tag{70}$$

must hold, where, given E and P's strategies in this equilibrium,

$$\begin{aligned} V^R &= \Big(\pi - (1+\beta)a + \delta V^R\Big) \cdot S + r \cdot (1-S) \\ &= \big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S) + \delta V^R \cdot S \\ &= \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} \end{aligned} \tag{71}$$

and

$$\begin{aligned} V^E &= a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot p\Big) + \gamma a \cdot \big(1 - F(\hat{x})\big) \cdot p + \delta V^E \cdot S + \delta V^R \cdot (1-S) \\ &= a \cdot \Big(1 - (1-\gamma) \cdot \big(1 - F(\hat{x})\big) \cdot p\Big) + \delta V^E \cdot S + \delta V^R \cdot (1-S) \\ &= \frac{a \cdot \Big(1 - (1-\gamma) \cdot \big(1 - F(\hat{x})\big) \cdot p\Big) + \delta V^R \cdot (1-S)}{1-\delta S}, \end{aligned} \tag{72}$$

with

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p) \in [p, 1]. \tag{73}$$

The condition is then equivalent to, with some algebra,

$$g(\beta, \gamma) \equiv \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a > 0. \tag{74}$$

The proposition is then proven. □

Again, the intuition of Proposition E.2 follows Lemma E.1: the function $g(\beta, \gamma)$ indicates, given the specified strategies, how E's aspiration $V^R - V^E$ to replace R is compared with the punishment in case the challenge fails. The condition $g(\beta, \gamma) > 0$ then suggests that the aspiration dominates the punishment. Lemma E.1 then implies that we have the unsecured-R equilibrium.

Following the same argument as in Section 3.1.2, the unsecured-R equilibrium is empirically relevant and nontrivial. We thus now explore the conditions under which it always exists and is the unique equilibrium. The following result first shows that the secured-R equilibrium and the unsecured-R equilibrium cannot exist simultaneously:

Corollary E.1. *Given $r \leq \pi - 2a$, if $g(\beta, \gamma) > 0$, then $h(\beta, \gamma) > 0$, i.e., if the unsecured-R*

equilibrium exists, then the secured-R equilibrium does not exist.

Proof. Observe that, by $r \leq \pi - 2a$, for any $S \in [p,1]$, $g(\beta,\gamma) \leq g(\beta,\gamma)\big|_{S=1} = h(\beta,\gamma)$. Therefore, if $g(\beta,\gamma) > 0$, then $h(\beta,\gamma) > 0$. □

The intuition of Corollary E.1 is as follows. Since R is safer in the secured-R equilibrium than in the unsecured-R equilibrium, E's aspiration to replace R is stronger, too. Therefore, if E's aspiration is already so strong that the unsecured-R equilibrium is supported ($g(\beta,\gamma) > 0$), then given the strategies specified in the secured-R equilibrium, E's aspiration must be too strong to support the secured-R equilibrium ($h(\beta,\gamma) > 0$).

This corollary helps derive a set of conditions under which the unsecured-R equilibrium will generally exist and be the unique equilibrium, parallel to Proposition 1:

Proposition E.3 (Focus on unsecured-R equilibrium in the Markov game)**.** *If* $\big((1-\delta p)/(1-\delta)(1-p)\delta\big) \cdot a \leq r \leq \pi - 2a$*, then given any* $\beta \in [0,1]$ *and* $\gamma \in [0,1]$*, the unsecured-R equilibrium exists and is the unique Markov perfect equilibrium.*

Proof. For any $\beta \in [0,1]$ and $\gamma \in [0,1]$, by $0 < \Big((1-\delta p)/(1-\delta)(1-p)\delta\Big) \cdot a \leq r \leq \pi - 2a$ and $S \in [p,1]$, we have

$$\begin{aligned} g(\beta,\gamma) &\geq \frac{(\pi - 2a)\cdot S + r\cdot(1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta}\cdot a \\ &\geq \frac{r}{1-\delta S} - \frac{(1-p)\delta + p(1-\delta)}{(1-\delta)(1-p)\delta}\cdot a > \frac{r}{1-\delta p} - \frac{1-p+p}{(1-\delta)(1-p)\delta}\cdot a \\ &\geq \frac{r}{1-\delta p} - \frac{1}{(1-\delta)(1-p)\delta}\cdot a \geq 0. \end{aligned} \tag{75}$$

Therefore, $g(\beta,\gamma) > 0$, i.e., the unsecured-R equilibrium exists, and by Corollary E.1, the secured-R equilibrium does not exist. Therefore, the unsecured-R equilibrium is the unique equilibrium. □

In this result, $\Big((1-\delta p)/(1-\delta)(1-p)\delta\Big) \cdot a \leq r$ is parallel to $w > ap/(1-p)$ and $z > yp/(1-p)$ in Proposition 1, guaranteeing that E's aspiration to replace R is sufficiently strong so that E will challenge if P will not side with R.

E.2 Analysis of the Unsecured-R Equilibrium

To focus on the empirically relevant, nontrivial unsecured-R equilibrium in our analysis, from now on we assume that the condition in Proposition E.3 holds, i.e., $\big((1-\delta p)/(1-\delta)(1-p)\delta\big) \cdot a \leq r \leq \pi - 2a$, so that the unsecured-R equilibrium exists and is the unique Markov perfect equilibrium. Without losing generality, as in Section 3.2, we also assume

that the state of the world x's probability density function satisfies $f(x) \in [\underline{f}, \bar{f}] \subset (0, \infty)$ over $x \in [0, pa]$.

Now we can derive parallel results to Sections 3.1.3 and 3.2. First note that as in Section 3.1.3, in each period, the probability of challenge is still $1 - F(\hat{x})$ and the survival probability of the status quo is still

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1 - p), \tag{76}$$

where $\hat{x} \equiv (1 - \gamma)\beta p \cdot a$, so Proposition 2 still holds in this Markov game.

Now examine R's preference over γ and β. The net present value of R's payoffs in equilibrium is

$$V^R = \frac{\big(\pi - (1 + \beta)a\big) \cdot S + r \cdot (1 - S)}{1 - \delta S} = \frac{\big(\pi - (1 + \beta)a - r\big) \cdot S + r}{1 - \delta S}, \tag{77}$$

which differs from Equation (8) only at that it includes the future payoffs. Therefore, R will still prefer γ to be as low as possible.

On R's preference over β, first, since Proposition 2 still holds in this Markov game, the political–economic trade-off still appears and Lemma 1 still holds. We can then derive the following result parallel to Proposition 3:

Proposition E.4 (Institutional compatibility in the Markov game)**.** *If $\gamma < \underline{\bar{\gamma}} \equiv 1 - (1 - \delta)\big/\big((1 - \delta(1 - p))(\pi - 2a - r) + \delta r\big)\,\underline{c}$, then R will prefer β to be as high as possible; if $\gamma > \bar{\bar{\gamma}} \equiv 1 - (1 - \delta)p\big/\big(\pi - a - r(1 - \delta)\big)\,\bar{c}$, then R will prefer β to be as low as possible, where $\underline{\bar{\gamma}} < \bar{\bar{\gamma}} < 1$ and if $\pi > 2a + \left(\frac{1-\delta}{\underline{c}} + (1 - \delta(2 - p))r\right) \big/ \big(1 - \delta(1 - p)\big)$, then $\underline{\bar{\gamma}} > 0$.*

Proof. The marginal impact of β on R's net present value in equilibrium is

$$\frac{dV^R}{d\beta} = \frac{\left(\pi - (1 + \beta)a - r + \frac{\delta\left(\left(\pi - (1+\beta)a - r\right)S + r\right)}{1 - \delta S}\right) \cdot \frac{dS}{d\beta} - aS}{1 - \delta S}. \tag{78}$$

By Lemma 1, $\beta \in [0, 1]$, $S \in [p, 1]$, and $0 < \Big((1 - \delta p)\big/(1 - \delta)(1 - p)\delta\Big) \cdot a \leq r \leq \pi - 2a$, we have

$$\begin{aligned} \frac{dV^R}{d\beta} &\geq \frac{\left(\frac{(1-\delta(S-p))(\pi-2a-r)+\delta r}{1-\delta S}\right) \cdot \underline{c}a \cdot (1 - \gamma) - a}{1 - \delta S} \\ &\geq \frac{a}{1 - \delta S} \cdot \left(\frac{(1 - \delta(1 - p))(\pi - 2a - r) + \delta r}{1 - \delta} \cdot \underline{c} \cdot (1 - \gamma) - 1\right), \end{aligned} \tag{79}$$

so, if

$$\frac{\left(1-\delta(1-p)\right)\left(\pi-2a-r\right)+\delta r}{1-\delta}\cdot\underline{c}\cdot(1-\gamma)-1>0, \tag{80}$$

i.e.,

$$\gamma<1-\frac{1-\delta}{\left(\left(1-\delta(1-p)\right)\left(\pi-2a-r\right)+\delta r\right)\cdot\underline{c}}\equiv\underline{\underline{\gamma}}, \tag{81}$$

then $dV^R/d\beta>0$. At the same time, we have

$$\begin{aligned}\frac{dV^R}{d\beta}&\leq\frac{\left(\pi-a-r+\frac{\delta(\pi-a)}{1-\delta}\right)\cdot\bar{c}a\cdot(1-\gamma)-ap}{1-\delta S}\\&=\frac{a}{1-\delta S}\cdot\left(\left(\frac{\pi-a}{1-\delta}-r\right)\cdot\bar{c}\cdot(1-\gamma)-p\right),\end{aligned} \tag{82}$$

so, if

$$\left(\frac{\pi-a}{1-\delta}-r\right)\cdot\bar{c}\cdot(1-\gamma)-p<0, \tag{83}$$

i.e.,

$$\gamma>1-\frac{(1-\delta)p}{\left(\pi-a-r(1-\delta)\right)\cdot\bar{c}}\equiv\bar{\bar{\gamma}}, \tag{84}$$

then $dV^R/d\beta<0$. Finally, note $\underline{\underline{\gamma}}<\bar{\bar{\gamma}}<1$, and $\underline{\underline{\gamma}}>0$ is equivalent to $\pi>2a+\left(\frac{1-\delta}{\underline{c}}+(1-\delta(2-p))r\right)/\left(1-\delta(1-p)\right)$. The proposition is then proven. □

Proposition E.4 differs from Proposition 3 only at that $\underline{\underline{\gamma}}$ and $\bar{\bar{\gamma}}$ are differently defined from how $\underline{\gamma}$ and $\bar{\gamma}$ are defined, respectively, due to the change in the expression of V^R. This is then followed by parallel results to Section 3.3. We have then shown that we can derive all the parallel results to the main text from the Markov game.

E.3 Allowing for Mixed Strategies

Here we allow for mixed strategies by dropping the earlier assumption that E will not challenge and P will side with R if they are indifferent between their options. We then re-characterize all the Markov perfect equilibria of the game at Stage 2 and examine whether the main insights would remain. As in Online Appendix A, we assume $\gamma<1$; we also assume that x is a continuous random variable so that its distribution does not have any mass point, and that $F(ap)<1$ so that $1-F(\hat{x})>0$ always holds.

In any Markov perfect equilibrium, P's strategy is then "not to side with R when $x>\hat{x}$ and to side with R when $x<\hat{x}$." As x is a continuous random variable, we can leave P's strategy when $x=\hat{x}$ unspecified without much real consequence.

By $\gamma < 1$, given P's strategy and E's continuation strategy in the equilibrium, E's strategy is then not to challenge when $x < \hat{x}$; when $x > \hat{x}$, E will challenge with a given probability $q_E(x) \in [0,1]$, which is a function of $x > \hat{x}$, and we denote

$$\bar{q}_E \equiv \frac{\int_{\hat{x}}^{\infty} q_E(x) dF(x)}{1 - F(\hat{x})} \in [0,1]. \tag{85}$$

In particular, if in equilibrium

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{86}$$

then $q_E(x) = 1$ for any $x \geq \hat{x}$, with $\bar{q}_E = 1$; if in equilibrium

$$V^R - V^E < \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{87}$$

then $q_E(x) = 0$ for any $x \geq \hat{x}$, with $\bar{q}_E = 0$; if in equilibrium

$$V^R - V^E = \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{88}$$

then $q_E(x)$ should make this condition hold. Again, as x is a continuous random variable, we can leave E's strategy when $x = \hat{x}$ unspecified.

In the equilibrium with such strategies, we must have

$$V^R = \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1 - \delta S}, \tag{89}$$

$$\begin{aligned} V^E &= a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p\Big) + \gamma a \cdot \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p + \delta V^E \cdot S + \delta V^R \cdot (1-S) \\ &= \frac{a \cdot \Big(1 - (1-\gamma) \cdot \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p\Big) + \delta V^R \cdot (1-S)}{1 - \delta S}, \end{aligned} \tag{90}$$

and

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot (1-p). \tag{91}$$

By some algebra, the function that governs the existence of the equilibrium turns out to be

$$\begin{aligned} &V^R - V^E - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \\ &= \frac{1-\delta}{1-\delta S} \cdot \left(\frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \right). \end{aligned} \tag{92}$$

Now define

$$k(\beta,\gamma,\bar{q}_E) \equiv \frac{\big(\pi-(1+\beta)a\big)\cdot S + r\cdot(1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta}\cdot(1-\gamma)a, \tag{93}$$

where

$$S = 1-\big(1-F(\hat{x})\big)\cdot\bar{q}_E\cdot(1-p) \quad \text{and} \quad \hat{x} = (1-\gamma)\beta p a. \tag{94}$$

Note that by $F(pa)<1$ and $\pi-2a>r$, $k(\beta,\gamma,\bar{q}_E)$ is strictly decreasing over $\bar{q}_E\in[0,1]$. We can then characterize the Markov perfect equilibria in three scenarios, except for E and P's strategies when $x=\hat{x}$:

1. When $k(\beta,\gamma,0)<0$, the unique family of Markov perfect equilibria that can exist must satisfy:

   - P will side with R when $x<\hat{x}$ and will not side with R when $x>\hat{x}$;
   - E will never challenge when $x\neq\hat{x}$.

2. When $k(\beta,\gamma,1)>0$, the unique family of Markov perfect equilibria that can exist must satisfy:

   - P will side with R when $x<\hat{x}$ and will not side with R when $x>\hat{x}$;
   - E will not challenge when $x<\hat{x}$ and will challenge when $x>\hat{x}$.

3. When $k(\beta,\gamma,0)\geq 0$ and $k(\beta,\gamma,1)\leq 0$, there exists a unique $\bar{q}_E\in[0,1]$ such that

   $$k(\beta,\gamma,\bar{q}_E)=0, \tag{95}$$

   and the unique family of Markov perfect equilibria that can exist must satisfy:

   - P will side with R when $x<\hat{x}$ and will not side with R when $x>\hat{x}$;
   - E will challenge with a given probability $q_E(x)\in[0,1]$, where the function $q_E(x)$ satisfies
     $$\frac{\int_{\hat{x}}^{\infty} q_E(x)dF(x)}{1-F(\hat{x})} = \bar{q}_E, \tag{96}$$
     when $x>\hat{x}$ and will not challenge when $x<\hat{x}$.

Note that Scenario 1 corresponds to Proposition E.1, where $h(\beta,\gamma)\equiv k(\beta,\gamma,0)$, and Scenario 2 corresponds to Proposition E.2, where $g(\beta,\gamma)\equiv k(\beta,\gamma,1)$. Now examine whether our main messages remain in Scenario 3.

In Scenario 3, in equilibrium, we always have

$$k(\beta,\gamma,\bar{q}_E) \equiv \frac{\big(\pi-(1+\beta)a\big)\cdot S + r\cdot(1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta}\cdot(1-\gamma)a = 0, \tag{97}$$

i.e.,

$$\big(\pi-(1+\beta)a-r\big)\cdot S + r = \left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\cdot a\cdot(1-\delta S). \tag{98}$$

This implies

$$dS = \frac{\frac{pa(1-\delta S)}{(1-p)\delta}\cdot d(1-\gamma) + aS\cdot d\beta}{\pi-(1+\beta)a-r+\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\delta a}. \tag{99}$$

By $\pi - 2a > r$, we see that a higher β and a lower γ will increase in equilibrium the survival probability of the status quo S, corresponding to Proposition 2, which is for Scenario 2.

This result also suggests that in equilibrium

$$\begin{aligned}\frac{dS}{d\beta} &= \frac{aS}{\pi-(1+\beta)a-r+\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\delta a} \\ &= \frac{a\cdot\left(\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\cdot a - r\right)}{\left(\pi-(1+\beta)a-r+\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\delta a\right)^2}.\end{aligned} \tag{100}$$

This implies

$$\frac{dS}{d\beta} \le \frac{a\cdot\left(\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\cdot a - r\right)}{\left(\pi-2a-r+\frac{1}{1-\delta}\cdot\delta a\right)^2} \equiv \bar{b}(\gamma) \tag{101}$$

and

$$\frac{dS}{d\beta} \ge \frac{a\cdot\left(\left(\frac{1}{1-\delta}+\frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\cdot a - r\right)}{\left(\pi-a-r+\left(\frac{1}{1-\delta}+\frac{p}{(1-p)\delta}\right)\delta a\right)^2} \equiv \underline{b}(\gamma), \tag{102}$$

where both $\bar{b}(\gamma)$ and $\underline{b}(\gamma)$ are decreasing in γ. The insight in Lemma 1 then maintains. A result similar to Proposition E.4 would then follow.

To summarize, allowing for mixed strategies would allow the mixed-strategy, Scenario-3 equilibria to exist, in which the main insights from Scenario 2 would maintain, but with

more technical complexity. In light of this, we can rule out mixed strategies from the Markov game, gaining in simplicity without losing much intuition.

F Relevance of Elites and People in Conflicts

There existed a wide range of conflicts in both Chinese and European histories. Having carefully examined significant examples, we argue that the positions taken by the Elites and the People were critical in determining the outcome of the conflict. Below we discuss some examples.[30]

History has shown that given the Elites' political, economic, and military resources, whether they sided with the Ruler when the Ruler was challenged was critical to the outcome of the challenge. For example, the fate of the French throne during the Hundred Years' War closely followed whether the Duke of Burgundy, first John the Fearless and later his son Philip the Good, allied with the English or veered back to the French ruler (Seward, 1978). During the Wars of the Roses (1455–1485), "crucially, Thomas, Lord Stanley, refused to answer Richard [III of England]'s summons" in the Battle of Bosworth in 1485, and his brother "Sir William Stanley committed his men, tipping the battle decisively in Henry [Tudor, later Henry VII of England]'s favour," delivering the demise of Richard III and the coronation of Henry VII (Grummitt, 2014, p. 123). In China, during the civil war at the end of the Sui dynasty (611–618), Emperor Yang was killed in a coup by Yuwen Huaji, the commander of the royal guard and the son of Duke Yuwen Shu; during the late Tang dynasty, after Qiu Fu, Wang Xianzhi, and Huang Chao led peasants to revolt all over the country (859–884), it was the regional governors, such as Wang Chongrong and Li Keyong, who fought hard to recover Chang'an, defeated the uprisings, and restored the throne of Tang.

The People's position was more than often crucial, too, as we can see in the history of not only China but also Europe. In Chinese history, in the final years of the Qin, Xin, Sui, Tang, Yuan, and Ming dynasties, following the initial rebellion within the country or invasion from the outside, peasants revolted and contributed to the end of these dynasties. In Europe, for example, Morton (1938, p. 46, 63) comments on the English history: "the

[30] An incomplete list of the examples we examine include, for China, the Qin–Han turnover, Rebellion of the Seven Prince States, Western Han–Xin turnover, Xin–Eastern Han turnover, Eastern Han–Three Kingdoms turnover, Western–Eastern Jin turnover, Eastern Jin–Southern Dynasties turnover, Sui–Tang turnover, Tang–Zhou turnover, An Lushan Rebellion, Huang Chao Rebellion and Tang–Five Dynasties and Ten Kingdoms turnover, Northern–Southern Song turnover, Yuan–Ming turnover, Ming–Qing turnover, and Revolt of the Three Feudatories; for Europe, the Rebellion of Robert de Mowbray, Henry I's invasion of Normandy, 1215 Magna Carta, Second Barons' War, Hundred Years' War, Jacquerie, Wat Tyler's Rebellion, Richard II–Henry IV of England turnover, Jack Cade's Rebellion, Wars of the Roses, German Peasants' War, Dutch Revolt, and Thirty Years' War. Some examples include more than one entries of examination. These cover 15 and 14 entries for China and Europe, respectively, and 29 in total.

king was able to make use of the peasantry in a crisis when his position was threatened by a baronial rising," and "even the strongest combination of barons had failed to defeat the crown when, as in 1095 [Robert de Mowbray's rebellion] and in 1106 [the challenge of Duke Robert Curthose of Normandy over the throne of Henry I], it had the support of other classes and sections of the population." Finer (1997b, p. 901) also observes that the English fyrd, largely mobilized from the freemen, "was retained, and even called out by the Norman kings against their rebellious Norman barons." In the Hundred Years' War, the turning point toward the eventual French triumph was the rise of Joan of Arc, as she inspired the common people of France to join the war.[31] In England, shortly before and during the Wars of the Roses, popular support was generally important in determining how firmly Richard II, Henry IV, Henry VI, Edward IV, and Richard III could hold the throne (e.g., Morton, 1938; Bennett, 1999; Grummitt, 2014).[32] In the German Peasants' War, as

[31]For more details on the French throne's lack of popular support before Joan of Arc, the change after that, and the implications of the change on the development of the war, see Morton (1938) and Seward (1978).

[32]During the Richard II–Henry IV turnover, "Richard found himself without supporters," as the common "merchants [have been] alienated" (Morton, 1938, p. 115); Grummitt (2014, p. 5) comments that "Lancastrian legitimacy was based on an appeal to popular support," and Bennett (1999, p. 204) states that "it was widely believed that Henry had been raised to the throne on the basis of a covenant with the people." During Henry VI's reign, he relied heavily on the support of the people, i.e., "the willingness of the political nation to act for the common good," but later in his reign "[t]he Commons could have little confidence in the king" (Grummitt, 2014, p. 21). Over 1449–1454, "the defeat in France led to a popular groundswell of opinion against the Lancastrian regime; the appeal to the commons that had been one of the foundations of Lancastrian rule would, in part, prove its undoing" (Grummitt, 2014, p. 13, 21, 23); in 1450, Jack Cade's Rebellion broke out, during which Londoners played a decisive role by first siding with the revolt but later deserting them for looting, eventually "shutting off Cade and his men ... from the City" (Morton, 1938, p. 123); Grummitt (2014, p. 161) comments that "[i]n 1450 politics were driven by an agenda that was unmistakably set by the commons," which would continue "[t]hroughout the following two decades" when Richard, Duke of York, would have, "to all intents and purposes, become an opportunistic 'Cadist', jumping on a popular bandwagon to end his self-imposed political exile" (Grummitt, 2014, p. 161, also 24–28). Following the strategy, Edward IV's taking of the throne depended on the support of the common people: "[t]he most significant aspect of the series of events that led to Edward's accession was ... the judgment made by the assembled crowds at St John's Fields ... that Henry VI ... should be deposed for his recent misdeeds. ... The assembled crowd then merely assented to Edward's de jure right to be king of England. ... The notion of popular support for the new king emerged as an important principle of Yorkist propaganda," and "the Yorkists owed their success in no small part to the support they enjoyed from the commons. ... This popularity was also crucially transformed into tangible expressions of support in terms of men and, more importantly perhaps, money" (Grummitt, 2014, p. 71, 76–77). At that time, the importance of the people made them also a force to be used by other political players: for example, "[m]ore than any of his contemporaries perhaps, [Richard Neville] was able to exploit the idea of the commonweal and mobilise the commons' sophisticated understanding of the constitution and the nature and duties of kingship in support of his own aims" (Grummitt, 2014, p. 102). During the Edward IV–V–Richard III turnover, "[a]ll the chronicles agree that [Richard III's] usurpation evoked no popular enthusiasm" initially (Wood, 1975, p. 270), but what followed proved again what the common People could do: Buckingham's rebellion failed because it "failed to mobilise popular support," while Richard III's lack of popular support when he was killed in the decisive Battle of Bosworth suggested that "[t]he disengagement from politics [of the common People] was probably the most damning indictment of Richard III's short reign" (Grummitt, 2014, p. 117, 124). In general, the

the status quo was challenged by peasants across southwestern Germany, the uprisings were eventually defeated by the Swabian League, given that the support from the common people in cities were inconsistent.

These examples show that both the Elites and the People are highly relevant in conflicts. This gives us confidence to link the power structure among the Ruler and both the Elites and the People to the stability of autocratic rule.

G Dynastic Cycle in Chinese History

Our model also provides an interpretation of the dynastic cycle, a prominent observation of Chinese history. In brief, students of Chinese history often observe that each dynasty started with a relatively stable autocratic rule, but would see over generations declining power of the emperor, increasing dominance of the elites over the emperor and common people, increasing concentration of land ownership, and decreasing effectiveness of governance, eventually slipping to chaos and leading to the end of the dynasty (e.g., Skinner, 1985; Usher, 1989; Dillon, 1998).

To interpret the observation, guided by our model, we start by noting that the founding emperor of each dynasty often enjoyed a strong absolute power, i.e., a low γ, since he was bestowed with a high level of legitimacy by receiving himself the Mandate of Heaven to be the new Ruler (e.g., Zhao, 2009; Jiang, 2011). By Proposition 3 and Online Appendix D, he would have been more willing or able to restrict the asymmetry between the Elites and People, maintaining a relatively high β; as a result of the high β–low γ power structure, the stability of autocratic rule would have been relatively high. Over generations, however, later emperors became more and more distant from the act of receiving the Mandate of Heaven; an increasing number of precedents also placed further constraints on their behaviors; all these led to a decline of the Ruler's absolute power, i.e., a higher γ. By Proposition 3 and Online Appendix D, again, the later emperors would be less willing or able to enforce a more symmetric Elite–People relationship, leading to a lower β. The low β–high γ power structure would then lead to a lower stability of autocratic rule and, eventually, a downward spiral to its collapse.

Wars of the Roses "had witnessed the incorporation of the commons as a legitimate partner with the king, lords and parliamentary Commons in the process of politics and government" (Grummitt, 2014, p. 163).