A Theory of Power Structure and Political Stability: China vs. Europe Revisited*

Ruixue Jia† Gérard Roland‡ Yang Xie§

September 8, 2020

Abstract

A large literature in economics has emphasized the importance of rule of law and property rights for economic development. Yet, a comparative analysis of the historical trajectories of medieval Europe and imperial China raises puzzles that cannot be readily solved by the existing framework. Why was Europe, with stronger rule of law and property rights, mired in conflict during most of its history, while China experienced relatively higher political stability? We offer one answer by focusing on power structure: how power was shared among three estates – the Ruler, the Elites (lords or bureaucrats), and the People. Based on historical narratives, we emphasize two important differences: (1) the Ruler enjoyed less absolute power in Europe than in China; (2) the rights of the Elites and the People were more symmetric, i.e., less unbalanced in China than in Europe. Using a simple theoretical framework, we show that both differences led to higher political stability in China via two channels – a generic punishment channel and a strategic political alliance channel. The coexistence of both above differences can be explained on the basis of the same political–economic trade-off faced by the Rulers in China and in Europe.

*We are grateful to Joel Mokyr for his thorough and constructive feedback. We have also benefited from comments from Guido Tabellini.

†School of Global Policy and Strategy, University of California, San Diego, CIFAR and NBER; rxjia@ucsd.edu.

‡Department of Economics, University of California, Berkeley, CEPR, and NBER; groland@econ.berkeley.edu.

§Department of Economics, University of California, Riverside; yang.xie@ucr.edu.

1 Introduction

A large body of economic research has documented the importance of property rights for economic development. North (1989) argues that the success of industrialization in Great Britain was the result of better institutions (particularly property rights and constraints on the executive) compared to absolutist governments like Spain. These insights were confirmed by the pioneering empirical analysis in Acemoğlu et al. (2001) and the very influential subsequent literature. A comparative analysis of the historical trajectories of medieval Europe and imperial China raises, however, a puzzle that cannot be readily solved by the existing institutionalist framework. Why was Europe, with stronger rule of law and property rights, mired in conflicts during most of its history, while China, known for its lack of rule of law, enjoyed relatively less frequent major conflicts and higher stability of political centralization?[1] Are institutional differences of the two societies part of the answer?

It seems not *a priori* clear why less absolute power of the ruler and stronger property rights led to more conflicts in Europe. For instance, one could imagine that the incentives of the ruled were more aligned with the ruler in Europe thanks to stronger property rights, which could potentially contribute to more stability rather than conflict.[2] In this paper, we investigate this issue by characterizing the power structure in China and Europe. While acknowledging the differences in rule of law, we emphasize the role of power allocation among three estates: the Ruler, the Elites (lords or bureaucrats), and the People. Based on rich historical narratives, we emphasize two important differences in the theory we propose: (1) the Ruler enjoyed less absolute power in Europe than in China; (2) the rights between the Elites and the People were more symmetric (less unbalanced) in China than in Europe. We propose to explain why these differences in power structure led to disparate political trajectories.

The first difference – stronger or weaker absolute power of the ruler – has been well recognized by existing studies in political economy that emphasize stronger rule of law and property rights in Europe. Our contribution is to elucidate its relationship with the political stability. The second difference – the relative status of the Elites and the People – has been hardly paid any attention by economists and political scientists. Historians and sociologists, however, provide insights on the relationship between the Elites and the People (e.g., Levenson, 1965; Wickham, 2009). For instance, elite status in Europe was hereditary while it

[1] The political stability we are interested in is about conflicts and consequently the instability of the political status quo, rather than the regular discontinuation of a dynasty. It is true that due to marriages, European family lines could last despite many conflicts and regime changes.

[2] In fact, Blaydes and Chaney (2013) use this argument to explain why Christian kings became longer-lived than Muslim sultans from the 8th to 15th century.

was governed by a civil service exam in China. Several historical works have indicated that land ownership concentration in Europe was also higher than that in China (see Zhang, 2017 for many works on England and China), which is partly enforced by different inheritance rules – partible inheritance in China versus primogeniture in Europe. In China, the elites had less power than in Europe where landlords were absolute masters on their land, with little danger of encroachment from rulers. Moreover, ordinary peasants in Europe gradually lost all their land, their rights and power in the centuries following the Fall of the Roman Empire, and many of them became serfs (Wickham, 2009). In China, most peasants enjoyed de facto land user rights and a certain autonomy. We hope to bring this new perspective to the literature and show that this difference also contributed to different political stability and welfare outcomes in China and Europe.

We show that both differences matter for political stability. To accommodate different types of conflict (e.g., external wars, coups, and peasant rebellions), we lay out a general framework with four players: the Ruler, the Elites, the People, and the Challenger. An outside aggressor, defiant elite, or group of rebellious common people, the Challenger decides whether to challenge the status quo. Whether the status quo will survive depends on whether the Elites and People choose to side with the Ruler or not. Both stronger absolute power of the Ruler and more symmetric rights between the Elite and the People (like in China) facilitate political stability through two channels. Besides a generic punishment channel that makes the Elites and People less willing to support a challenge to the Ruler, there exists a strategic alliance channel, as the People are more likely to align with the Ruler in China than their counterparts in Europe (due to both the stronger absolute power of the Ruler and the more symmetric Elites–People relationship). This alliance further decreases the Elites' willingness to support a challenge to the Ruler. Expecting more (and less) political alliance among the Ruler, Elites, and People, the Challenger is less (and more) likely to initiate a conflict. The importance of political alliance between the Ruler and the People in shaping political stability is the first key insight from our model. In fact, as remarked by Orwell (1947, p. 17), this idea of the ruler and the common people "being in a sort of alliance against the upper classes" is "almost as old as history."[3] Our model thus illustrates why it was easier for the Chinese rulers than for the European rulers to succeed in such alliance.

A second important message emerges from our model: the co-existence of the two differences can be explained as a result of the same political–economic trade-off faced by the Rulers in China and in Europe. The Ruler in both societies cares about the expected payoff

[3]For example, *Han Feizi*, the most representative text in the Chinese Legalist tradition from the third century BC, emphasizes the stabilizing effect of the political alliance between the Ruler and the People against the Elites (Watson, 1964, p. 87).

from maintaining the status quo and staying in power, which depends both on political stability and the size of the economic surplus left after sharing with the Elites and the People. In the Chinese scenario, because the Ruler enjoys sufficiently strong absolute power, leveraging the rights between the Elites and the People can be sufficiently effective in increasing political stability that the political stability concern dominates the economic concern at the margin. In contrast, due to too weak absolute power of the European Ruler, leveraging the rights between the Elites and the People could help him too little in deterring the Elites from supporting a challenge to his rule. Consequently, the economic concern dominates, which gives the Ruler incentives to choose more asymmetric Elites–People rights. The symmetry or asymmetry of Elite–People rights is thus endogenous.

A few additional implications arise from our model. For instance, the size of the economy also matters. The Ruler of a sufficiently large economy prefers more symmetric Elites–People rights, as the political concern dominates at the margin, while the opposite is true for the ruler of a sufficiently small economy. Moreover, because the tradeoff between the Ruler's absolute power and the Elites–People rights, it is possible for the People to prefer a weaker rule of law, which speaks to a debate on the living standard of the people in historical China and Europe (see Pomeranz, 2000 for more related historical studies). We further show that our simple model can be extended by allowing the current political stability to affect future power structure. This speaks to why European Rulers may have hoped to choose more symmetric Elites–People rights but lacked the capacity because of the dynamic link between current stability and future power structure.

Our study contributes to the political economy literature exploring the links between formal and informal institutions, political stability and conflicts, as well as economic development. In this line of research, influential studies have shown the importance of institutions for development (e.g., North, 1989; Acemoğlu et al., 2001; North et al., 2009; Besley and Persson, 2011; Acemoğlu and Robinson, 2012, 2019; Mokyr, 2016; Cox et al., 2019) and how conflicts contributed to the rise of state capacity in Europe (e.g., Tilly, 1990). Yet, there is a clear tension between the "better" institutions, for example, stronger rule of law and property rights, and the lower political stability in Europe, compared with China. In our framework, we interpret these better institutions as weaker absolute power of the Ruler, and we model it as how much of the Elites and People's power or rights in the status quo would remain after they have defied the Ruler's will. We show that the weaker the Ruler's absolute power and thus the better these institutions, the lower political stability and the more frequent conflicts, resolving the tension.

On another aspect, the literature often analyzes a society by categorizing it into two estates (e.g., the ruler vs ruled, state vs society, elites vs non-elites, those with vs without access to political and economic resources and decisions). Persistent pressure and struggle from the lower estate is thus needed to achieve more "open-access" or "inclusive" institutions. We extend the two-estate framework into a three-estate framework. This helps us show that less power asymmetry between the Elites and People can help the Ruler strengthen political stability by strategically forging an alliance in support of the Ruler. Because of this, the Ruler may actively choose to co-opt the People by strategically reducing the Elites–People power asymmetry. Our model shows that this is especially true when the Ruler has strong absolute power, explaining why many co-opting measures were more present in a more absolutist regime as in China than in Europe.

It is well documented that autocratic political unification has been much more stable in the history of China than in Europe. Scholars have argued that this political divergence has fundamentally shaped the post-industrial economic development paths of these two regions (e.g., Mokyr, 2016). On the origin of the political divergence, several inspiring explanations have been put forward.[4] Related to ours, Acemoğlu and Robinson (2019) argue that the German-Roman tradition of a balanced state–society relationship put Europe in the narrow corridor of social, political, and economic development, whereas the state has been too dominant since the early history of China. Stasavage (2020) underscores that a strong bureaucracy in ancient China led it to a path different from Europe, favoring a stronger autocratic regime. In line with these efforts, we provide a power structure explanation to the political divergence between China and Europe. Moreover, without necessarily modeling details of various specific institutions, our model can be useful in interpreting the roles of these institutions. For instance, we can interpret a strong bureaucracy (instead of inheritable elite status) in China as reducing the Elites–People asymmetry and helping the Ruler align with the People; the same interpretation can apply to the chartering of cities in Europe, against the backdrop of the Ruler–Elites competition for political alliance with the People.

China and Europe exhibited two ways of organizing the allocation of power. By linking power structure with political stability, our study suggests that it is unclear whether one way dominates the other or whether the two ways would converge. Indeed, historians comparing pre-industrial living standards of people, such as life expectancy and consumption, often find it difficult to believe in the superiority of European pre-modern institutions – which are usually caricatured as "stronger property rights" – and turn to non-institutional factors like ecological advantages, colonialism and more wars to explain the rise of Europe (e.g., Pomeranz, 2000; Rosenthal and Wong, 2011; McNeill, 1982; Hoffman, 2017).[5] We hope that

[4] These explanations are based on geography (e.g., Turchin, 2009; Ko et al., 2018; Scheidel, 2019; Roland, 2020; Fernández-Villaverde et al., 2020), hydraulics (e.g., Wittfogel, 1957), culture and other institutions (e.g., Fukuyama, 2011; Greif and Tabellini, 2017).

[5] Pomeranz (2000) discusses many related studies and provides various numbers for comparison, including

our study speaks to these insights by characterizing both the relationship between the Ruler and the ruled and, within the ruled, the relationship between the Elites and the People.

In Section 2, we present historical narratives to motivate our assumptions. We describe our theoretical settings in Section 3. We characterize the equilibria in Section 4 and analyze the unique, theoretically nontrivial and empirically relevant equilibrium in Section 5, which delivers the two key insights of our framework. Section 6 concludes the paper.

2 China versus Europe: Historical Narratives

In this section, we provide historical narratives on which we base our model. For simplicity, we sometimes refer to "Europe" as if it were a single entity or discuss a specific country as an example for Europe. Of course, there exists important variation within Europe: Eastern Europe is different from Western Europe; France is different from England within Western Europe, and so forth. Our model is more about building an "ideal type" representation of differences in power structure between China and Europe, and it can be applied to interpret intermediate cases. Moreover, between the Roman Empire and the end of the Middle Ages, institutions and power structure were not invariant. Similar remarks can also be made about China in its millennial history. We follow the *longue durée* approach by focusing on important differences in power structure that persist over time. The key differences in the power structure that we emphasize seem to largely hold between the 5th and the 17th century.

Section 2.1 describes differences in political stability. Sections 2.2 and 2.3 present qualitative evidence on the two dimensions of power structure. Section 2.4 discusses the relevance of the elites and people in conflicts.

2.1 Political Stability

Since the difference in political stability has been well documented by a large literature, it seems not necessary to argue at length that China was more stable than Europe. Nevertheless, we discuss three pieces of evidence here. The first two are the number of states and the

life expectancy, ordinary and luxury consumption as well as access to transportation networks. Across all these measures, no indicator suggests that pre-industrial China performed worse than pre-industrial Europe. He turns to ecological advantages and the discovery of the New World as explanations for why the Industrial Revolution happened in Europe. Rosenthal and Wong (2011) emphasize that, if anything, tax rates were lower while public good provision was higher in China, compared with Europe. They argue that a greater need for war financing in Europe explains such differences, which is related to the idea that military competition led to the rise of Europe (McNeill, 1982). Hoffman (2017) provides systematic evidence on this point.

population claimed by one single polity, both of which reflect more stability of political centralization in China. A third proxy is the number of conflicts, which indicates the stability of political status quo.

Existing studies have documented that between the 5th and the 17th century, the number of states in Europe, which was typically above 25, was much higher than that in China, which was usually below 5 (e.g., Ko et al., 2018). In addition, Scheidel (2019) describes the proportion of the population of Europe and East Asia claimed by the largest polity in that area over time. In post-Roman Europe, 20% or less of the population has ever been claimed by one single polity. In contrast, the number is over 70% for East Asia due to the presence of the Chinese empire.

Most existing studies on conflict relies on Brecke (1999), who provides the arguably most reliable historical data on both major wars and local conflicts since 1400. To complement this data, we digitalize the war data during the 5–17th century recorded in Jaques (2007). Realizing that these sources may cover more information on Europe than China, we employ two ways to increase the comparability. First, since major wars are more likely to be recorded, we focus on wars that lasted three years or longer. Second, we also report the wars in China based on the *Chronicle of Wars in the Chinese History,* part of the Chinese Military History Project (2003). As reported in Table 1, all comparisons suggest that there were more major conflicts and wars, both civil and external ones, in Europe.

Table 1: Number of Major Conflicts and Wars that Lasted 3 Years or Longer

	China	Europe	Source
Major civil conflicts, 1400–1700	21	71	Brecke (1999)
Major external conflicts, 1400–1700	19	172	Brecke (1999)
Major civil wars, 500–1700	3	48	Jaques (2007)
	27	-	Chinese Military History Project (2003)
Major external wars, 500–1700	15	134	Jaques (2007)
	10	-	Chinese Military History Project (2003)

All the three pieces of evidence suggest that China was politically more stable than Europe in the period of interest. To explain this difference, we turn to how power and rights are allocated across the most important players in China and Europe.

2.2 The Absolute Power of the Ruler

Our approach is to rely on rich historical narratives where we can compare the two societies based on various measures. We summarize these narratives in Table 2 and elaborate on them below.

The first difference we emphasize is that Chinese rulers enjoyed more absolute power than their European counterpart. Chinese emperors were above the law. Traditional Chinese political theory held that "all lands under Heaven belong to the Emperor, all people under Heaven are subjects of the Emperor." As put by Fukuyama (2011, p. 290), "in dynastic China, no emperor ever acknowledged the primacy of any legal source of authority; law was only the positive law that he himself made." In contrast, European rulers enjoyed less absolute power, as they faced strong constraints from the Christian church that could be more important in legitimizing the privileged status of the lords (Mann, 1986; Fukuyama, 2011). The Pope had the power both to legitimize a ruler, but also to excommunicate him. When the German Emperor Henry IV decided to take from the Church the power to nominate bishops, he was excommunicated by Pope Gregory VII and made the humiliating trip to Canossa to ask for forgiveness. The threat of excommunication was real throughout European history. Rulers also faced legal constraints. Even in absolutist France under Louis XIV, provinces had their own laws, taxes and customs authority, making the King centralize only partially. When coming to power, French kings were constrained by existing laws made before them. Their power over the judicial system was also limited, as the latter derived its authority from local parliaments.

The difference in the absolute power is also reflected by the ultimate ownership of the most important assets in historical societies: land and population. Below, we discuss some evidence.

While land could be owned by individuals in normal times in China (Zhao and Chen, 2006), the Emperor could re-centralize the ownership when he deemed it necessary.[6] Land confiscation happened repeatedly in Chinese history since the Qin united China in 221 BC. The Qin state (221–206 BC) confiscated land from the feudal nobles and shared it among the peasants. During the Han dynasty that followed the Qin dynasty, Emperor Wu of Han (141–87 BC) confiscated land from nobles to raise additional revenue to fund the Han–Xiongnu War. Similarly, in 843 CE, the Tang government (618–907 CE) ordered that the property of all Manichaean monasteries be confiscated in response to the outbreak of war with the

[6] An oversimplified view of land ownership in China is that peasants only had rights of use but not ownership. However, as pointed out by Zhao and Chen (2006), land transfers in historical China were commonly observed. Thus, we do not make a strong claim about ownership in normal times. Our argument is that the Emperor could claim his ultimate ownership in critical moments including using confiscation as a punishment channel.

Table 2: The Power Structure in Historical China and Europe

	China	Europe	Related Works
Absolute Power of the Ruler			
Legal constraints	Little	Constrained by the Church and by law (e.g., Magna Carta)	Mann (1986), Fukuyama (2011), Acemoğlu and Robinson (2019)
Ownership of land by the Ruler	Ultimate ownership that may not matter in normal times	Very limited	Zhao and Chen (2006)
Confiscation of land	Widespread	Very constrained and exceptional	Ebrey and Walthall (2013)
Direct power over the population	Widespread	Rare	Reynolds (2010), McCollim (2012)
Loyalty requirement to the Ruler	Absolute	Less severe punishment for disloyalty	Bloch (1939), Lander (1961), Tuchman (1978), Mann (1986)
Relative Rights of Elites and People			
Access to elite status	Through the Civil Service Exam	Hereditary nobility	Ho (1959), Levenson (1965)
Land concentration	Less concentrated, <45% owned by landlords since the 6th century	More concentrated, e.g., 70% owned by landlords in 17th-century England	Beckett (1984), Zhang (2017)
Inheritance rule	Partible inheritance	Primogeniture more common	Goldstone (1991)
Tax rates of peasants (after 1500)	Lower	Higher	Rosenthal and Wong (2011), Ma and Rubin (2019)

Uyghurs. With the rise of the Civil Service Exam system in the Song dynasty (960–1279 CE), the gentry class gradually replaced the noble families. Nevertheless, the central government continued to confiscate land owned by the landed gentry in order to raise revenue for multiple military projects (Ebrey and Walthall, 2013).[7]

In contrast, when European Rulers needed revenues, they could usually not confiscate land from the Elites or the Church. Instead, they had to exchange rights with revenues. For instance, in 1189, Richard I of England, intent on raising money for his crusade, had to allow William I of Scotland to annul the Treaty of Falaise and to buy back Scotland's independence. Even under Louis XIV in France, the multiple wars were financed by higher taxes, but not by expropriation of land from the nobles. He imposed for the first time taxes on the nobility, but only at the end of his reign. Moreover, the latter were insignificant in size and were subject to numerous exemptions (McCollim, 2012). Expropriations did happen but mostly under Eminent Domain, more rarely after the execution of an aristocrat for serious crimes (Reynolds, 2010). The creation of the first central bank in England – an institution designed to lend to the government – was initially an expedient by William III of England for the financing of his war against France.

The Chinese Emperor also possessed rights of access to the population. As all people under Heaven were his subjects, he could reward or punish anyone arbitrarily, which precisely reflected his absolute power. One person's crime or disobedience to the Emperor could lead to the eradication of the whole family line. In 1042, Fang Xiaoru refused to write an inaugural address for the Yongle Emperor. In addition to Fang's own execution, his blood relatives and their spouses were killed along with all of his students and peers.

In feudal Europe, the ruler did not have direct access to peasants who were controlled by their landlords. The former could be punished by local courts controlled by landlords, and the ruler did not have control over these local courts (Bloch, 1939). The well-known Magna Carta is regarded as the foundation of the freedom of the individual against the arbitrary authority of the despot. But as recognized by scholars, Magna Carta was intended not to make new laws but to ensure that respect was paid to the good laws of the past which were more ancient and binding (e.g., Vincent, 2012).

Loyalty of the Elites and People to the Chinese Emperor is essential in Confucianism. Loyalty could not be transferred from a deposed regime to its successor. Zhao Meng-fu, the famous artist who served as an official in the Song and the Yuan was condemned by Chinese scholars of his own time and of later dynasties. Coups were not common in Chinese history (which is an equilibrium outcome as we will show), and those involved in plotting coups were

[7] In the same vein, one may say that the Chinese Communist Party's confiscation and redistribution of the landlords' land in the 1950s is not too different from dynastic China, even though its ideology is different.

heavily punished. In 713, Emperor Xuan of Tang, believing that his aunt was planning to overthrow him, acted first, executing a large number of her allies and forcing her to commit suicide.

Although loyalty was also emphasized in Europe and enforced through mechanisms like oaths, an important difference is that treason was punished less harshly. First, despite the notion of "corruption of blood" punishing in different ways the family of a treacherous aristocrat, this very rarely entailed killing the family. Moreover, confiscation of the land of a traitor (known as attainder) was often undone later on (e.g., Lander, 1961).[8]

Formalization in our model. Motivated by these narratives, we distinguish between two scenarios: one is the status quo where the ruled have not defied the Ruler; the other is when they have done so. We conceptualize that greater absolute power of the Ruler means that more of the power and rights of the ruled in the status quo depends on the condition that they have not defied the Ruler, and less would remain if they have been against the Ruler's will. To model this idea, we assume that the Ruler, the Elites and the People share a pie of size π in the status quo; when the Ruler has survived the status quo after the ruled had not sided with him, he can punish the defier by having her enjoy only γ of her "status quo" share of the pie. Compared with the European Ruler, the Chinese emperor is assumed to be capable of enforcing much severer punishment, i.e., a lower γ.

2.3 Power Asymmetry between the Elites and the People

The differences in power structure in China and Europe lied not only in the greater or less absolute power of the Ruler but also in the relationship between the Elites and the People. This second difference is reflected by differences in (1) access to elite status, (2) land concentration, (3) inheritance rule and (4) tax rates of peasants, as summarized in Table 2. Below, we discuss related evidence.

As pointed out by Levenson (1965, p. 39), the infinite power of the Chinese Emperor is his ability to "raise and lower his subjects at will" that renders the relative symmetric Elites–People power structure. Lü Simian, a prominent historian, summarized the Chinese scenario elegantly: "when fathers and elder brothers possess the Empire, younger sons and brothers are low common men" (Lü, 1939, p. 347). As an important institution to facilitate the fluid change between the Elites and the People, China invented the Civil Service Examination to

[8] Moreover, the lords in Europe could often owe allegiance to more than one ruler for different parts of his estates. In the words of Bloch (1939), they were "the man of several masters". In a conflict, the lord chose which superior to follow (e.g., Tuchman, 1978; Mann, 1986). Even though the share of these lords was difficult to tell, it is difficult to find any similar pattern in China.

regulate elite status in the 6th century, and elite status gained via success in the exam could not be inherited.

Different from China, the difference between the Elites and the People was less fluid since elite status was mostly inherited. Statistics on social mobility, despite being scarce, appear consistent with this difference. Ho (1959) provides a comprehensive description of social mobility in China between the 13th and the 19th century based on data from the Civil Service Exam. It is difficult to find comparable data in Europe. He compares China with Cambridge students during 1752–1938. With such an unfair comparison, he still finds a higher social mobility in China: 78–88% of Cambridge students came from elite families whereas only 50–65% of the highest degree holders (Jinshi) came from elite families in China.[9]

Social mobility in Europe, to a large degree, relied on hereditary nobility. Government positions, especially in courts and in the army were reserved for aristocrats. While in the early middle ages, ordinary peasants routinely performed military service, which was seen as a privilege, this stopped to be the case later and was reserved for the nobility (knights and higher titled nobles – see Wickham, 2009 for more discussion).

Statistics on land ownership inequality are also suggestive. In the early Middle Ages, mostly between the eighth and tenth century, 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. 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; Zhao and Chen, 2006).

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, while primogeniture was more common in Europe. The consequence of these rules on elite privilege is intuitive: partible inheritance makes it more difficult for

[9]Bai and Jia (2016) show empirically that the abolition of the exam system as the mobility channel partly contributed to the political instability of China in the early 20th century. One interpretation of this is that the alliance between the ruler and common people was temporarily broken. Recently, Huang and Yang (2020) argue that the exam system contributed to China's imperial longevity.

elite families to accumulate assets across generations. As Goldstone (1991, p. 380) observed, "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."

In addition, scholars have also noticed that tax rates of the peasants were lower in China despite the existence of a strong Chinese state during the Ming and Qing dynasties (e.g., Rosenthal and Wong, 2011; Ma and Rubin, 2019).[10] For us, this serves as another piece of evidence on the difference in relative status – at least official status – between the Elites and the People, since the Elites were usually exempt from taxes in both societies.

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 pie of size π mentioned above, the Elites get a and the People get βa, where $0 \leq \beta \leq 1$ and a higher β indicates a more symmetric Elites–People relationship.

Remarks. To be sure, both China and Europe experienced changes and challenges of the power structure over the centuries. In fact, it should not be surprising that multiple Rulers in Europe attempted to form an alliance with the People against the Elites, as for example Louis XIV's insistence on depriving the nobility of actual power after the rebellions of the Fronde. He attempted to choose ministers and officials on merit, using commoners to replace aristocrats. Nevertheless, the weaker power of the Ruler and the multiple checks on executive power by the Elites in Europe generally make it less possible for the Ruler to consistently succeed in these kinds of endeavors. For instance, even though Louis XIV succeeded temporarily, access to nobility through a judiciary and administrative office became practically barred in the 18th-century France. In a similar vein, the rise of cities in Europe was another phenomenon related to the ruler's hope to enlist cities as allies to centralize power, and the charters that guaranteed certain city rights were usually issued by the king

[10]We would like to be cautious about the definition of taxes. If one only counts taxes to the state, the tax rate was lower in Europe before 1500 (Stasavage, 2020). However, peasants in Europe also paid taxes to the church and their lords. Because the tithe tax to the church alone was similar to the level paid to the Chinese state, the total tax rate in Europe, which is more relevant for our discussion, was likely to be higher even before 1500, and this is not counting the numerous in-kind duties serfs had towards landlords: forced labor and at will military service for the lord plus many other obligations, including fees for using the lord's mills, ovens, etc.

against other local powers.[11] The population that could enjoy cities' privileges was relatively small, however, reflecting the limited power of the Ruler (Cantor, 1964). In Appendix C, We show that our main model can be extended to accommodate this interpretation, where we allow the current political stability to affect future power structure.

2.4 Relevance of Elites and People in Conflicts

Before modeling the implications of the power structure among the Ruler and both the Elites and the People on political stability, it would be reassuring to confirm that both the Elites and the People are relevant in conflicts. To be sure, there existed a wide range of conflicts in both Chinese and European histories. Having carefully examined significant examples,[12] 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.

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 (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

[11] For example, "Louis [VII of France] gave encouragement to the commune movement and received reciprocal support from the communities, at the expense of local lords" (Bradbury, 1998, p. 32); following Louis VII, "Philip [II] 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 [of Philip II] ... 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).

[12] 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.

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 commented on English history: "the 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." 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.[13] In England, shortly before and during the Wars of the Roses, popular support was generally important in determining the fates of Richard II, Henry IV, Edward IV, and Richard III (Grummitt, 2014). In the German Peasants' War, as 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.

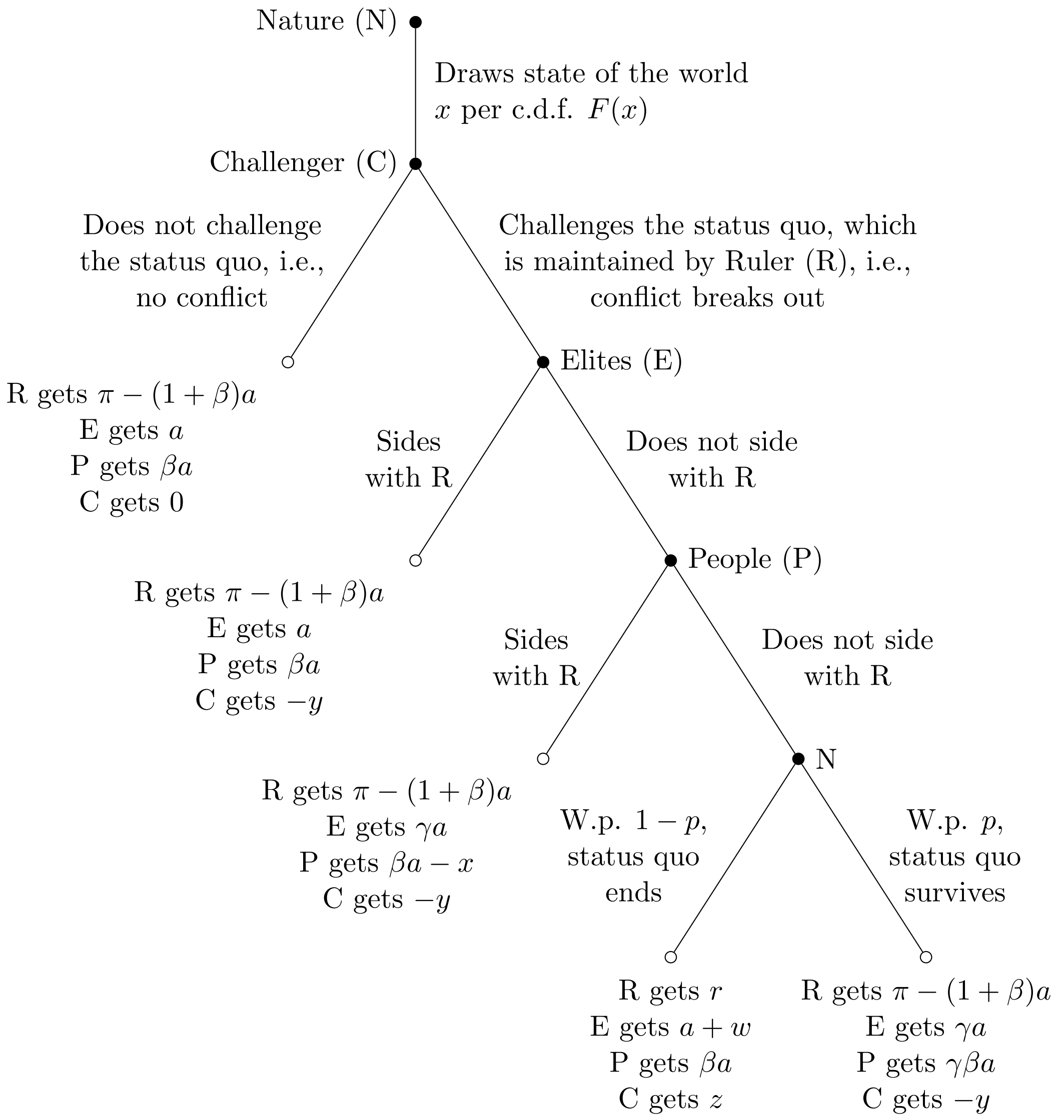
These examples show that both the Elites and the People are highly relevant in conflicts. This gives us confidence to link the power structure among the Ruler and both the Elites and the People to political stability.

3 A General Framework

To accommodate different types of conflicts, we start with a general framework, which is a sequential game with four players: the Ruler (R), the Challenger (C), the Elites (E), which represents the nobles, lords, and bureaucrats, and the People (P), which includes peasants and urban commoners. C could be an outsider, in which case the conflict would be an interstate war; C could be one or a group of elites, in which case E would be the other elites and the conflict would be an elite rebellion or coup; C could also be a group of members of the people, in which case P would be the other members of the people and the conflict

[13]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).

would be an uprising of peasants or urban commoners. Naturally, unanimous actions were rare in reality both within E and within P. Therefore, we interpret their 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.



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

Figure 1: A general framework for comparative institutional analysis

Figure 1 presents the extensive form of the framework. At the beginning of the game, 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 strike, then no conflict will happen, and C will get her default payoff 0; E will get her status quo payoff $a > 0$, which is exogenous; P will get βa, where $\beta \in [0, 1]$ is exogenous and measures the power symmetry between E and P in the status quo; 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. The game then ends there.

If C instead does challenge, then a conflict will happen and E will decide whether to side with R. If E sides with R, then the status quo will survive. The game 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 her 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 $\gamma \in [0, 1]$ is exogenous. A lower γ measures the stronger absolute power of R to punish its subjects who have defied her. The game then ends there.

If P does not side with R either, then R will be left on her 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 her status quo payoff $\pi - (1 + \beta)a$; E will be punished, getting γa; P will be punished, too, getting $\gamma\beta a$. The game 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 reward $w > 0$ for having not sided with R, in addition to her status quo payoff a, so her total payoff will be $a + w$. The game then ends there.

About the random elements of the game, we assume that N's draws of x and whether the status quo will survive on R's own are mutually independent. We assume all the payoffs are von Neumann–Morgenstern utilities so that all players maximize their own expected payoff.

About the informational environment, we assume that the game has 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. Appendix A shows that the insights from our main results would remain robust if mixed strategies were allowed.

Before analyzing the framework, we make a few remarks on the conceptual and technical issues around the framework.

Remarks. First, when mapping our model to history, we interpret that China has a higher β and a lower γ than Europe. The β–γ characterization of 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 the power and rights of the ruled differ between the status quo and the scenario when they have defied R. As we will show, first recognizing the estate- and scenario-specificness and then characterizing power structure this way are instrumental in understanding how power structure determines political stability, since both β and γ shape P, E, and C's strategies in equilibrium, affecting the fate of R and the political status quo. In this sense, β and γ indeed characterize the structure of R's *power* over the others: as Dahl (1957, p. 203) famously puts, "A has power over B to the extent that he can get B to do something that B would not otherwise do."

Second, as mentioned, C can be an outsider or an elite member or part of the people; the reward for E not to side with R also depends on the specific context. Thus, for generality and simplicity, we have left C's identity unspecified and modeled incentives of C and E via exogenous parameters, i.e., w, y, and z. This treatment makes these incentives independent of the β–γ power structure and the strategies of all the players in equilibrium. To address this limitation, in Appendix D, we collapse C and E into one player E from the inside of the status quo, make her look forward infinitely in a Markov game, and allow her to replace R. The β–γ power structure thus determines the punishment upon E in case her challenge fails, and her aspiration for challenge is thus the difference between the equilibrium values of being R and being E, in turn determined by all players' strategies in equilibrium. Therefore, Appendix D can be seen as the fullest yet simplest extension of the current framework. We show parallel results in Appendix D to all results in the current framework. Relatedly, note that R does not make any decision in this game. That said, in the current framework, we will examine R's preference for β and γ; in Appendix D, as we endogenize the incentives of C and E, R's payoff in equilibrium will affect other players' strategies in equilibrium.

Third, having the random variable x is a simple yet useful way to model the cost/reward for P's choice. 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 incentive as a single, exogenously drawn, state-of-the-world variable, i.e., the random cost for siding with R, x. Modeling it alternatively as a reward for not siding with R would not affect our analysis.

Fourth, in the framework, 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's equilibrium strategy by affecting P's equilibrium strategy and they affect C's equilibrium strategy by affecting E and P's equilibrium strategies. Assuming an alternative sequence of moves or simultaneous moves 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. 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, 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 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.

4 Equilibrium Characterization

We start the backward induction from P's strategy. In any subgame perfect equilibrium, P will not side with R if and only if

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

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:

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

Now consider E's strategy while expecting this strategy of P in equilibrium. 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 reward 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 $x \leq \hat{x}$, 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 and no conflict will ever break out. Similar to the equilibria of little relevance we mentioned 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*

$$w > \frac{p}{1-p} \cdot a \quad \textit{and} \quad z > \frac{p}{1-p} \cdot y, \tag{7}$$

then there exists a unique subgame perfect equilibrium, 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

$$\hat{x} \equiv p \cdot (1-\gamma)\beta a. \tag{8}$$

This equilibrium is indeed theoretically non-trivial, since in the equilibrium, whether C will challenge the status quo and start a conflict 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, it is possible for a conflict to break out and for E and P not to side with R, i.e., the probability of conflicts $1 - F(\hat{x})$ can be strictly positive and the survival probability of the status quo

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1 - p) \tag{9}$$

can be strictly lower than one.

5 Analysis of the Equilibrium

We attempt to answer two questions in this section. First, how do the absolute power of the Ruler (γ) and the Elites–People relationship (β) affect political stability, i.e., the probability of conflicts and the survival probability of the status quo? Second, from an institutional design perspective, how would R prefer β and γ, respectively, and could these preferences shed light on the institutional compatibility between a low/high γ and a high/low β?

To focus on the empirically relevant, theoretically nontrivial equilibrium in Proposition 1, from now on we assume that the condition in Proposition 1 holds, i.e., $w > a \cdot p/(1-p)$ and $z > y \cdot p/(1-p)$. Without losing generality, 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 $x \in [0, pa]$.

5.1 Political Stability

We first analyze the impact of the power structure on political stability:

Proposition 2. *In equilibrium, a higher β and a lower γ decrease the probability of conflicts and increase the survival probability of the status quo.*

Proof. By Proposition 1, the probability of conflicts 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 political stability 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: both a higher β and a lower γ impose a heavier punishment $(1 - \gamma)\beta a$ on P for not siding with R in case C's challenge fails, 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.[14] 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., 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.[15]

Third, the impact of β on E's strategy in equilibrium has only the political alliance channel: a higher β imposes a heavier punishment $(1-\gamma)\beta a$ on P for not siding with R in case C's challenge fails, but does not change the punishment $(1-\gamma)a$ on E. Therefore, it would make P more willing to side with R, lowering the chance for C's challenge to succeed, and making E more willing to side with R in the first place.

Finally, the impacts of β and γ on C's strategy in equilibrium has only the political alliance channel, too: a higher β and a lower γ will not affect C's payoffs at any of the five ending nodes in the game, but they will encourage E and L to side with R, therefore lowering C's chance to succeed in her challenge. Expecting this, C will be more reluctant in equilibrium to challenge the status quo.

To summarize, Proposition 2 reveals the first key insight from our model: 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

[14] 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.

[15] 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.

probability of conflicts 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 political stability.[16] Compared with Europe, both a higher β and a lower γ in China first makes P more aligned with R, then E more aligned with R, too, and finally C less likely to initiate a challenge.

5.2 Institutional Compatibility: R's Perspective on γ and β

The equilibrium strategies imply that R's expected payoff is

$$V^R = (\pi - (1+\beta)a) \cdot S + r \cdot (1 - S). \tag{10}$$

R's preference over γ is then 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.

It is also clear that R faces a political–economic trade-off around β:

- a higher β increases the survival probability S of the status quo, which is political;
- a higher β decreases the status quo payoff $\pi - (1+\beta)a$, 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 straightforward, 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 to be 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,

[16] Chapter 17 in *Han Feizi* argues that "too much compulsory labor service" upon 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 "long lasting benefit" of the ruler (low S) (Watson, 1964, p. 87). This argument follows exactly the modeled impact of β on political stability via the political alliance channel in this analysis and Appendix D.

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 R's 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*

$$\underline{c}a \cdot (1-\gamma) \leq \frac{dS}{d\beta} \leq \bar{c}a \cdot (1-\gamma). \tag{11}$$

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

$$\frac{dS}{d\beta} = (1-p) \cdot \frac{dF(\hat{x})}{d\beta} = (1-p)pf(\hat{x}) \cdot a \cdot (1-\gamma), \tag{12}$$

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. □

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

Proposition 3 (R's perspective on β governed by γ, i.e., institutional compatibility)**.** *If* $\gamma < \underline{\gamma} \equiv 1 - 1/(\pi - 2a - r)\underline{c}$, *then R will prefer β to be as high as possible; if* $\gamma > \bar{\gamma} \equiv 1 - p/(\pi - a - r)\bar{c}$, *then R will prefer β to be as low as possible, where* $\underline{\gamma} < \bar{\gamma} < 1$ *and if* $\pi > 2a + r + 1/\underline{c}$, *then* $\underline{\gamma} > 0$.

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

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS. \tag{13}$$

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

$$\begin{aligned} \frac{dV^R}{d\beta} &\geq \big(\pi - (1+\beta)a - r\big) \cdot \underline{c}a \cdot (1-\gamma) - aS \\ &\geq \big((\pi - 2a - r) \cdot \underline{c} \cdot (1-\gamma) - 1\big) \cdot a, \end{aligned} \tag{14}$$

so if

$$(\pi - 2a - r) \cdot \underline{c} \cdot (1-\gamma) - 1 > 0, \tag{15}$$

i.e.,

$$\gamma < 1 - \frac{1}{(\pi - 2a - r) \cdot \underline{c}} \equiv \underline{\gamma}, \tag{16}$$

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

$$\frac{dV^R}{d\beta} \leq \big(\pi - (1+\beta)a - r\big) \cdot \bar{c}a \cdot (1-\gamma) - aS$$
$$\leq \big((\pi - a - r) \cdot \bar{c} \cdot (1-\gamma) - p\big) \cdot a, \tag{17}$$

so if

$$(\pi - a - r) \cdot \bar{c} \cdot (1-\gamma) - p < 0, \tag{18}$$

i.e.,

$$\gamma > 1 - \frac{p}{(\pi - a - r) \cdot \bar{c}} \equiv \bar{\gamma}, \tag{19}$$

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 political stability will be significant; therefore, the political side of R's trade-off around β will always be dominant; R then prefers the highest possible β. If γ is sufficiently high, the opposite will happen, and the economic cost of a higher β will dominate the political gain.

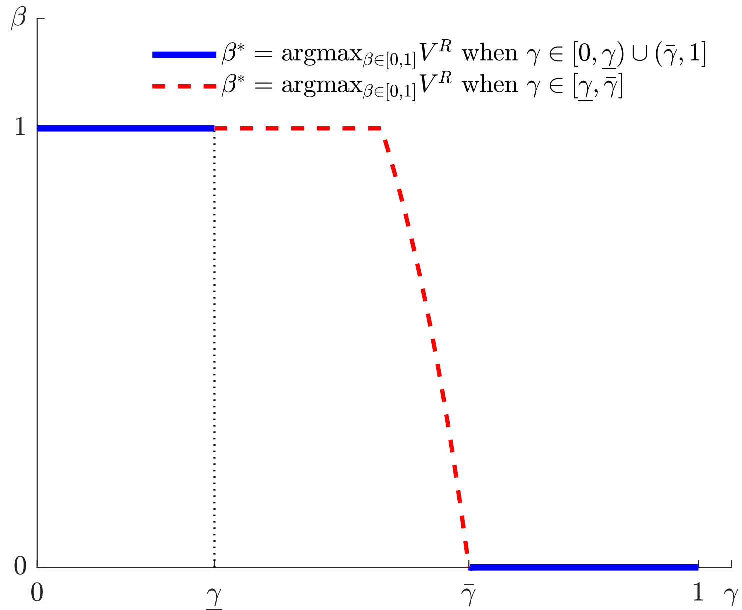
Proposition 3 reveals the second key insight of institutional complementarity derived from our model: as in European history, a high γ and a low β are compatible, while as in Chinese history, a low γ and a high β are compatible.

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

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS \quad \text{and} \quad \frac{dS}{d\beta} = (1-p)pf(\hat{x}) \cdot a \cdot (1-\gamma). \tag{20}$$

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, 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.

That said, we provide a numerical example in Figure 2. We plot R's choice of $\beta^* \equiv \arg\max_{\beta \in [0,1]} V^R$ against γ: consistent with Proposition 3, $\beta^* = 1$ if $\gamma < \underline{\gamma}$, while $\beta^* = 0$ if



Specification: $F(x) = 1 - e^{-x}$, $p = 0.8$, $\pi = 20$, $a = 0.6$, $r = 5$. Under this specification, $\pi - 2a > r$. The blue line plots β^* when $\gamma \in [0, \underline{\gamma}) \cup (\bar{\gamma}, 1]$, which is consistent with Proposition 3. The red line plots β^* when $\gamma \in [\underline{\gamma}, \bar{\gamma}]$, about which Proposition 3 is silent.

Figure 2: R's choice of $\beta \in [0, 1]$ that maximizes V^R in equilibrium as a function of $\gamma \in [0, 1]$

$\gamma > \bar{\gamma}$; silent in Proposition 3, given the specification of the example, β^* weakly decreases with γ over $\gamma \in [\underline{\gamma}, \bar{\gamma}]$.

5.3 Additional Implications

R's perspective on β governed by π. As Equation (13) suggests, a greater total surplus π for R to enjoy will increase the weight of the political side $dS/d\beta$ relative to the economic side in R's trade-off. We can then derive an equivalent result to Proposition 3 but with respect to π:

Corollary 1 (R's perspective on β governed by π)**.** *If $\pi > 2a + r + 1\big/(1-\gamma)\underline{c}$, then R will prefer β to be as high as possible; if $\pi < a + r + p\big/(1-\gamma)\bar{c}$, then R will prefer β to be as low as possible.*

Proof. Following the proof of Proposition 3,

$$(\pi - 2a - r) \cdot \underline{c} \cdot (1 - \gamma) - 1 > 0 \tag{21}$$

is equivalent to

$$\pi > 2a + r + \frac{1}{\underline{c} \cdot (1-\gamma)}; \tag{22}$$

$$(\pi - a - r) \cdot \bar{c} \cdot (1-\gamma) - p < 0 \tag{23}$$

is equivalent to

$$\pi < a + r + \frac{p}{\bar{c} \cdot (1-\gamma)}. \tag{24}$$

The corollary then follows. □

This perspective can also be relevant in the China vs. Europe comparison. A larger size of China (partly thanks to higher political stability) further strengthens the Ruler's political incentives, pushing her to favor a more symmetric Elites–People relationship.

P's perspective on γ given institutional compatibility. Given the institutional compatibility between γ and β, how would P prefer a high or low γ?

Corollary 2. *Taking into R's preference of β into consideration, P prefers any $\gamma < \underline{\gamma}$ over any $\gamma > \bar{\gamma}$.*

Proof. In equilibrium, P's expected payoff is

$$\begin{aligned} V^P &= \gamma\beta a \cdot \big(1 - F(\hat{x})\big) \cdot p + \beta a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot p\Big) \\ &= \Big(1 - \big(1 - F(\hat{x})\big) \cdot p \cdot (1-\gamma)\Big) \cdot \beta a \end{aligned} \tag{25}$$

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

$$V^P\Big|_{\gamma<\underline{\gamma},\beta=1} > 0 = V^P\Big|_{\gamma>\bar{\gamma},\beta=0}. \tag{26}$$

The corollary is then proven. □

The intuition is as follows. On the equilibrium path, 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 $\beta = 0$, so P will receive exactly a zero payoff; any sufficiently low $\gamma < \underline{\gamma}$ will induce R to choose $\beta = 1$, granting P a strictly positive payoff. P will then prefer any sufficiently low $\gamma < \underline{\gamma}$ over the sufficiently high $\gamma > \bar{\gamma}$.

To clarify, we focus on the extreme case to highlight the tradeoff faced by the People. Naturally, in less extreme cases when $\beta > 0$, this comparison is less stark. Our main point is that it is not always the case that the People strongly prefer a high γ. This perspective partly speaks to the debate on the welfare of the people in China and Europe. While lack of rule of law has been characterized as repression of the ruled in the political economy literature, many historians, especially the California school, have documented that pre-industrial living standard of Chinese peasants was not lower than that in Europe (e.g., Pomeranz, 2000; Rosenthal and Wong, 2011; Hoffman, 2017).

Allowing current stability to shape future power. About the institutional compatibility, one may argue that it was not that the European rulers did not want to raise β but was that they were not be able to do so. Using Proposition 3, Appendix C provides a response to this argument by first deriving a result that when the total surplus is extremely big, any R with $\gamma < 1$ will prefer the highest possible β. Second, using this result, we can justify a mechanical, monotonic link from the current S to the future β and γ, finally creating an institutional compatibility in dynamics. By introducing this mechanical dynamics, one can interpret the institutional difference between China and Europe as different stable steady states given the same primitives but different initial strengths of the ruling position in history, which is compatible with different β and γ at very early times.

6 Conclusion

In recent years, economists have made a lot of progress in understanding the importance of rule of law and property rights. While we agree with the existing literature on their importance, we believe other dimensions of institutions are also worth understanding and analyzing, which may help avoid an oversimplified view on institutions and speak to scholars in other disciplines. In this paper, we characterize the power structure in historical China and medieval Europe and relate it to the frequency of major conflicts and (in)stability of the political status quo in the two societies. We acknowledge the difference in the power of the ruler (and thus relatedly rule of law and property rights) and offer a new perspective on the relationship between the elites and people, which, to our knowledge, had not been done before.

We show that both differences contribute to the higher political stability in China than that in Europe via two channels. On top of the intuitive punishment channel (due to the power of the ruler), there exists an important political alliance channel. The stronger absolute power of the ruler and the more symmetric elites–people relationship in China make it not only easier for the ruler to form alliance with the people, but also desirable from the ruler's stability perspective.

Moreover, in our framework, the coexistence of absolute power of the ruler and a more symmetric elites–people relationship is an *equilibrium* outcome, stemming from the political–economic tradeoff faced by the ruler. Generally speaking, political stability concerns dominate the economic concerns in the Chinese scenario and vice versa in the European scenario, exactly due to the ruler's effectiveness or ineffectiveness in raising political stability by leveraging the power structure.

The comparative economic history literature has highlighted the fact that the higher political stability in China, together with the absolute power of the ruler, contributed to the relative lower rate of fundamental innovation (e.g., Mokyr, 2016). Our paper provides a new perspective on what generated higher political stability in China relative to Europe. Although we do not model innovation directly, one can see through the lens of our framework why political concerns dominate economic concerns – including innovation and beyond – for the Chinese ruler.

Admittedly, our model is highly stylized as we capture the power structure with only two parameters, and we only examine political stability as the outcome. The benefit of doing so is that we can deliver our key insights in a simple manner. That said, there can be more insights to gain if one applies our framework of power structure to other parts of the world, e.g., the Muslim world, or if one employs different frameworks to link the power structure with other political, economic and social outcomes (e.g., taxation and social mobility). We thus hope that our study opens new avenues for future research.

References

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

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

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

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.

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

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

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. 1939. *La Société Féodale: La Formation des Liens de Dépendance, Volume 1.* Paris: Éditions Albin Michel.

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.

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

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

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

Dahl, Robert A. 1957. The concept of power. *Behavioral Science* **2**: 201–215.

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

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

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.

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

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.

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.

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

Hoffman, Philip T. 2017. *Why Did Europe Conquer the World?* Princeton: Princeton University Press.

Huang, Yasheng, and Clair Yang. 2020. A longevity mechanism of Chinese absolutism. Working paper.

Jaques, Tony. 2007. *Dictionary of Battles and Sieges*, volume 3. Westport: Greenwood Publishing Group.

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

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.

Lü, Simian. 1939. *A General History of China, Volume 2*. Shanghai: East China Normal University Press. Reprint, 1992.

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.

McNeill, William H. 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.

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

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

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

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

Pomeranz, Kenneth. 2000. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton: Princeton University Press.

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.

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

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990–1990*. Cambridge: Basil Blackwell.

Tuchman, Barbara W. 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.

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

Watson, Burton. 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.

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

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

Zhao, Gang, and Zhongyi Chen. 2006. *History of China's Land Institutions*. Beijing: New Star Press.

Online Appendix

A Allowing for Mixed Strategies

In this section we allow for mixed strategies 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 C to challenge and 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, 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

$$w < \frac{p}{1-p} \cdot (1-\gamma)a; \tag{27}$$

E will not side with R if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a; \tag{28}$$

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

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

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

$$w < \frac{p}{1-p} \cdot (1-\gamma)a; \tag{30}$$

C will also not challenge if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z < \frac{p}{1-p} \cdot y; \tag{31}$$

C will challenge if

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z > \frac{p}{1-p} \cdot y; \tag{32}$$

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

$$w > \frac{p}{1-p} \cdot (1-\gamma)a \text{ and } z = \frac{p}{1-p} \cdot y, \tag{33}$$

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

$$w = \frac{p}{1-p} \cdot (1-\gamma)a, \tag{34}$$

however, C will compare

$$0 \text{ vs. } q_E(x) \cdot (-y) + (1-q_E(x)) \cdot \big((1-p) \cdot z - p \cdot y\big), \tag{35}$$

i.e.,

$$0 \text{ vs. } (1-q_E(x))(1-p) \cdot z - \big(1-(1-q_E(x))(1-p)\big) \cdot y, \tag{36}$$

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

$$z > \frac{1-(1-q_E(x))(1-p)}{(1-q_E(x))(1-p)} \cdot y, \tag{37}$$

$q_C(x) = 0$ if

$$z < \frac{1-(1-q_E(x))(1-p)}{(1-q_E(x))(1-p)} \cdot y, \tag{38}$$

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

$$z = \frac{1-(1-q_E(x))(1-p)}{(1-q_E(x))(1-p)} \cdot y. \tag{39}$$

We have then specified all equilibrium strategies at any $x \neq \hat{x}$. Therefore, 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 conflicts is $\int_{\hat{x}}^{\infty} q_C(x) dF(x)$, while the survival probability of the status quo is

$$S = 1 - \int_{\hat{x}}^{\infty} q_C(x) dF(x) \cdot (1-p). \tag{40}$$

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

$$z \text{ vs. } \frac{1-(1-q_E(x))(1-p)}{(1-q_E(x))(1-p)} \cdot y, \tag{41}$$

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 conflicts is $\int_{\hat{x}}^{\infty} q_C(x) dF(x)$, while the survival probability of the status quo is

$$S = 1 - \int_{\hat{x}}^{\infty} q_C(x)(1-q_E(x))(1-p) dF(x). \tag{42}$$

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

Proposition 4. *If the distribution of x satisfies*

$$\epsilon \equiv -\frac{x \cdot f'(x)}{f(x)} \leq \bar{\epsilon} \equiv 1 - \frac{a}{\pi - 2a - r} \tag{43}$$

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

$$\frac{dV^R}{d\beta} = \big(\pi - (1+\beta)a - r\big) \cdot \frac{dS}{d\beta} - aS, \quad \frac{dS}{d\beta} = (1-p)pf(\hat{x}) \cdot a \cdot (1-\gamma), \tag{44}$$

and

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p). \tag{45}$$

Therefore,

$$\begin{aligned}
\frac{\partial^2 V^R}{\partial\gamma\partial\beta} &= \big(\pi-(1+\beta)a-r\big)\cdot\frac{\partial S}{\partial\gamma\partial\beta} - a\cdot\frac{dS}{d\gamma} \\
&= -\big(\pi-(1+\beta)a-r\big)\cdot(1-p)pa\cdot\big((1-\gamma)f'(\hat{x})\cdot pa\beta + f(\hat{x})\big) - a\cdot\frac{dS}{d\gamma} \\
&= -\big(\pi-(1+\beta)a-r\big)\cdot(1-p)pa\cdot\big(f'(\hat{x})\cdot\hat{x} + f(\hat{x})\big) + a\cdot(1-p)f(\hat{x})p\beta a \\
&= -(1-p)pa\cdot\Big(\big(\pi-(1+\beta)a-r\big)\cdot\big(f'(\hat{x})\cdot\hat{x}+f(\hat{x})\big) - f(\hat{x})\beta a\Big) \\
&= -(1-p)pa\cdot\Big(\big(\pi-(1+\beta)a-r\big)\cdot f'(\hat{x})\cdot\hat{x} + \big(\pi-(1+2\beta)a-r\big)\cdot f(\hat{x})\Big). \qquad (46)
\end{aligned}$$

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

$$\big(\pi-(1+\beta)a-r\big)\cdot f'(\hat{x})\cdot\hat{x} + \big(\pi-(1+2\beta)a-r\big)\cdot f(\hat{x}) \geq 0, \qquad (47)$$

i.e.,

$$\epsilon \equiv -\frac{f'(\hat{x})\cdot\hat{x}}{f(\hat{x})} \leq \frac{\pi-(1+2\beta)a-r}{\pi-(1+\beta)a-r} = 1 - \frac{\beta a}{\pi-(1+\beta)a-r}. \qquad (48)$$

Since

$$\frac{\beta a}{\pi-(1+\beta)a-r} \in \left[0, \frac{a}{\pi-2a-r}\right], \qquad (49)$$

we have

$$1-\frac{\beta a}{\pi-(1+\beta)a-r} \in \left[1-\frac{a}{\pi-2a-r}, 1\right]. \qquad (50)$$

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

$$\epsilon \leq 1-\frac{a}{\pi-2a-r} \equiv \bar{\epsilon}, \quad \text{where} \quad \bar{\epsilon} < 1. \qquad (51)$$

The proposition then follows. □

C Allowing Current Stability to Shape Future Power

Proposition 3 implies:

Corollary 3 (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 game 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}, 1) affect S_t but 2) do so not through γ_t or β_t.

- At $t + 1$: The same happens.

The dynamics then follows

$$\beta_t = \beta(S_{t-1}), \quad \gamma_t = \gamma(S_{t-1}), \quad S_t = S(\beta_t, \gamma_t, \theta), \tag{52}$$

or just

$$S_t = S\left(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\right). \tag{53}$$

Steady states are then defined by

$$\begin{cases} S^* = S(\beta^*, \gamma^*, \theta): & \text{steady-state political stability;} \\ \beta^* = \beta(S^*): & \text{steady-state right symmetry;} \\ \gamma^* = \gamma(S^*): & \text{steady-state rule of law,} \end{cases} \tag{54}$$

or just

$$S^* = S\left(\beta(S^*), \gamma(S^*), \theta\right). \tag{55}$$

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 2. *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 3. *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

$$\frac{dS\left(\beta\left(s\right), \gamma\left(s\right), \theta\right)}{ds} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \geq 0, \tag{56}$$

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

$$0 \leq \frac{dS\left(\beta\left(s\right), \gamma\left(s\right), \theta\right)}{ds} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{57}$$

□

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

Lemma 4. *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

$$-1 < \frac{dS\left(\beta\left(S^*\right), \gamma\left(S^*\right), \theta\right)}{dS^*} = S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{58}$$

Note that

$$S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \geq 0, \tag{59}$$

so

$$0 \leq S_\beta \cdot \beta_S + S_\gamma \cdot \gamma_S \leq 1. \tag{60}$$

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 5. *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. We can say something about institutional compatibility under multiple steady states: a higher β^*, a lower γ^*, and a higher S^* are compatible, while a lower $\beta^{*\prime}$, a higher $\gamma^{*\prime}$, and a lower $S^{*\prime}$ are compatible. 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 6. *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 7. *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 $S_t = S\big(\beta(S_{t-1}), \gamma(S_{t-1}), \theta\big)$ 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 divides 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 of the ruling position in history, which is compatible with different β and γ at very early times.

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

In this extension we collapse C and E in the general framework 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 3 shows each period of the Markov game.

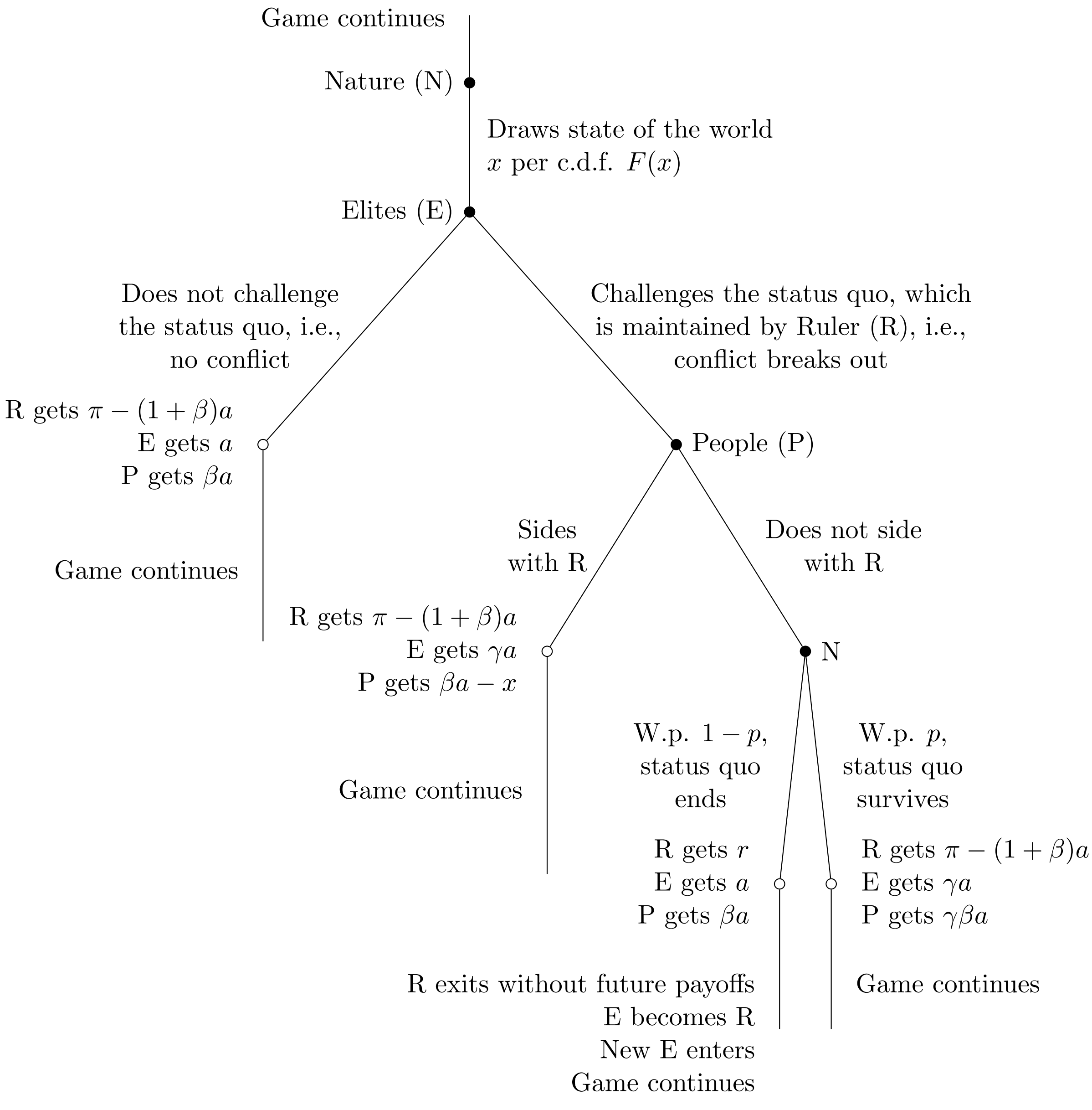


Figure 3: Each period in the Markov game

Compared with Figure 1, the game will now continue after each period; the prize z for

C to challenge and the reward 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 coup on her 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 general framework maintain 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. Appendix D.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.

D.1 Equilibrium Characterization

Now we analyze the model 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 5. *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 inspiration to replace R in equilibrium dominates the probability-adjusted punishment in case of a failed coup:*

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a. \tag{61}$$

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

$$\beta a - x + \delta V^P \geq (\beta a + \delta V^P) \cdot (1-p) + (\gamma \beta a + \delta V^P) \cdot p, \tag{62}$$

i.e.,

$$x \leq (1-\gamma)\beta p \cdot a \equiv \hat{x}. \tag{63}$$

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

$$a + \delta V^E \geq \gamma a + \delta V^E \tag{64}$$

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

$$a + \delta V^E < (a + \delta V^R) \cdot (1 - p) + (\gamma a + \delta V^E) \cdot p, \tag{65}$$

i.e.,

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1 - \gamma)a. \tag{66}$$

The lemma is then proven. □

Note that the analysis is parallel to Section 4, the definition of $\hat{x}$ is the same as in Section 4, and Condition (61) is parallel to Conditions (4) and (6).

By Lemma 5, only two Markov perfect equilibria are possible. The first one is a secured-R equilibrium:

Proposition 8 (Secured-R equilibrium in the Markov game)**.** *If*

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

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

$$V^R - V^E \leq \frac{p}{(1-p)\delta} \cdot (1 - \gamma)a \tag{68}$$

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

$$V^R = \frac{\pi - (1 + \beta)a}{1 - \delta} \quad \text{and} \quad V^E = \frac{a}{1 - \delta}. \tag{69}$$

The condition is then equivalent to

$$\frac{\pi - (1 + \beta)a}{1 - \delta} - \frac{a}{1 - \delta} \leq \frac{p}{(1-p)\delta} \cdot (1 - \gamma)a, \tag{70}$$

i.e.,

$$h(\beta, \gamma) \equiv \frac{\pi - (2 + \beta)a}{1 - \delta} - \frac{p}{(1-p)\delta} \cdot (1 - \gamma)a \leq 0. \tag{71}$$

The proposition is then proven. $\square$

The intuition of the result is as follows: the function $h(\beta, \gamma)$ measures E's inspiration $V^R - V^E = (\pi - (2+\beta)a)/(1-\delta)$ 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 coup fails. The condition $h(\beta, \gamma) \leq 0$ then suggests that the inspiration cannot dominate the punishment. Lemma 5 then implies that we have the secured-R equilibrium.

Note that this equilibrium is parallel to the scenario in Section 4 when Conditions (4) and (6) do not hold. Following the same argument as in Section 4, 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 9 (Unsecured-R equilibrium in the Markov game)**.** *If*

$$g(\beta, \gamma) \equiv \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a > 0, \tag{72}$$

where

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1-p) \in [p, 1] \quad \text{and} \quad \hat{x} \equiv (1-\gamma)\beta p \cdot a, \tag{73}$$

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

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \tag{74}$$

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

$$\begin{aligned} V^R &= \Big(\pi - (1+\beta)a + \delta V^R\Big) \cdot S + r \cdot (1-S) \\ &= \big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S) + \delta V^R \cdot S \\ &= \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} \end{aligned} \tag{75}$$

and

$$
\begin{aligned}
V^E &= a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot p\Big) + \gamma a \cdot \big(1 - F(\hat{x})\big) \cdot p + \delta V^E \cdot S + \delta V^R \cdot (1 - S) \\
&= a \cdot \Big(1 - (1 - \gamma) \cdot \big(1 - F(\hat{x})\big) \cdot p\Big) + \delta V^E \cdot S + \delta V^R \cdot (1 - S) \\
&= \frac{a \cdot \Big(1 - (1 - \gamma) \cdot \big(1 - F(\hat{x})\big) \cdot p\Big) + \delta V^R \cdot (1 - S)}{1 - \delta S}, \qquad (76)
\end{aligned}
$$

with

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot (1 - p) \in [p, 1]. \qquad (77)$$

The condition is then equivalent to, with some algebra,

$$g(\beta, \gamma) \equiv \frac{\big(\pi - (1 + \beta)a\big) \cdot S + r \cdot (1 - S)}{1 - \delta S} - \frac{a}{1 - \delta} - \frac{p}{(1 - p)\delta} \cdot (1 - \gamma)a > 0. \qquad (78)$$

The proposition is then proven. □

Again, the intuition of Proposition 9 follows Lemma 5: the function $g(\beta, \gamma)$ indicates, given the specified strategies, how E's inspiration $V^R - V^E$ to replace R is compared with the punishment in case the coup fails. The condition $g(\beta, \gamma) > 0$ then suggests that the inspiration dominates the punishment. Lemma 5 then implies that we have the unsecured-R equilibrium.

Following the same argument as in Section 4, 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 4. *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 4 is as follows. Since R is safer in the secured-R equilibrium than in the unsecured-R equilibrium, E's inspiration to replace R is stronger, too. Therefore, if E's inspiration 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 inspiration 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 10 (Focus on unsecured-R equilibrium in the Markov game)**.** *If* $\big((1-\delta p)/(1-\delta)(1-p)\delta\big)\cdot a \le r \le \pi - 2a$*, 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 $0 < \Big((1-\delta p)/(1-\delta)(1-p)\delta\Big)\cdot a \le r \le \pi - 2a$ and $S \in [p,1]$, we have

$$\begin{aligned} g(\beta,\gamma) &\ge \frac{(\pi-2a)\cdot S + r\cdot(1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta}\cdot a \\ &\ge \frac{r}{1-\delta S} - \frac{(1-p)\delta + p(1-\delta)}{(1-\delta)(1-p)\delta}\cdot a > \frac{r}{1-\delta p} - \frac{1-p+p}{(1-\delta)(1-p)\delta}\cdot a \\ &\ge \frac{r}{1-\delta p} - \frac{1}{(1-\delta)(1-p)\delta}\cdot a \ge 0. \end{aligned} \tag{79}$$

Therefore, $g(\beta,\gamma) > 0$, i.e., the unsecured-R equilibrium exists, and by Corollary 4, the secured-R equilibrium does not exist. Therefore, the unsecured-R equilibrium is the unique equilibrium. □

In this result, $\Big((1-\delta p)/(1-\delta)(1-p)\delta\Big)\cdot a \le r$ 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.

D.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 10 holds, i.e., $\big((1-\delta p)/(1-\delta)(1-p)\delta\big)\cdot a \le r \le \pi - 2a$, so that the unsecured-R equilibrium exists and is the unique Markov perfect equilibrium. Without losing generality, as in Section 5, 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 Section 5. First note that as in Section 5, the probability of conflicts is still $1 - F(\hat{x})$ and the survival probability of the status quo is still

$$S = 1 - \big(1 - F(\hat{x})\big)\cdot(1-p), \tag{80}$$

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

$$V^R = \frac{\big(\pi - (1+\beta)a\big)\cdot S + r\cdot(1-S)}{1-\delta S} = \frac{\big(\pi - (1+\beta)a - r\big)\cdot S + r}{1-\delta S}, \tag{81}$$

which differs from Equation 10 only at it including 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 11 (Institutional compatibility in the Markov game)**.** *If* $\gamma < \underline{\gamma} \equiv 1 - (1-\delta)/\left((1-\delta(1-p))(\pi - 2a - r) + \delta r\right)\underline{c}$*, then R will prefer* β *to be as high as possible; if* $\gamma > \bar{\gamma} \equiv 1 - (1-\delta)p/\left(\pi - a - r(1-\delta)\right)\bar{c}$*, then R will prefer* β *to be as low as possible, where* $\underline{\gamma} < \bar{\gamma} < 1$ *and if* $\pi > 2a + \left(\frac{1-\delta}{\underline{c}} + (1-\delta(2-p))r\right) / \left(1-\delta(1-p)\right)$*, then* $\underline{\gamma} > 0$*.*

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

$$\frac{dV^R}{d\beta} = \frac{\left(\pi - (1+\beta)a - r + \frac{\delta\left(\left(\pi-(1+\beta)a-r\right)S+r\right)}{1-\delta S}\right)\cdot\frac{dS}{d\beta} - aS}{1-\delta S}. \tag{82}$$

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

$$\begin{aligned}\frac{dV^R}{d\beta} &\ge \frac{\left(\frac{(1-\delta(S-p))(\pi-2a-r)+\delta r}{1-\delta S}\right)\cdot\underline{c}a\cdot(1-\gamma) - a}{1-\delta S} \\ &\ge \frac{a}{1-\delta S}\cdot\left(\frac{(1-\delta(1-p))(\pi-2a-r)+\delta r}{1-\delta}\cdot\underline{c}\cdot(1-\gamma) - 1\right),\end{aligned} \tag{83}$$

so, if

$$\frac{(1-\delta(1-p))(\pi-2a-r)+\delta r}{1-\delta}\cdot\underline{c}\cdot(1-\gamma) - 1 > 0, \tag{84}$$

i.e.,

$$\gamma < 1 - \frac{1-\delta}{\left((1-\delta(1-p))(\pi-2a-r)+\delta r\right)\cdot\underline{c}} \equiv \underline{\gamma}, \tag{85}$$

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

$$\begin{aligned}\frac{dV^R}{d\beta} &\le \frac{\left(\pi - a - r + \frac{\delta(\pi-a)}{1-\delta}\right)\cdot\bar{c}a\cdot(1-\gamma) - ap}{1-\delta S} \\ &= \frac{a}{1-\delta S}\cdot\left(\left(\frac{\pi-a}{1-\delta} - r\right)\cdot\bar{c}\cdot(1-\gamma) - p\right),\end{aligned} \tag{86}$$

so, if

$$\left(\frac{\pi - a}{1-\delta} - r\right) \cdot \bar{c} \cdot (1-\gamma) - p < 0, \tag{87}$$

i.e.,

$$\gamma > 1 - \frac{(1-\delta)p}{\big(\pi - a - r(1-\delta)\big) \cdot \bar{c}} \equiv \bar{\gamma}, \tag{88}$$

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

Proposition 11 differs from Proposition 3 only at $\underline{\gamma}$ and $\bar{\gamma}$ are differently defined, respectively, due to the change in the expression of V^R. This is then followed by parallel results to Section 5.3. We have then shown that we can derive all the parallel results to the general framework from the Markov game.

D.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 and examine whether the main insights would remain. As in 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 to 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

$$\bar{q}_E \equiv \frac{\int_{\hat{x}}^{\infty} q_E(x) dF(x)}{1 - F(\hat{x})} \in [0,1]. \tag{89}$$

In particular, if in equilibrium

$$V^R - V^E > \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{90}$$

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

$$V^R - V^E < \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{91}$$

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

$$V^R - V^E = \frac{p}{(1-p)\delta} \cdot (1-\gamma)a, \tag{92}$$

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

$$V^R = \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1 - \delta S}, \tag{93}$$

$$\begin{aligned} V^E &= a \cdot \Big(1 - \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p\Big) + \gamma a \cdot \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p + \delta V^E \cdot S + \delta V^R \cdot (1-S) \\ &= \frac{a \cdot \Big(1 - (1-\gamma) \cdot \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot p\Big) + \delta V^R \cdot (1-S)}{1 - \delta S}, \end{aligned} \tag{94}$$

and

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot (1-p). \tag{95}$$

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

$$\begin{aligned} &V^R - V^E - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \\ &= \frac{1-\delta}{1-\delta S} \cdot \left(\frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1-S)}{1-\delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a \right). \end{aligned} \tag{96}$$

Now define

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

where

$$S = 1 - \big(1 - F(\hat{x})\big) \cdot \bar{q}_E \cdot (1-p) \quad \text{and} \quad \hat{x} = (1-\gamma)\beta pa. \tag{98}$$

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

$$k(\beta, \gamma, \bar{q}_E) = 0, \tag{99}$$

   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

$$\frac{\int_{\hat{x}}^{\infty} q_E(x) dF(x)}{1 - F(\hat{x})} = \bar{q}_E, \tag{100}$$

     when $x > \hat{x}$ and will not challenge when $x < \hat{x}$.

Note that Scenario 1 corresponds to Proposition 8, where $h(\beta, \gamma) \equiv k(\beta, \gamma, 0)$, and Scenario 2 corresponds to Proposition 9, where $g(\beta, \gamma) \equiv k(\beta, \gamma, 1)$. Now examine whether our main messages maintain in Scenario 3.

In Scenario 3, in equilibrium, we always have

$$k(\beta, \gamma, \bar{q}_E) \equiv \frac{\big(\pi - (1+\beta)a\big) \cdot S + r \cdot (1 - S)}{1 - \delta S} - \frac{a}{1-\delta} - \frac{p}{(1-p)\delta} \cdot (1-\gamma)a = 0, \tag{101}$$

i.e.,

$$\big(\pi - (1+\beta)a - r\big) \cdot S + r = \left(\frac{1}{1-\delta} + \frac{p \cdot (1-\gamma)}{(1-p)\delta}\right) \cdot a \cdot (1 - \delta S). \tag{102}$$

This implies

$$dS = \frac{\frac{pa(1-\delta S)}{(1-p)\delta} \cdot d(1-\gamma) + aS \cdot d\beta}{\pi - (1+\beta)a - r + \left(\frac{1}{1-\delta} + \frac{p \cdot (1-\gamma)}{(1-p)\delta}\right) \delta a}. \tag{103}$$

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

$$\frac{dS}{d\beta} = \frac{aS}{\pi - (1+\beta)a - r + \left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\delta a}$$
$$= \frac{a \cdot \left(\left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right) \cdot a - r\right)}{\left(\pi - (1+\beta)a - r + \left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right)\delta a\right)^2}. \tag{104}$$

This implies

$$\frac{dS}{d\beta} \leq \frac{a \cdot \left(\left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right) \cdot a - r\right)}{\left(\pi - 2a - r + \frac{1}{1-\delta} \cdot \delta a\right)^2} \equiv \bar{b}(\gamma) \tag{105}$$

and

$$\frac{dS}{d\beta} \geq \frac{a \cdot \left(\left(\frac{1}{1-\delta} + \frac{p\cdot(1-\gamma)}{(1-p)\delta}\right) \cdot a - r\right)}{\left(\pi - a - r + \left(\frac{1}{1-\delta} + \frac{p}{(1-p)\delta}\right)\delta a\right)^2} \equiv \underline{b}(\gamma), \tag{106}$$

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 11 would then follow.

To summarize, allowing for mixed strategies would allow the mixed-strategy, Scenario-3 equilibria, in which the main insights from Scenario 2 would maintain, but with more technical complicity. In light of this, we can rule out mixed strategies from the Markov game, gaining in simplicity without losing much intuition.